catalysts

Novel Enzyme and Whole-Cell Biocatalysts

Edited by
Anwar Sunna and Richard Daniellou
Printed Edition of the Special Issue Published in *Catalysts*

MDPI

Novel Enzyme and Whole-Cell Biocatalysts

Novel Enzyme and Whole-Cell Biocatalysts

Editors

Anwar Sunna
Richard Daniellou

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

MDPI

Editors
Anwar Sunna
Macquarie University
Australia

Richard Daniellou
Université d'Orléans UMR-CNRS 7311
France

This is a reprint of articles from the Special Issue published online in the open access journal *Catalysts* (ISSN 2073-4344) (available at: https://www.mdpi.com/journal/catalysts/special_issues/enzyme_whole_cell_biocataly).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Article Number*, Page Range.

Contents

About the Editors

Anwar Sunna received his PhD in Technical Microbiology from the Hamburg University of Technology (TUHH) in Germany and has held research and teaching positions at the TUHH and Potsdam University. He joined Macquarie University (MQ) in 2005 at the Department of Molecular Sciences. He was manager of the Environmental Biotechnology Cooperative Research Centre at MQ and later was the recipient of the prestigious MQ Vice-Chancellor's Innovation Fellowship. His research is strongly driven by industrial and biomedical applications using Synthetic Biology and NanoBiotechnology to address current biotechnological and biomedical challenges. His current research integrates enzyme technology and biomolecule–inorganic interactions with synthetic biology for the construction of multi-enzyme biocatalytic modules and the assembly of synthetic pathways. His research is aimed at addressing current relevant issues including valorization of organic waste, and biomanufacturing bulk and specialty chemicals. He is currently the Director of the MQ Biomolecular Discovery Research Centre, one of the Directors of Synthetic Biology Australasia, and the Director of the Master of Biotechnology and Master of Biotechnology and Business programs at Macquarie University.

Richard Daniellou, An internationally recognized expert in Glycosciences, Prof. Richard Daniellou received a BSc in Biochemistry and a PhD (2003) in Organic Chemistry from Paris XI. After two years as a postdoctoral researcher at the University of Saskatchewan (Canada), he was offered an Assistant Professor position at the ENSC of Rennes (France). In 2010, he was promoted Full Professor of Biochemistry at ICOA (France). His main interest for carbohydrate-active enzymes as biocatalysts for chemo-enzymatic synthesis of glycoconjugates led him to the creation of the research group named Enzymology and Glycobiochemistry. He is currently the co-author of 75 publications and 3 patents.

Editorial Catalysts: Special Issue on Novel Enzyme and Whole-Cell Biocatalysts

Anwar Sunna [1],* and Richard Daniellou [2],*

1 Department of Molecular Sciences, Macquarie University, Sydney, NSW 2109, Australia
2 ICOA UMR CNRS 7311, University of Orleans, rue de Chartres, BP 6759, CEDEX 2, 45067 Orléans, France
* Correspondence: anwar.sunna@mq.edu.au (A.S.); richard.daniellou@univ-orleans.fr (R.D.)

Global trends emphasizing the reduction of organic waste, carbon capture and landfill avoidance are driving the demand for greener products with improved properties. Recent advances in synthetic biology, molecular biology, computational tools and metabolic engineering have promoted the discovery of new enzymes and the rational design of whole-cell biocatalysts. Accordingly, with increased demand for sustainable and environmentally-friendly biomanufacturing, the field of enzyme technology and biocatalysis (multi-enzymes and whole-cells) has become a primary focus for the synthesis of bio-based chemicals and high-value compounds.

In this Special Issue, we would like to highlight these current advances in the field of biocatalysis, with special emphasis on novel enzymes and whole-cell biocatalysts for applications in industry, health, or cosmetics.

Over the past decades, biodiesel has attracted great interest as a sustainable alternative for fossil fuels. Two research papers focused on this important challenge. The enzymatic production of biodiesel from waste cooking oil that could contribute to resolve the problems of energy demand and environment pollution. Yang et al. [1] report on the activation of *Burkholderia cepacia* lipase by surface imprinting and its immobilization in magnetic cross-linked enzyme aggregates, thus exhibiting a significant increase in biodiesel yield and tolerance to methanol. Shafiq et al. [2] describe the use of response surface methodology to optimize the reaction parameters of bioproduction of biodiesel from waste chicken fat oil and demonstrated that optimal yield can be obtained in their conditions using immobilized *Aspergillus terreus* lipase on Fe_3O_4 nanoparticles with a methanol-to-oil ratio of 6:1 at 42 °C for 36 h.

However, bioproduction is certainly not limited to solving energy and environmental troubles and there is also a challenging area in the field of biomass valorization, as depicted by an excellent review from Martinez et al. [3]. The authors outlined the challenges and opportunities in the discovery of original keratinases for value-added peptide production. Indeed, keratins represent millions of tons of protein wastes and their enzymatic hydrolysates can generate valuable industrial applications. To do so, the search for original, innovative and robust biocatalysts is a key step, and extremophile sources are a good starting point. Therefore, a large part of this special issue has been devoted to this immense field of research. First, Kopp et al. [4] identify and characterize the first archaeal mannonate dehydratase from *Thermoplasma acidophilum* and demonstrate it to have an original physiological role in this extremophile. Then, Burkhardt et al. [5] show interest in two endo-β-1,3-glucanases from the thermophilic bacterium *Fervidobacterium* sp. These two enzymes proved to be highly specific to laminarin and tolerant to high temperature, and are good candidates for application in biomass conversion. In addition, Tahir et al. [6] clone and overexpress a novel hormone-sensitive lipase-like

esterase from *Glaciozyma antartica*. Unlike other known enzymes, this protein demonstrates higher activity towards medium-chain ester substrates, rather than shorter chain esters, and increased stability at 60 °C, as well as alkaline pH conditions. In addition, asymmetric catalysis is evoked by the article of Aregger et al. [7], in which the authors report the characterization of the novel ene reductase Ppo-Er1 from *Paenibacillus polymyxa*. This biocatalyst exhibits enantioselective activity towards a large panel of substrates, a large range of temperature and co-solvents, making it a promising tool for industrial bioconversions. Finally, an excellent review from Shah et al. [8] presents the cofactor F420 both (i) as an alternative to nicotinamide cofactors implicated in asymmetric reduction of enoates, imines and ketones, and (ii) as an underexplored resource for asymmetric redox biocatalysis at the industrial level.

Besides revealing their original activities, many processes also can be improved so to increase industrial applications of enzymes. One elegant and powerful example relies in the machine learning approach described by Nagaraja et al. [9] for the efficient selection of enzyme concentration and its application for flux optimization. Pham et al. [10] demonstrate the immobilization of β-galactosidase on the *Lactobacillus* cell surface using LysM, the common peptidoglycan-binding motif, thus facilitating many uses of the biocatalyst and showing its potential for applications in the synthesis of prebiotic galacto-oligosaccharides.

Understanding the intrinsic mechanism of the enzyme also can help in improvement of the biocatalyst. Bracco et al. [11] demonstrated one decisive criterion to differentiate between esterase and lipase, with the latter being the only one active in dry organic solvents. Substrates also can give important clues on the mechanism, as demonstrated by Guillotin and co-workers [12] when using glycosyl thioimidates with biologically-relevant glycoside hydrolases. Fine tuning of amino acids is another level of improvement. Liu et al. [13] show the role of V88L substitution in increasing the enzyme activity and decreasing the protein stability in the New Dehli metallo-β-lactamase-1 family. Mutagenesis of enzymes is a further powerful tool as shown by Schwardmann et al. [14]. The authors studied the well-defined outer ring of the substrate groove of a non-specific nuclease from *Pseudomonas syringae* and defined it as a potential target for modulation of the enzymatic performance. Perugino et al. [15] demonstrated that random mutagenesis and biological selection allowed the identification of residues that are critical in determining thermal activity, stability and substrate recognition of a β-glycosidase from the thermoacidophile *Saccharolobus solfataricus*. Cyclodextrin transferase's product specificity was changed finally by Sonnendecker et al. [16] by semi-rational mutagenesis, obtaining larger cyclodextrin's rings of up to 12 units.

Whole-cell biocatalysed reactions are also discussed broadly. First, the impact of low-level organic solvents on engineered *Escherichia coli* strains was studied in a model reaction of multi-enzyme whole-cell biocatalysts by Yang and co-workers [17]. Secondly, Milessi-Esteves et al. [18] described the production of ethanol from xylo-oligomers by native *Saccharomyces cerevisiae*. The authors investigate a new concept of biocatalysts to overcome the ease of the contamination of the bioreactor by bacteria that metabolize xylose. In addition, Cang et al. [19] showed that the extremely radiation resistant *Deinococcus wulumquiensis* R12 was a new and robust biocatalyst for selective oxidation of 5-hydroxymethylfurfural to 5-hydroxymethyl-2-furancaroxylic acid. In the final contribution, Noshahri et al. [20] examined the Iranian soil to locate robust microorganisms with ω-transaminase activities.

In conclusion, this Special Issue showcases a large panel of techniques and tools in biocatalysis, so as to transform biomass into valuable energy and other bioproducts.

References

1. Yang, H.; Zhang, W. Surfactant imprinting hyperactivated immobilized lipase as efficient biocatalyst for biodiesel production from waste cooking oil. *Catalysts* **2019**, *9*, 914. [CrossRef]

2. Shafiq, F.; Mumtaz, M.W.; Mukhtar, H.; Touqeer, T.; Raza, S.A.; Rashid, U.; Nehdi, I.A.; Choong, T.S.Y. Response surface methodology approach for optimized biodiesel production from waste chicken fat oil. *Catalysts* **2020**, *10*, 633. [CrossRef]

3. De Oliveira Martinez, J.P.; Cai, G.; Nachtschatt, M.; Navone, L.; Zhang, Z.; Robins, K.; Speight, R. Challenges and opportunities in identifying and characterising keratinases for value-added peptide production. *Catalysts* **2020**, *10*, 184. [CrossRef]

4. Kopp, D.; Willows, R.; Sunna, A. Characterisation of the first archaeal mannonate dehydratase from *Thermoplasma acidophilum* and its potential role in the catabolism of D-Mannose. *Catalysts* **2019**, *9*, 234. [CrossRef]

5. Burkhardt, C.; Schäfers, C.; Claren, J.; Schirrmacher, G.; Antranikian, G. Comparative analysis and biochemical characterization of two endo-β-1,3-glucanases from the thermophilic bacterium *Fervidobacterium* sp. *Catalysts* **2019**, *9*, 830. [CrossRef]

6. Tahir, H.M.; Raja Abd Rahman, R.N.Z.; Leow, A.T.C.; Ali, M.S.M. Expression, characterisation and homology modelling of a novel hormone-sensitive lipase (HSL)-like esterase from *Glaciozyma antarctica*. *Catalysts* **2020**, *10*, 58. [CrossRef]

7. Aregger, D.; Peters, C.; Buller, R.M. Characterization of the novel ene reductase Ppo-Er1 from *Paenibacillus polymyxa*. *Catalysts* **2020**, *10*, 254. [CrossRef]

8. Shah, M.V.; Antoney, J.; Kang, S.W.; Warden, A.C.; Hartley, C.J.; Nazem-Bokaee, H.; Jackson, C.J.; Scott, C. Cofactor F_{420}-dependent enzymes: An under-explored resource for asymmetric redox biocatalysis. *Catalysts* **2019**, *9*, 868. [CrossRef]

9. Nagaraja, A.A.; Charton, P.; Cadet, X.F.; Fontaine, N.; Delsaut, M.; Wiltschi, B.; Voit, A.; Offmann, B.; Damour, C.; Grondin-Perez, B.; et al. A machine learning approach for efficient selection of enzyme concentrations and its application for flux optimization. *Catalysts* **2020**, *10*, 291. [CrossRef]

10. Pham, M.L.; Tran, A.M.; Kittibunchakul, S.; Nguyen, T.T.; Mathiesen, G.; Nguyen, T.H. Immobilization of β-galactosidases on the *Lactobacillus* cell surface using the peptidoglycan-binding motif LysM. *Catalysts* **2019**, *9*, 443. [CrossRef] [PubMed]

11. Bracco, P.; van Midden, N.; Arango, E.; Torrelo, G.; Ferrario, V.; Gardossi, L.; Hanefeld, U. *Bacillus subtilis* lipase A—Lipase or esterase? *Catalysts* **2020**, *10*, 308. [CrossRef]

12. Guillotin, L.; Assaf, Z.; Pistorio, S.G.; Lafite, P.; Demchenko, A.V.; Daniellou, R. Hydrolysis of glycosyl thioimidates by glycoside hydrolase requires remote activation for efficient activity. *Catalysts* **2019**, *9*, 826. [CrossRef]

13. Liu, Z.; Piccirilli, A.; Liu, D.; Li, W.; Wang, Y.; Shen, J. Deciphering the role of V88L substitution in NDM-24 metallo-β-lactamase. *Catalysts* **2019**, *9*, 744. [CrossRef]

14. Schwardmann, L.S.; Schmitz, S.; Nölle, V.; Elleuche, S. Decoding essential amino acid residues in the substrate groove of a non-specific nuclease from *Pseudomonas syringae*. *Catalysts* **2019**, *9*, 941. [CrossRef]

15. Perugino, G.; Strazzulli, A.; Mazzone, M.; Rossi, M.; Moracci, M. Effects of random mutagenesis and in vivo selection on the specificity and stability of a thermozyme. *Catalysts* **2019**, *9*, 440. [CrossRef]

16. Sonnendecker, C.; Zimmermann, W. Change of the product specificity of a cyclodextrin glucanotransferase by semi-rational mutagenesis to synthesize large-ring cyclodextrins. *Catalysts* **2019**, *9*, 242. [CrossRef]

17. Yang, Y.; Liu, M.Z.; Cao, Y.S.; Li, C.K.; Wang, W. Low-level organic solvents improve multienzyme whole-cell catalytic synthesis of myricetin-7-*O*-glucuronide. *Catalysts* **2019**, *9*, 970. [CrossRef]

18. Milessi-Esteves, T.S.; Corradini, F.A.S.; Kopp, W.; Zangirolami, T.C.; Tardioli, P.W.; Giordano, R.C.; Giordano, R.L.C. An innovative biocatalyst for continuous 2G ethanol production from xylo-oligomers by *Saccharomyces cerevisiae* through simultaneous hydrolysis, isomerization, and fermentation (SHIF). *Catalysts* **2019**, *9*, 225. [CrossRef]

19. Cang, R.; Shen, L.Q.; Yang, G.; Zhang, Z.D.; Huang, H.; Zhang, Z.G. Highly selective oxidation of 5-hydroxymethylfurfural to 5-hydroxymethyl-2-furancarboxylic acid by a robust whole-cell biocatalyst. *Catalysts* **2019**, *9*, 526. [CrossRef]
20. Noshahri, N.G.; Fooladi, J.; Syldatk, C.; Engel, U.; Heravi, M.M.; Mehrjerdi, M.Z.; Rudat, J. Screening and comparative characterization of microorganisms from iranian soil samples showing ω-transaminase activity toward a plethora of substrates. *Catalysts* **2019**, *9*, 874. [CrossRef]

Article

Surfactant Imprinting Hyperactivated Immobilized Lipase as Efficient Biocatalyst for Biodiesel Production from Waste Cooking Oil

Huixia Yang and Weiwei Zhang *

State Key Laboratory of High-efficiency Utilization of Coal and Green Chemical Engineering, School of Chemistry and Chemical Engineering, Ningxia University, Yinchuan 750021, China; yhx6668297@sina.com
* Correspondence: zhangww@nxu.edu.cn; Tel.: +86-0951-2062004

Abstract: Enzymatic production of biodiesel from waste cooking oil (WCO) could contribute to resolving the problems of energy demand and environment pollutions. In the present work, *Burkholderia cepacia* lipase (BCL) was activated by surfactant imprinting, and subsequently immobilized in magnetic cross-linked enzyme aggregates (mCLEAs) with hydroxyapatite coated magnetic nanoparticles (HAP-coated MNPs). The maximum hyperactivation of BCL mCLEAs was observed in the pretreatment of BCL with 0.1 mM Triton X-100. The optimized Triton-activated BCL mCLEAs was used as a highly active and robust biocatalyst for biodiesel production from WCO, exhibiting significant increase in biodiesel yield and tolerance to methanol. The results indicated that surfactant imprinting integrating mCLEAs could fix BCL in their active (open) form, experiencing a boost in activity and allowing biodiesel production performed in solvent without further addition of water. A maximal biodiesel yield of 98% was achieved under optimized conditions with molar ratio of methanol-to-WCO 7:1 in one-time addition in hexane at 40 °C. Therefore, the present study displays a versatile method for lipase immobilization and shows great practical latency in renewable biodiesel production.

1. Introduction

Over the past decades, biodiesel has attracted great interest as a sustainable alternative for fossil fuels in virtue of the depletion of fossilized fuel resources and their environmental impacts [1]. Biodiesel is a renewable and clean energy, and possess favorable advantages in combustion emission like low emissions of CO, sulfur free, low hydrocarbon aroma, high cetanenumber, and high flash point [2].

The conventional chemical technologies for biodiesel production involve the use of acid or basic catalysts (i.e., NaOH, KOH, and H_2SO_4), thus numerous disadvantages are inescapable, for example, high corrosive procedure, high energy consumption, high quantities of waste pollution, and costly in efficient product separation processes [3]. Furthermore, in order to prevent the hydrolysis reaction and saponification, high quality oils are required, with low contents of water and free fatty acids [4].

Feedstocks used for biodiesel can be allocated five categories, including edible vegetable oils, non-edible plant oils, animal fats, microalgae oils, and waste oils [5]. The global application of first-generation biodiesel produced by using edible oils, was restricted due to food scarcity and high cost of the edible oils [6]. Biodiesel production from waste cooking oils (WCO) could be a promising and cost effective candidate in handling issues associated with energy crisis, environmental concerns, and total cost reduction of biodiesel production [7]. Moreover, 15 million tons of WCO are produced annually throughout the world [8], bringing great challenge in reasonable management

of such oils on account of environment concerns [9]. However, using WCO as raw material is quite challenging as it contains a high amount of free fatty acids (FFAs) and water which could hinder the homogeneous alkaline-catalyzed transesterification in conventional biodiesel production processes [10]. Complete conversion of these low-quality feedstocks like WCO could be accomplished in enzymatic biodiesel production without saponification. Therefore, enzyme-catalyzed transesterification has become a laudable potential alternative for biodiesel synthesis.

Particularly, lipases are foremost and efficient enzymes implemented in biodiesel production. Lipase-catalyzed process exhibits key advantages such as no soap formation, high-purity products, easy product removal, adaptable to different biodiesel feedstock, environmentally friendly, and mild operating conditions [5]. However, the commercialization of enzymatic biodiesel production remains complicated, because of high price and low stability of lipases as well as low reaction rate of biocatalysis. Heterogeneous enzyme-catalyzed transesterification using immobilized lipases is a possible solution to these problems [11].

Immobilization of enzymes has been investigated for many years, and lipase can generally be immobilized by various techniques such as cross-linking, adsorption, entrapment and encapsulation [12,13]. Thereinto, cross-linked enzyme aggregates (CLEAs) is a cheap and efficient strategy for enzyme immobilization, which has broad applicability over numerous enzyme classes. Owing to its outstanding resistance to organic solvents, extreme pH, and high temperatures, CLEAs has attracted growing attention in cost effective biocatalysis [14]. Nevertheless, small particle size and low mechanical stability of CLEAs could directly affect mass transfer and stability under operational conditions, thus accordingly cause problems in practical use [15]. An alternative approach for circumventing compressed construction of CLEAs is to use "smart" magnetic CLEAs (mCLEAs). Magnetic nanoparticles (MNPs) could provide enhanced stability over repeated uses, especially for enzymes having low amount of lysine residues on their surface. Besides, mCLEAs could perform easily separation using a permanent magnet, affording novel combinations of bioconversions and down-streaming processes, thus provide the necessary reduction in enzymecosts to enable commercial viability.

Among various types of nanomaterials, MNPs have attracted substantial attention in enzyme immobilizations. However, bare MNPs tend to aggregate due to their high surface energy and are easily oxidized in the air leading to loss of magnetism and dispersibility, thus limiting their exploitation in practical applications [16]. The surface modification with an organic or an inorganic shell is an appropriate strategy to address these issues. Due to their excellent biocompatibility, slow biodegradation, high surface area-to-volume-ratio, and unique mechanical stability, Hydroxyapatite (HAP) could be a proper inorganic surface coating for MNPs [17]. Moreover, HAP-coated MNPs can be easily functionalized with organosilanes, and consequently has great application potential in enzyme immobilization.

Burkholderia cepacia lipase (BCL)is one of the most widely used lipases in biocatalysis [18]. On account of its versatility to accommodate a wide variety of substrates, high heat resistance, and good tolerance to polar organic solvents, BCL has been extensively used in various biotechnological processes, especially for biodiesel production. The active site of BCL is shielded by a mobile element, called the lid or flap [19]. The displacement of lid or flap to closed or open position, which directly impacts the accessibility of active site, determines the enzyme in an in active or active conformation. In general, substrate access to the underlying active site is prohibited in its closed configuration. As the stabilization of the open conformation of all lipases could remarkably increase their catalytic activity, a favorable method to obtain highly active biocatalysts should try to immobilize lipases in their most active form (open conformation).

Generally, the preparation of immobilized enzyme with enhanced activity and stability is a persistent goal of the biotechnology industry to seek maximum profit. Therefore, developing a simple and efficient approach for lipase interfacial activation in immobilization is highly desirable. Bioimprinting is a commonly used method for achieving hyperactivation of lipases in organic media.

The principle of bioimprinting is to "anchor" the enzyme in its active form, which could be achieved by binding with imprint molecules (such as surfactants, natural substrates, substrate analogs etc.). From an applied point of view, the dramatic hyperactivation of lipases by low concentrations of surfactants is an expeditious and facile method for lipase interfacial activation [20].

To develop an efficient and environmentally benign process for the biodiesel production from waste cooking oils, in the present study surfactant imprinting strategy on BCL was implemented in combination with mCLEAs immobilization using HAP-coated MNPs. Subsequent cross-linking could "lock" BCL in its favorable conformation, while HAP-coated MNPs could facilitate the recovery of immobilized BCL and simplify the biodiesel purification. To the best of our knowledge, this is the first report on BCL immobilization integrating surfactant imprinting and mCLEAs. The optimal conditions for mCLEAs preparation, along with the effect of different surfactants (anionic, cationic, and non-ionic) on the catalytic activity of BCL mCLEAs in transesterification were studied. The optimized surfactant-activated BCL mCLEAs was further used in transesterification of waste cooking oils to biodiesel. In addition, a detailed analysis of solvents, methanol-to-oil molar ratio, and temperatures on the yield of biodiesel production was presented. The results obtained in the research are expected to provide a reliable basis for further exploration of lipase immobilization and efficient biodiesel production in industry.

2. Results and Discussion

2.1. Preparation and Characterization of Immobilized Lipase

In this study, the prepared MNPs encapsulated by hydroxyapatite (HAP) were used as immobilization supports. The amino functionalization of HAP-coated MNPs was carried out using 3-aminopropyltrimethoxysilane (APTES) for efficient enzyme attachment. Typically, the preparation procedure of immobilization supports and surfactant-activated BCL mCLEAs were performed according to the scheme shown in Scheme 1. The prepared magnetic supports and immobilized BCL were characterized by fourier transform infrared spectroscopy (FT-IR), transmission electron microscope (SEM) and vibrating sample magnetometer (VSM).
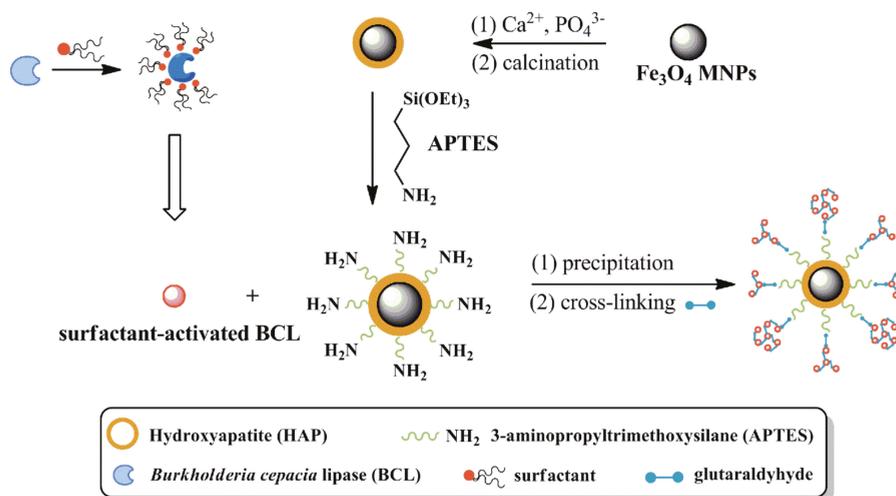
Scheme 1. Preparation procedure of immobilization supports and surfactant-activated *Burkholderia cepacia* lipase (BCL) magnetic cross-linked enzyme aggregates (mCLEAs).

FTIR characterization was performed to investigate the chemical composition of functionalized MNPs and immobilized BCL. Spectra were recorded on over the region from 4000 to 400 cm^{-1}.

As shown in Figure 1, the strong peak at 588 and 639 cm^{-1} corresponds to the stretching vibration of Fe-O bond. The characteristic absorption bands related to the HAP appease at 565 and 1044 cm^{-1}, which are assigned to phosphate groups [21]. In the IR spectrum of modified MNPs and BCL mCLEAs, the characteristic absorption bands related to the functional groups of HAP emerged clearly, which demonstrated the successful incorporation of MNPs with HAP. For all immobilized lipases, including BCL CLEAs, BCL mCLEAs and surfactant-activated BCL mCLEAs, the typical IR bands responsible for the lipase that were chemically covalent-bonded to the functionalized MNPs were observed at 1642 cm^{-1} for amide I (C=O stretching vibration) and at 1539 cm^{-1} for amide II (N-H bending vibration), respectively. Besides, compared with the results shown in Figure 1, aliphatic C-H stretch band at 2859 and 2927 cm^{-1}, corresponding to C-H stretching vibrations, are clearly observed in all immobilized lipases, which also indicated the successful loading of lipase.
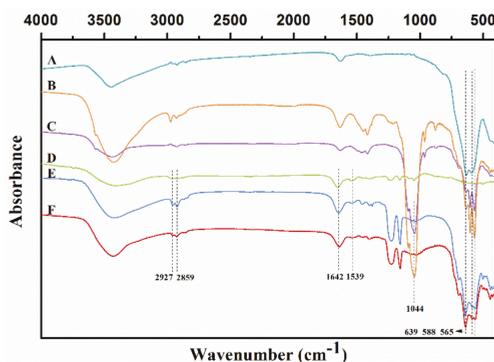
Figure 1. Spectra of (**A**) Fe$_3$O$_4$ MNPs, (**B**) hydroxyapatite coated magnetic nanoparticles (HAP-coated MNPs), (**C**) 3-aminopropyltrimethoxysilane (APTES)-HAP-coated MNPs, (**D**) BCL CLEAs, (**E**) BCL mCLEAs, (**F**) Triton-activated BCL mCLEAs.

In order to assess morphology, size and composition of functionalized MNPs and immobilized BCL, SEM images were collected and illustrated in Figure 2. As seen in Figure 2, bare Fe$_3$O$_4$ MNPs formed significantly dense agglomeration, because of their high surface energy and strong dipole-dipole interactions. It is obvious that the structure of Fe$_3$O$_4$ MNPs becomes looser and more evenly distributed after being functionalized with HAP (Figure 2B) and APTES (Figure 2C), suggesting that surface modification is favorable for preventing aggregation of Fe$_3$O$_4$ MNPs. At the same time, the rough surface of Fe$_3$O$_4$ MNPs also increased the surface area for attachment of enzyme.

The crucial structure factors in aggregated-based enzyme immobilization, including morphological topographies, structural arrangement and size, play an important role in affecting substrate affinity and stability of biocatalyst [22]. Besides, the particle size of enzymes is an important property of any heterogeneous catalysis since it can directly affect the diffusion of substrates and catalytic efficiency, especially in the internal enzymes of highly compact aggregates [23]. SEM images (Figure 2D) of standard BCL CLEAs revealed no defined morphologies and large size particles. Moreover, standard BCL CLEAs presented a uniform and compact surface with the presence of few tiny pores. On the contrary, after the incorporation of functionalized MNPs, BCL mCLEAs formed spherical structures and small particle sizes, which could reduce inner steric hindrance in closely packed CLEAs. It is noteworthy that the presence of functionalized MNPs displayed large active surface available for lipase immobilization, therefore were important for development of a stabilized enzyme-matrix. Furthermore, a loose and homodispersed structure of Triton-activated BCL mCLEAs was found in Figure 2F, suggesting that the formation of large aggregates were forbidden by the imprinting of surfactants. From the SEM outcomes, it can be discerned that, thanks to the coating of surfactants, lipase could be uniformly dispersed on functionalized MNPs, which could contribute to a wider

surface area with more catalytic sites and decrease the diffusion limit. Consequently, compared with standard BCL CLEAs, Triton-activated BCL mCLEAs could perform superior catalytic efficiency.
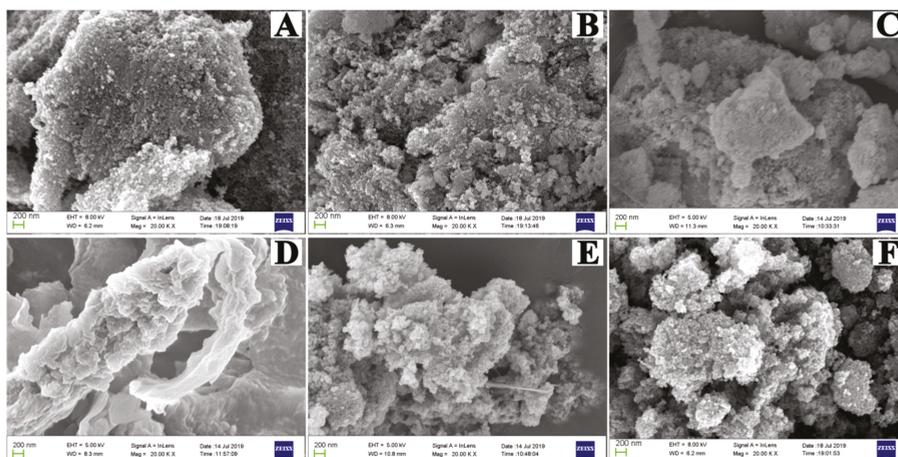
Figure 2. Images of(**A**) Fe_3O_4 MNPs, (**B**) HAP-coated MNPs, (**C**) APTES-HAP-coated MNPs, (**D**) BCL CLEAs, (**E**) BCL mCLEAs, (**F**) Triton-activated BCL mCLEAs.

The magnetic property of functionalized MNPs and immobilized BCL were measured using VSM. The hysteresis curves of the Fe_3O_4 MNPs, HAP-coated MNPs, APTES-HAP-coated MNPs, BCL mCLEAs and Triton-activated BCL mCLEAs shown in Figure 3, exhibited a perfect sigmoidal behavior, corresponding to a typical superparamagnetism.
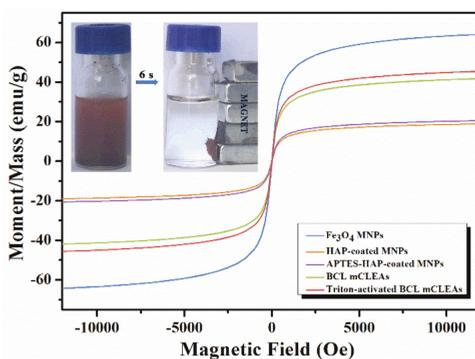
Figure 3. Hysteresis loops of Fe_3O_4 MNPs, HAP-coated MNPs, APTES-HAP-coated MNPs, BCL mCLEAs and Triton-activated BCL mCLEAs. The inner shows the easy magnetic separation of Triton-activated BCL mCLEAs in reaction mixture.

With further functionalization of MNPs, the saturation magnetization value decreased and correlated with the increase of the core-shell layer. Interestingly, it is obviously observed that the saturation magnetization value of Triton-activated BCL mCLEAs increased visibly compared to BCL mCLEAs. It might be due to the uniform dispersion of lipase on MNPs and availability of large surface area which decreased the shielding-effect of the out layer substances. As seen in Figure 3 (inner), Triton-activated BCL mCLEAs showed fast response (6s) to the external magnetic field andcould be easily recovered from the reaction mixture. After removing the external magnetic field, the magnetic

immobilized BCL redispersed rapidly by a slight shake, indicating good dispersion and efficient recyclability in industrial application.

2.2. Optimization of the Immobilization Conditions

In this study, the enzymes were precipitated by adding water-miscible organic solvents (acetone, ethanol and 2-propanol), PEG 800 and ammonium sulfate. The optimum precipitant was selected by measuring the transesterification activity of the corresponding BCL mCLEAs. Compared with free BCL, all BCL mCLEAs prepared using different precipitants performed higher transesterification activity in organic solvent. Among the protein precipitants evaluated, ammonium sulfate showed maximum recovery of activity (Figure 4a), therefore was further used in BCL immobilization.
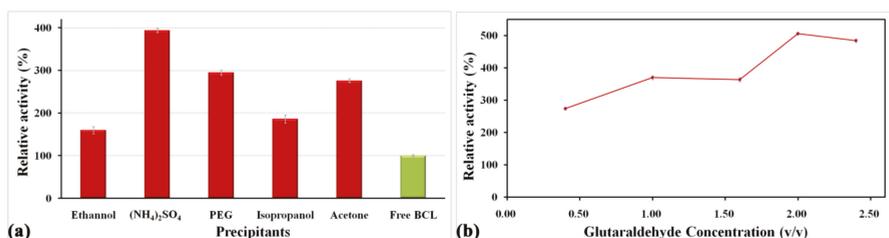
Figure 4. (**a**) Precipitant type and (**b**) glutaraldehyde concentration on the activities of BCL mCLEAs.

Traditionally, glutaraldehyde has been extensively used as the cross-linking agent to prepare CLEAs of various enzymes and exhibited a strong effect on activity and particle size of enzyme aggregates. The activity recovery of CLEAs greatly depends on the type of enzyme and the concentration of glutaraldehyde [24]. Lower glutaraldehyde concentration affects the cross-linking efficiency, which might result in enzyme leakage in immobilization, while excessive glutaraldehyde can induce the flexibility of enzymes and the active site availability, consequently, decreasing the activity recovery of CLEAs [25,26]. In this study, the influence of glutaraldehyde concentration on activity of BCL mCLEAs was investigated by using various concentrations of glutaraldehyde in cross-linking. As shown in Figure 4b, the optimum glutaraldehyde concentration of BCL mCLEAs was 2.0% (v/v).

2.3. Hyperactivation of BCL mCLEAs with Surfactants

A pivotal challenge in lipase immobilization is to open the lid of lipases and fix their open form for the exposure of active site. Surfactant imprinting is an efficient approach to activate lipases by facilitating lid-opening. Like other lipases, BCL also consists of a mobile element at the surface, which composed of two helical elements (a5- and a9-helices) and covers the active site [18]. To improve the catalytic performance of BCL mCLEAs, BCL was imprinted in the presence of surfactants prior to immobilization. Thus, four different surfactants with different properties (cationic, anionic and non-ionic) were investigated for modulating the activity of BCL mCLEAs in biodiesel production. As seen in Figure 5, Triton X-100 exhibited maximum effect on the enhancement of lipase activity in low surfactant concentration, while the addition of sodium bis-2-(ethylhexyl) sulfosuccinate (AOT) showed the least influence. The optimal surfactant in proper concentration acting as a bipolar agent, could simulate the amphiphilic environment to benefit the exposure of hydrophobic regions in the active site. Meanwhile, surfactants may also promote the dissociation of large aggregates formed, thus slightly increase the enzymatic activity of lipase (Figure 2).
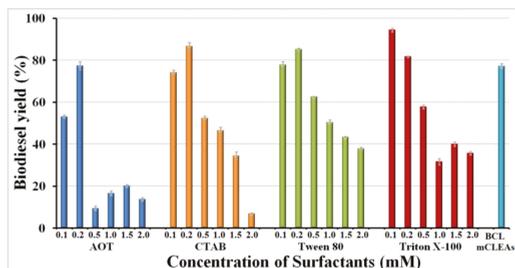
Figure 5. Four different surfactants activation on activity of surfactant-activated BCL mCLEAs in biodiesel production.

However, the increase of surfactant concentration led to gradual decrease of biodiesel yield in all cases, indicating that surfactants showed positive and negative effect on the activity of lipase. Additional detergent molecules may bind to the active site region of lipase, blocking the substrate access, inducing inhibition [27]. Compared with ionic surfactants (AOT and cetrimonium bromide (CTAB)), nonionic surfactants (Triton X-100 and Tween 80) were preferred aiming at regulating the activity of BCL (Figure 5). As the main interaction between the enzyme and nonionic surfactants is hydrophobic interaction while anionic or cationic surfactants perform electrostatic interactions [28], mild hydrophobic interaction between BCL and the surfactant might be important to trigger the interfacial activation mechanism. Therefore, nonionic Triton X-100 and Tween 80 were further studied to confirm the optimal amphiphile and surfactant concentration. As performed in Figure 6, the maximum hyperactivation of BCL mCLEAs was observed in the pretreatment of BCL with 0.1 mM Triton X-100, and the optimal Triton-activated BCL mCLEAs were used for further experiments.
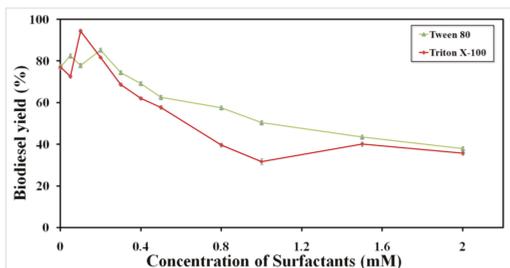
Figure 6. Surfactants (Triton X-100 and Tween 80) concentration in surfactant-activated BCL mCLEAs preparation.

2.4. Biodiesel Production

For the economic feasibility of biodiesel production, solvents, methanol-to-oil molar ratio, and reaction temperature are important variables to optimize for transesterification step. As a result of oxidative reactions occurring during cooking and long-term storage in air, WCO generally exhibits a dramatic increase in viscosity and saponification value [7]. Compared to the fresh oil, high viscose WCO is not favored in biodiesel production. Using solvents could reduce the viscosity of the reaction medium and decrease the diffusion limitations, while it might also directly affect the enzyme structure and activity. In general, hydrophobic solvents could promote the interface and stabilize lipases on their open assembly, causing the hyperactivation of these enzymes. To select the most suitable medium, five hydrophobic solvents commonly used in transesterification were tested in biodiesel production (Figure 7a). It can be clearly seen that biodiesel yield is remarkably dependent on the type of solvent. Overall, Triton-activated BCL mCLEAs exhibited higher activity than BCL mCLEAs and free BCL in all the solvents tested, and the changing trend of their activity in various solvents was accord with

BCL mCLEAs and free BCL. In case of Triton-activated BCL mCLEAs, the best results were achieved using n-hexane with a yield of up to 94% biodiesel, which was 3.3-fold higher than that in free BCL catalyzed reaction. Interestingly, surfactant hyperactivation in combination with immobilization could fasten lipase in their active conformation, allowing biodiesel production performed in solvent without further addition of water, which was in accordance with earlier reports [29,30].
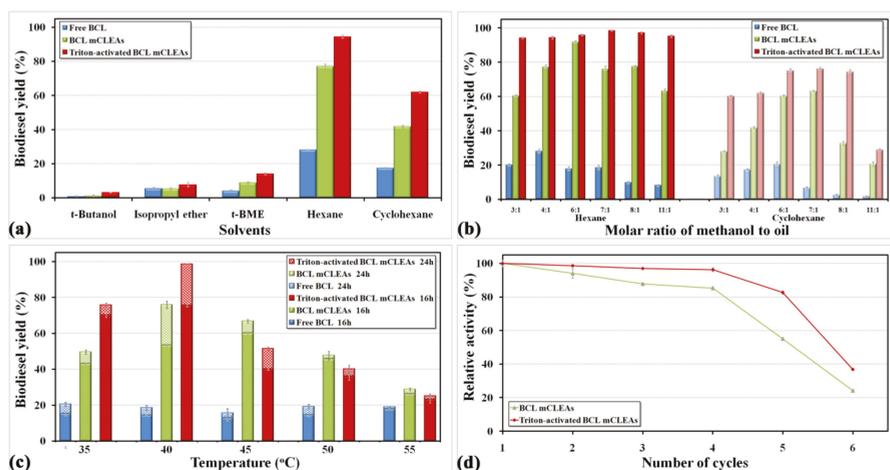
Figure 7. Reaction parameters on biodiesel production catalyzed by free BCL, BCL mCLEAs and Trion-activated BCL mCLEAs and reusability of immobilized BCL. (**a**) Solvents, (**b**) molar ratio of methanol to oil, (**c**) temperature, (**d**) reusability.

The methanol:oil molar ratio can have a significant effect on the reaction yield because excess methanol increases the reaction rate and drives high yield of biodiesel, while a high concentration of methanol leads to inactivation of lipases. In this study, experiments were performed at different molar ratios of methanol to WCO ranging from 3:1 (stoichiometric ratio) to 11:1 both in hexane and cyclohexane with methanol added only once. As shown in Figure 7b, Triton-activated BCL mCLEAs exhibited higher biodiesel yields in one-time addition of methanol under all the experimental conditions, especially when hexane was used as solvent. Meanwhile, owing to the significant inhibitory effect of excess methanol in one time addition, free BCL showed low activity in biodiesel production. It is worth noting that yields of Triton-activated BCL mCLEAs catalyzed reactions in hexane exceeded 90% in a wide range of methanol-to-WCO ratio, while the reaction yield of BCL mCLEAs catalyzed reaction decreased with methanol-to-WCO ratio exceeding 6:1. Consequently, it can be confirmed that surfactants pretreatment provided not only hyperactivation but also protection to lipases from denaturation in excess methanol. In addition, the maximum biodiesel yield was observed at a methanol-to-WCO ratio of 7:1 for Triton-activated BCL mCLEAs. Consequently, the minimal stoichiometric methanol-to-WCO ratio of 7:1 was chosen in further experiments.

Most of the enzymatic transesterification depends on temperature, which could enhance reaction rate and improve the dispersion of immobilized particles in reaction medium with better mass transfer between the reactants [31]. However, thermal denaturation of the enzyme might occur with elevation of temperature, typically according to the property of enzyme and immobilized methods. The effect of temperature on the yield of biodiesel during the transesterification of WCO has been investigated over a temperature range from 35 to 55 °C. According to Figure 7c, the optimum operational temperature for Triton-activated BCL mCLEAs and BCL mCLEAs was 40 °C, with biodiesel yields of 98% and 76% respectively after 24 h, and further increase of temperature will result in decrease of biodiesel yields simultaneously. Besides, Triton-activated BCL mCLEAs showed better activity below 40 °C, while BCL mCLEAs performed higher biodiesel yield over 40 °C. The suitable covalent cross-linking

with functionalized MNPs provided extra structure stabilization in mCLEAs, requiring much more energy to the disruption of this stable structure than free enzyme [32]. Nevertheless, the accessible active site of lipases achieved by surfactants pretreatment might be more sensitive to high temperature denaturation [33].

In summary, the optimal reaction conditions for Triton-activated BCL mCLEAs catalyzed transesterification of WCO are as follows: hexane used as solvent, molar ratio of methanol-to-WCO 7:1 in one-time addition, reaction temperature 40 °C. To verify the feasibility of the whole process at a larger scale, transesterification of WCO were performed under optimal conditions adding proper amount of Triton-activated BCL mCLEAs (the initial content of BCL was 240 mg in immobilization) to a mixture of 1 g WCO in 20 mL hexane. The biodiesel yield reached 94% after shaking at 40 °C for 48 h. Triton-activated BCL mCLEAs showed good activity and stability under higher oil content, indicating the possibility of its scale-up application in bioreactor systems.

2.5. Reusability

Reusability of immobilized enzyme is a chief criterion for its cost-effective use for potential industry applications. The utilization of functionalized MNPs facilitates the consequent reuse of immobilized enzyme. To investigate the reusability of BCL mCLEAs and Triton-activated BCL mCLEAs, the immobilized lipases were recovered by magnetic separation, and applied in the consecutive batches of biodiesel reactions under optimized conditions. Assessments of the operational stability were analyzed for 6 cycles and presented in Figure 7d. As observed, Triton-activated BCL mCLEAs showed no significant loss in the catalytic activity after subsequent consecutive reuse for 4 cycles, and kept 82% relative activity after continuous running 5 cycles. Meanwhile, the relative activity of BCL mCLEAs was 55% after 5 cycles, implying that BCL could possess good long-term stability with surfactant pretreatment. The protein denaturation in one time addition of methanol and byproduct inhibition might be account for the decrease in biodiesel yield in long-term reuses [34].

3. Materials and Methods

3.1. Materials

Burkholderia cepacia lipase (powder, Amano Lipase PS, ≥3000 U/g) and fatty acid methyl ester standards were purchased from Sigma-Aldrich (St. Louis, MO, USA). Also, 3-aminopropyl triethoxysilane (APTES), glutaraldehyde (25%, v/v) and 2-phenyl ethanol (>98%, CP) were supplied by Aladdin (Shanghai, China). Sodium bis-2-(ethylhexyl) sulfosuccinate (AOT) were procured from Acros (USA). Waste cooking oil (WCO) was obtained from local restaurant around Ningxia University campus (Yinchuan, China) with the following fatty acid compositions: 10.48% palmitic acid, 15.04% stearicacid, 38.44% oleic acid, 23.76% linoleic acid, and 1.72% linolenic acid. The WCO sample was filtered to separate impurities and solids in the oil. The physical properties of WCO are saponification value of 197.3 mg KOH/g, acid value of 4.37 mg KOH/g, and average molecular weight of 870.9 g/mol. All other chemicals were of analytical or chromatographical grade and used as purchased.

3.2. Preparation of Magnetic Support

Preparation of HAP-coated MNPs was carried out according to the previously reported method [35]. Initially, MNPs cores were prepared by the conventional co-precipitation method. Typically, $FeCl_2 \cdot 4H_2O$ (1.1 g) and (3.0 g) of $FeCl_3 \cdot 6H_2O$ were dissolved in 90 mL deionized water under the protection of argon, with subsequent addition of 25% ammonia solution (30 mL) under vigorous stirring at room temperature. After stirring for 30 min, a 60 mL aqueous solution composed of $Ca(NO_3)_2 \cdot 4H_2O$ (7.1 g) and $(NH_4)_2HPO_4$ (2.3 g) adjusted to pH=11 was added drop wise to the above suspension under continuous stirring. Subsequently, the resultant mixture was heated to 90 °C and stirred for 2 h. After cooling to room temperature and aging in the mother solution overnight, the obtained

precipitates were washed several times with deionized water until neutral and lyophilized for 12 h. The HAP-coated MNPs were obtained by calcining the materials in air at 300 °C for 3 h.

To obtain 3-aminopropyl trimetoxysilane functionalized HAP-coated MNPs (APTES-HAP-coated MNPs), HAP-coated MNPs (1.0 g) were suspended in a solution composed of 30 mL anhydrous toluene and 0.44 g of APTES. The mixture was refluxed under Ar atmosphere for 12 h, and then washed several times with ethanol, magnetically separated, and subsequently lyophilized prior to use.

3.3. Lipase Immobilization

BCL mCLEAs were produced according to the procedure described in Scheme 1. Firstly, 10 mg of APTES-HAP-MNPs were dispersed in 1 mL of BCL solution (10 mg/mL, 0.1 M phosphate buffer, pH 7.0) and shaken for 15 min at 30 °C. Then 5 mL of precipitant was added with stirring at 4 °C for 30 min. After precipitation, glutaraldyhyde was added drop wise into the suspension and stirred for 3 h at 30 °C. Afterwards, BCL mCLEAs were collected by centrifugation and washed thrice with phosphate buffer and deionized water, lyophilized and finally stored at 4 °C.

During optimization of the immobilization conditions, the effects of precipitants (acetone, ethanol, isopropanol, PEG 800 (1 g/mL), and saturated ammonium sulfate solution) and concentration of glutaraldehyde on the activity recovery of BCL mCLEAs were investigated.

The surfactant-activated BCL mCLEAs was prepared using cationic (CTAB), anionic (AOT) and nonionic (Tween 80 and Triton Triton X-100) surfactants at various concentrations. Then, 1 mL of BCL solution and appropriate amount of surfactant were mixed and stirred at 4 °C for 30 min. After incubated for 24 h at 4 °C, the suspended solution was sequentially used for BCL mCLEAs preparation under optimal conditions.

3.4. Characterization

The prepared support matrix and immobilized lipase described above were characterized using FTIR, SEM and VSM. The Fourier transform infrared (FTIR) spectra were acquired using a Perkin Elmer Frontier spectrometer (Spectrum Two, Waltham, MA, USA) equipped with an Attenuate Total Reflection (ATR) accessory. Samples were analyzed as KBr pellets in the range of 400 to 4000 cm^{-1} at a resolution of 0.5 cm^{-1}. The morphology of the particle surface was observed using a scanning electron microscope (SEM, Sigma HD, ZEISS, Germany), with deposition of a thin coating of gold onto the samples prior to analyses. The magnetic properties were detected by a vibrating sample magnetometer (VSM, MicroSense EZ9, Lowell, MA, USA) at room temperature.

3.5. Activity Assay

In studying the optimal conditions for BCL mCLEAs preparation, the enzymatic transesterification activities of free lipase and immobilized BCLs were assayed via transesterification reaction of 2-phenyl ethanol with vinyl acetate according to the method introduced previously [36]. The reaction mixture contained 10 mg of 2-phenylethanol, 1 mL of vinyl acetate and 10 mg of lipase (the initial content of BCL was 10 mg in preparing BCL CLEAs and BCL mCLEAs), and the reactions were carried out at 30 °C with continuous shaking at 220 rpm. After 24 h of reaction, samples were withdrawn and analyzed by high-performance liquid chromatography (HPLC). All experiments were repeated at least three times. The relative activity of BCL mCLEAs was calculated with the following equation:

$$\text{Relative activity } (\%) = \frac{\text{Transesterification yield of immobilized BCL}}{\text{Trasesterification yield of free BCL}} \times 100$$

3.6. Enzymatic Transesterification for Biodiesel Production

The transesterification of WCO were carried out at 40 °C in a 10 mL screw-capped vessel for 24 h with continuous shaking at 220 rpm. Unless otherwise stated, a typical reaction mixture consisted of 50 mg WCO, 2.0 mL hexane, 10 mg of lipase (the initial content of BCL was 10 mg in preparing

BCL CLEAs and BCL mCLEAs) and methanol using methanol: oil molar ratio of 4:1. Single factor optimization was conducted to determine optimal reaction parameters for transesterification of WCO to biodiesel. Various conditions including kinds and concentration of surfactants, solvents, molar ratio of methanol to oil and temperature (°C) were investigated. The transesterification reaction of large scale with 1 g WCO were carried out as described in Section 2.4. All biodiesel reactions were performed in dried solvents without any water added. The yield of biodiesel (20 μL) was analyzed in different time intervals using gas chromatography.

3.7. Analytical Methods

HPLC was conducted with Shimadzu LC-2010A HT apparatus using C18 column (UltimateXB-C18, 5 μm, 4.6 × 150 mm, Welch). The samples were analyzed with a mixture of MeOH/water = 80:20 (v/v) as eluent at 0.8 mL/min for 9 min at 254 nm.

Fatty acid methyl esters (FAMEs) were analyzed by a Fuli9790 plus gas chromatography (Fuli, Zhejiang, China) fitted with a flame ionization detectorcity (FID, Zhejiang, China), and a KB-FFAP column (30 m × 0.32 mm × 0.25 μm). Nitrogen gas was a carrier at continues flow of 1.0 mL/min. The oven (Zhejiang, China) temperature was set and at 160 °C maintained for 2 min, then a heating ramp was applied up to 240 °C at a rate of 10 °C /min, and the temperature of the oven was maintained at 240 °C for 15 min. The temperatures of the injector (Zhejiang, China) and the detector (Zhejiang, China) were set at 270 and 280 °C, respectively. Methyl tridecanoate was used as internal standard, and the biodiesel yield (%) was calculated by peaks area of standard FAME peaks.

3.8. Reusability

The reusability of Triton-activated BCL mCLEAs and BCL mCLEAs for the transesterification of WCO were also investigated under optimal conditions. After each batch reaction, immobilized BCL was recovered by magnetic separation and washed with n-hexane. The washed biocatalyst was reused consecutively in repetitive cycles. The biodiesel yield of the first reaction was set as 100% and the FAMEs yield in the subsequent reactions was calculated accordingly.

4. Conclusions

A facile and effectual surfactant imprinting method to expose the lipase active site integrating amino functionalized HAP-coated MNPs was established to immobilize CLEAs of BCL attaining enhanced activity and stability. The as-prepared Triton-activated BCL mCLEAs was subsequent processed in enzymatic transesterification of waste cooking oil for biodiesel production, and showed 98% biodiesel yield under optimal conditions, which was 5.3-fold higher than the free lipase. This study proved that hyperactivation with surfactant could significantly improve the resistance of lipase to methanol in one-time addition, when compared to BCL mCLEAs and free BCL. In addition, surfactant imprinting in combination with immobilization could fasten lipase in their active conformation, allowing biodiesel production performed in solvent without further addition of water, and thus displayed priority in downstream purification of biodiesel over ordinary immobilization methods. Besides, the green immobilization with functionalized MNPs facilitates fast and easy recovery of lipase, and the corresponding immobilized BCL was reused for 4 cycles without significant loss in the catalytic activity. Furthermore, this work provides a promising approach for immobilization of other lipases, which can be used with success in green and clean production processes.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mahmudul, H.M.; Hagos, F.Y.; Mamat, R.; AbdulAdam, A.; Ishak, W.F.W.; Alenezi, R. Production, characterization and performance of biodiesel as an alternative fuel in diesel engines-a review. *Renew. Sustain. Energy Rev.* **2017**, *72*, 497–509. [CrossRef]

2. Laesecke, J.; Ellis, N.; Kirchen, P. Production, analysis and combustion characterizationof biomass fast pyrolysis oil-biodiesel blends for use in diesel engines. *Fuel* **2017**, *199*, 346–357. [CrossRef]

3. Jamil, F.; Alhaj, L.; Almuhtaseb, A.H.; Baawain, M.; Rashid, U.; Ahmad, M.N.M. Current scenario of catalysts for biodieselproduction: A critical review. *Rev. Chem. Eng.* **2018**, *34*, 267–297. [CrossRef]

4. Banković-Ilić, I.B.; Stamenković, O.S.; Veljković, V.B. Biodiesel production fromnon-edible plant oils. *Renew. Sustain. Energy Rev.* **2012**, *16*, 3621–3647. [CrossRef]

5. Hama, S.; Noda, H.; Kondo, A. How lipase technology contributes to evolution ofbiodiesel production using multiple feedstocks. *Curr. Opin. Biotechnol.* **2018**, *50*, 57–64. [CrossRef]

6. Gebremariam, S.N.; Marchetti, J.M. Economics of biodiesel production: Review. *Energy Convers. Manag.* **2018**, *168*, 74–84. [CrossRef]

7. Moazeni, F.; Chen, Y.-C.; Zhang, G. Enzymatic transesterification for biodiesel production from usedcooking oil, a review. *J. Clean. Prod.* **2019**, *216*, 117–128. [CrossRef]

8. Lopresto, C.; Naccarato, S.; Albo, L.; De Paola, M.; Chakraborty, S.; Curcio, S.; Calabro, V. Enzymatic transesterification of waste vegetable oil to producebiodiesel. *Ecotoxicol. Environ. Saf.* **2015**, *121*, 229–235. [CrossRef]

9. Chhetri, A.; Watts, K.; Islam, M. Waste cooking oil as an alternate feedstock forbiodiesel production. *Energies* **2008**, *1*, 3–18. [CrossRef]

10. Sabudak, T.; Yildiz, M. Biodiesel production from waste frying oils and itsquality control. *Waste Manag.* **2010**, *30*, 799–803. [CrossRef]

11. Kim, K.H.; Lee, O.K.; Lee, E.Y. Nano-immobilized biocatalysts for biodiesel production from renewable and sustainable resources. *Catalysts* **2018**, *8*, 68. [CrossRef]

12. Facin, B.R.; Melchiors, M.S.; Valério, A.; Oliveira, J.V.; Oliveira, D. Driving immobilized lipases as biocatalysts: 10 years state of the art and future prospects. *Ind. Eng. Chem. Res.* **2019**, *58*, 5358–5378. [CrossRef]

13. Filho, D.G.; Silva, A.G.; Guidini, C.Z. Lipases: Sources, immobilization methods, and industrial applications. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 7399–7423. [CrossRef]

14. Xu, M.-Q.; Wang, S.-S.; Li, L.-N.; Gao, J.; Zhang, Y.-W. Combined cross-linked enzyme aggregatesas biocatalysts. *Catalysts* **2018**, *8*, 460. [CrossRef]

15. Sheldon, R.A. CLEAs, combi-CLEAs and 'smart' magnetic CLEAs: Biocatalysis in a bio-based economy. *Catalysts* **2019**, *9*, 261. [CrossRef]

16. Liu, D.-M.; Chen, J.; Shi, Y.-P. Advances on methods and easy separated support materials forenzymes immobilization. *TrAC-Trends Anal. Chem.* **2018**, *102*, 332–342. [CrossRef]

17. Izadia, A.; Meshkinia, A.; Entezari, M.H. Mesoporous superparamagnetic hydroxyapatite nanocomposite: Amultifunctional platform for synergistic targeted chemo-magnetotherapy. *Mater. Sci. Eng. C Mater.* **2019**, *101*, 27–41. [CrossRef]

18. Sánchez, D.A.; Tonetto, G.M.; Ferreira, M.L. Burkholderia cepacia lipase: A versatile catalyst in synthesisreactions. *Biotechnol. Bioeng.* **2018**, *115*, 6–24. [CrossRef]

19. Barbe, S.; Lafaquiere, V.; Guieysse, D.; Monsan, P.; Remaud-Siméon, M.; Andre, I. Insights into lid movements of Burkholderia cepacialipase inferred from molecular dynamics simulations. *Proteins* **2009**, *77*, 509–523. [CrossRef]

20. Gabriele, F.; Spreti, N.; Giacco, T.D.; Germani, R.; Tiecco, M. Effect of surfactant structure on the superactivity of Candida rugosalipase. *Langmuir* **2018**, *34*, 11510–11517. [CrossRef]

21. Zhang, N.; Gao, T.; Wang, Y.; Wang, Z.; Zhang, P.; Liu, J. Environmental pH-controlledloading and release of protein on mesoporous hydroxyapatite nanoparticles forbone tissue engineering. *Mater. Sci. Eng. C Mater.* **2015**, *46*, 158–165. [CrossRef]

22. Kannan, S.; Marudhamuthu, M. Development of chitin cross-linked enzyme aggregates of L-methioninase for upgraded activity, permanence andapplication as efficient therapeutic formulations. *Int. J. Biol. Macromol.* **2019**, *141*, 218–231. [CrossRef]

23. Shaarani, S.M.; Jahim, J.M.; Rahman, R.A.; Idris, A.; Murad, A.M.A.; Illias, R.M. Silanized maghemite for cross-linked enzyme aggregates of recombinantxylanase from Trichoderma reesei. *J. Mol. Catal. B Enzym.* **2016**, *133*, 65–76. [CrossRef]

24. Rehman, S.; Bhatti, H.N.; Bilal, M.; Asgher, M. Cross-linked enzyme aggregates (CLEAs) of Pencilluim notatum lipaseenzymewith improved activity, stability and reusabilitycharacteristics. *Int. J. Biol. Macromol.* **2016**, *91*, 1161–1169. [CrossRef]

25. Cui, J.; Lin, T.; Feng, Y.; Tan, Z.; Jia, S. Preparation of spherical cross-linked lipaseaggregates with improved activity, stabilityand reusability characteristic in water-in-ionicliquid microemulsion. *J. Chem. Technol. Biotechnol.* **2017**, *92*, 1785–1793. [CrossRef]

26. Yamaguchi, H.; Kiyota, Y.; Miyazaki, M. Techniques for preparation of cross-linked enzyme aggregates and their applications in bioconversions. *Catalysts* **2018**, *8*, 174. [CrossRef]

27. Bañó, M.C.; González-Navarro, H.; Abad, C. Long-chain fatty acyl-CoA esters induce lipase activation in the absence of a water-lipid interface. *BBA Mol. Cell Biol. Lipids* **2003**, *1632*, 55–61. [CrossRef]

28. Delorme, V.; Dhouib, R.; Canaan, S.; Fotiadu, F.; Carriere, F.; Cavalier, J.F. Effects of surfactants on lipase structure, activity, and inhibition. *Pharm. Res.* **2011**, *28*, 1831–1842. [CrossRef]

29. Mukherjee, J.; Gupta, M.N. Molecular bioimprinting of lipases with surfactantsand its functional consequences in low water media. *Int. J. Biol. Macromol.* **2015**, *81*, 544–551. [CrossRef]

30. Zhang, W.-W.; Yang, X.-L.; Jia, J.Q.; Wang, N.; Hu, C.-L.; Yu, X.-Q. Surfactant-activated magnetic cross-linked enzyme aggregates (magnetic CLEAs) of Thermomyces lanugunosus lipase for biodiesel production. *J. Mol. Catal. B Enzym.* **2015**, *115*, 83–89. [CrossRef]

31. Lee, H.V.; Yunus, R.; Juan, J.C.; Taufiq-Yap, Y.H. Process optimization design forjatropha-based biodiesel production using response surface methodology. *Fuel Process. Technol.* **2011**, *92*, 2420–2428. [CrossRef]

32. Dong, T.; Zhao, L.; Huang, Y.; Tan, X. Preparation of cross-linked aggregates ofaminoacylase from Aspergillus melleus by using bovine serum albumin as aninert additive. *Bioresour. Technol.* **2010**, *101*, 6569–6571. [CrossRef]

33. Rehm, S.; Trodler, P.; Pleiss, J. Solvent-induced lid opening in lipases: A molecular dynamics study. *Protein Sci.* **2010**, *19*, 2122–2130. [CrossRef]

34. Xie, W.; Huang, M. Immobilization of Candida rugosa lipase onto graphene oxide Fe3O4nanocomposite: Characterization and application for biodiesel production. *Energy Convers. Manag.* **2018**, *159*, 42–53. [CrossRef]

35. Xie, W.; Zang, X. Covalent immobilization of lipase onto aminopropyl-functionalizedhydroxyapatite-encapsulated-γ-Fe$_2$O$_3$ nanoparticles: A magneticbiocatalyst for interesterification of soybean oil. *Food Chem.* **2017**, *227*, 397–403. [CrossRef]

36. Zhang, W.-W.; Yang, H.-X.; Liu, W.-Y.; Wang, N.; Yu, X.-Q. Improved performance of magnetic cross-linked lipase aggregates by interfacial activation: A robustand magnetically recyclable biocatalyst for transesterification of Jatropha oil. *Molecules* **2017**, *22*, 2157. [CrossRef]

Article

Response Surface Methodology Approach for Optimized Biodiesel Production from Waste Chicken Fat Oil

Fatima Shafiq [1], Muhammad Waseem Mumtaz [1,*], Hamid Mukhtar [2], Tooba Touqeer [1],
Syed Ali Raza [3], Umer Rashid [4,*], Imededdine Arbi Nehdi [5,6] and Thomas Shean Yaw Choong [7]

[1] Department of Chemistry, University of Gujrat, Gujrat 50700, Pakistan; fatimashafiq369@gmail.com (F.S.); tuba.toqir@gmail.com (T.T.)

[2] Institute of Industrial Biotechnology, Government College University, Lahore 54000, Pakistan; hamidwaseer@yahoo.com

[3] Department of Chemistry, Government College University, Lahore 54000, Pakistan; chemstone@yahoo.com

[4] Institute of Advanced Technology, Universiti Putra Malaysia, UPM Serdang, Selangor 43400, Malaysia

[5] Chemistry Department, College of Science, King Saud University, Riyadh 1145, Saudi Arabia; inahdi@ksu.edu.sa

[6] Laboratoire de Recherche LR18ES08, Chemistry Department, Science College, Tunis El Manar University, Tunis 2092, Tunisia

[7] Department of Chemical and Environmental Engineering, Universiti Putra Malaysia, UPM Serdang, Selangor 43400, Malaysia; csthomas@upm.edu.my

* Correspondence: muhammad.waseem@uog.edu.pk (M.W.M.); umer.rashid@upm.edu.my (U.R.); Tel.: +60-3-9769-7393 (U.R.)

Abstract: Biodiesel is gaining acceptance as an alternative fuel in a scenario where fossil fuel reserves are being depleted rapidly. Therefore, it is considered as the fuel of the future due to its sustainability, renewable nature and environment friendly attributes. The optimal yield of biodiesel from cheap feed stock oils is a challenge to add cost effectiveness without compromising the fuel quality. In the current experiment, waste chicken fat oil was taken as the feedstock oil to produce biodiesel through the chemical and enzymatic route of transesterification. The process of chemical transesterification was performed using KOH and sodium methoxide, while enzymatic transesterification was done by using free *Aspergillus terreus* lipase and *Aspergillus terreus* lipase immobilized on functionalized Fe_3O_4 nanoparticles ($Fe_3O_4\_PDA\_Lipase$) as biocatalysts. The physico-chemical properties of the understudy feedstock oil were analyzed to check the feasibility as a feedstock for the biodiesel synthesis. The feedstock oil was found suitable for biodiesel production based upon quality assessment. Optimization of various reaction parameters (the temperature and time of reaction, catalyst concentration and methanol-to-oil mole ratio) was performed based on the response surface methodology (RSM). The maximum yield of biodiesel (90.6%) was obtained from waste chicken fat oil by using $Fe_3O_4\_PDA\_Lipase$ as an immobilized nano-biocatalyst. Moreover, the above said optimum yield was obtained when transesterification was done using 6% $Fe_3O_4\_PDA\_Lipase$ with a methanol-to-oil ratio of 6:1 at 42 °C for 36 h. Biodiesel production was monitored by FTIR spectroscopic analysis, whereas compositional profiling was done by GC–MS. The measured fuel properties—cloud point, pour point, flash point, fire point and kinematic viscosity—met the biodiesel specifications by American Society for Testing and Materials (ASTM).

Keywords: biodiesel; transesterification; immobilized lipase; RSM; fuel properties

1. Introduction

The rapid industrial growth and population explosion have built an immense pressure on natural resources, including fossil fuels. The whole world is determined to find suitable solutions in context with the forthcoming energy crisis. The world is in search of alternate sources of fuel to reduce its dependency on conventional fuels. Biodiesel has emerged as a promising alternative fuel in recent years due to its renewable nature and environment friendly attributes. Biodiesel may be characterized as alkyl esters of fatty acids and may be utilized easily in diesel engines without major alterations [1]. The emissions of CO and NOx from diesel burning are issues of keen interest as both are greenhouse gases and responsible for tropospheric ozone formation. It is an established fact based on the work of many researchers that, comparative to conventional diesel, combustion of biodiesel produces less CO and unburnt hydrocarbons but higher NOx emissions, probably due to a higher oxygen content in biodiesel [2,3].

Initially, synthesis of biodiesel was extensively carried out using vegetable oils and seed oils of non-edible origin. Usually, the production of biodiesel from edible oils is not cost-effective and these vegetable oils are used in food, hence are valuable. To avoid the problems associated with cost, edibility and food shortage, biodiesel production from non-edible fractions of food and related wastes is gaining sound gravity. Similarly, the biodiesel preparation from non-edible seeds like *Jatropha* is not completely feasible as the cultivation of non-edible seed oil plants may create competition with edible crops on shrinking fertile agriculture land [4]. Recently many studies were carried out for biodiesel synthesis from low cost vegetal and animal-based feed stocks like waste cooking oils and animal fats by reducing their viscosity through the transesterification process [5]. The chemical and enzymatic transesterification processes are adopted to convert fatty acids of feedstock into their alkyl esters. Both transesterification modes have their own modalities and advantages but may be optimized for high quality biodiesel [6].

Chicken fat is a poultry waste that can be used to produce biodiesel. The fat content in chicken is about 10% by weight, which is very high, and its cost is low. Commercial broiler chicken meat was reported to have relatively high contents of polyunsaturated lipids as compared with organic chicken [7]. Researchers have reported that chicken fat constitute about 25% to 35% saturated and 40% to 75% unsaturated fatty acids. Palmitic acid, along with stearic acid, linoleic acid and oleic acid, are major fatty acids in chicken fat [8,9]. The fats can be converted into alkyl esters by the process of transesterification. In an alkali-catalyzed transesterification reaction, both the glyceride and alcohol should be extensively free of water contents as the water compels the reaction to partly change into a saponification reaction, resulting in soap formation [6]. Sodium hydroxide and potassium hydroxide are commonly used as alkali catalysts, but they result in water formation during transesterification, that is why sodium and potassium methoxides are preferred for biodiesel production. Alkali catalysts are good especially for those feed stocks that contain minimal acid value. However, if the acid value of the feed stock is high, then it is recommended to perform pre-treatment acid esterification to reduce the free fatty acid contents before performing base-catalyzed transesterification of the feedstock [10].

On the other hand, enzyme-catalyzed transesterification is gaining acceptance and is considered technically comparable to alkali transesterification. This method normally employs lipase as a catalyst. Lipase-catalyzed transesterification of feedstock oils with a relatively higher free fatty acid content can be carried out without performing any pre-treatment acid esterification step that is normally required in case of alkaline transesterification [11]. However enzymatic transesterification is a high-cost process, because enzymes can be denatured easily in the presence of short-chain alcohols and it is difficult to recover [12]. To cope with these problems, enzymes are immobilized on various supports to enhance their durability. Immobilized enzymes are adoptable to harsh conditions as compared to the free enzymes and are easy to recover. Immobilization of enzymes on the matrix and beads may reduce the enzyme activity by blocking its active site and lowering the mass transfer. However, due to very small size and Brownian movement of nanoparticles, these are a potent choice for enzyme immobilization [13]. There are few reports on enzymatic transesterification of chicken fat oil [14,15].

In the present work the transesterification process was optimized to synthesize biodiesel from a cheaper source in the form of waste chicken fat. The relative effects of various catalysts and their concentration were studied and optimized for improved yields of biodiesel by involving the methanol-to-oil ratio along with reaction time and temperature. The synthesized biodiesel was also analyzed for fuel properties to check its feasibility for use in a compression ignition (CI) engine.

2. Results and Discussion

2.1. Physico-Chemical Characterization of Waste Chicken Fat Oil (WCFO)

The pre-analysis tests of WCFO revealed that the acid value of the oil was 6.56 ± 0.05 mg KOH/g, saponification value 200 ± 7.50 mg KOH/g, refractive index 1.46 ± 0.01, density 0.85 ± 0.07 g/cm^3, iodine number 75 ± 10.70 g iodine/100 g and the peroxide value was computed as (5.5 ± 0.50 meqO$_2$/kg). These values were depicted comparable with that reported by previous studies [16]. Chicken fat oil has a high acid value, which is why the acid esterification of the feedstock was done prior to the alkaline transesterification to reduce the free fatty acid content and avoid saponification.

2.2. Optimization of Biodiesel Production Process

The experimental results obtained after performing reactions as per CCRD were statistically analyzed to select the most appropriate model from the linear, 2F1, cubical and quadratic models. The model that was best suited was chosen by considering the *p*-values, R^2 values, lack-of-fit tests and adjusted R^2 values. It was observed that the quadratic model was most suited for both the chemical and enzymatic routes of biodiesel production (Table 1).

Table 1. Summary of selected quadratic models.

Feedstock	Catalysts/Biocatalysts		Selected Models	Sequential *p*-Value	Lack-of-Fit *p*-Value	Adjusted R-Squared
WCFO	Enzymes	Fe$_3$O$_4$_PDA_Lipase	Quadratic	<0.0001	0.0701	0.9713
WCFO		*Aspergillus terreus* lipase	Quadratic	<0.0001	0.1276	0.9679
WCFO	Chemicals	CH$_3$ONa	Quadratic	<0.0001	0.4021	0.9519
WCFO		KOH	Quadratic	<0.0001	0.0916	0.9518

The summary statistics clearly determined the fitness of quadratic models for chemical as well as enzymatic biodiesel production process for WCFO.

2.3. Graphs of Predicted vs. Actual Values

The predicted vs. actual value graphs for biodiesel yield depicts the fitness of the selected quadratic model. The graphs of predicted vs. actual values are shown in Figure 1, where Figure 1a–d describes the predicted vs. actual graphs based on experimental data about yield of biodiesel obtained through the transesterification of waste chicken fat oil by Fe$_3$O$_4$_PDA_Lipase (Figure 1a), *Aspergillus terreus* lipase (Figure 1b), sodium methoxide (Figure 1c) and KOH (Figure 1d). The distribution of the data along the straight line and the small difference between the predicted and actual value reveals the fitness of the quadratic model for all four experimental designs.
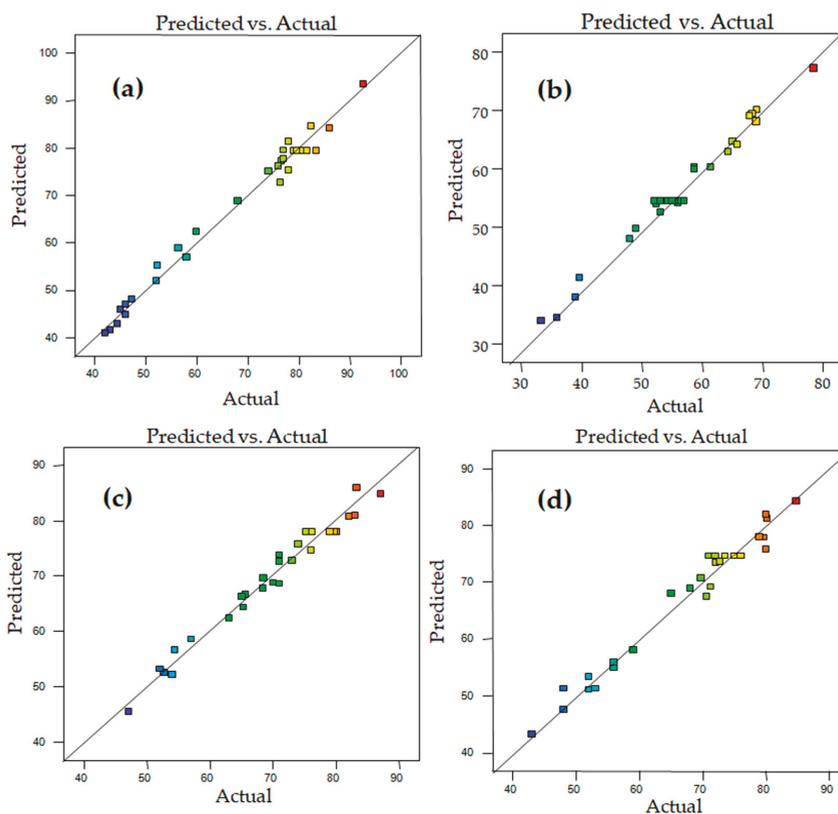
Figure 1. Graphs of predicted vs. actual yield for waste chicken fat oil-based biodiesel by Fe_3O_4_PDA_Lipase (**a**), *Aspergillus terreus* lipase (**b**), Sodium methoxide (**c**) and KOH (**d**).

2.4. Optimization of Reaction Parameters for Manufacturing of Biodiesel Using Chicken Fat Oils

The enzymatic transesterification of waste chicken oil using Fe_3O_4_PDA_Lipase as a bio-catalyst resulted in optimal biodiesel yield when transesterification reactions were performed by employing 6% Fe_3O_4_PDA_Lipaseconcentration with a 6:1 molar ratio of methanol to oil, at 42 °C for 36 h. While in case of enzymatic transesterification by *Aspergillus terreus* lipase, a 1% enzyme concentration, methanol-to-oil ratio of 6:1 and reaction temperature of 35 °C for 36 h were the optimal process conditions. However, when the sodium methoxide-catalyzed transesterification of WCFO was conducted, the optimum conditions for the reaction were a 1% catalyst level and a 6:1 methanol:oil mole ratio at 60 °C for a 1.25 h reaction time (Table 2). The optimum biodiesel yield in case of a potassium hydroxide (KOH)-catalyzed reaction was obtained at a 1% catalyst concentration, 1 h of reaction time, a methanol-to-oil ratio of 6:1 and 60 °C. Highest biodiesel yield was obtained for the nano-biocatalyst (Fe_3O_4_PDA_Lipase), which might be due to the high stability and activity of the immobilized enzyme at an elevated temperature and adoptability towards harsh conditions [17]. Moreover, the lipase can also convert the free fatty acids present in the feedstock to FAMEs. Lower yield obtained by the free lipase can be explained by the reduction of enzyme activity due to denaturation of the free enzymes at a higher temperature, which is required for biodiesel production from chicken fat oil, and the presence of short-chain alcohol [18]. For chemical transesterification, sodium methoxide was proven to be better than KOH, because sodium methoxide did not produce water, which might be responsible for saponification, and the separation of glycerol from biodiesel could be difficult, thus reducing process

efficiency. Comparable results for enzymatic and chemical transesterification of waste chicken fat oil have been reported in the published literature. Coppini et al. has reported a 90.61% biodiesel yield from chicken fat by using a 0.3 wt % NS-40116 enzyme, 1.5 of methanol:oil and 1.5 wt % water at 45 °C for 24 h [11]. Da Silva et al. has reported a 77% esterification yield by using 0.3 wt % lipase, 1:4.5 methanol:oil and 2 wt % water at 30 °C in 24 h [15]. Alptekin et al. has reported an 87.4% biodiesel yield from waste chicken fat using a 1% concentration of a KOH catalyst and a 6:1 methanol-to-oil ratio at 60 °C [19]. Mata et al. has reported a 76.8% biodiesel yield by transesterification of chicken fat using a 0.8% KOH catalyst, 6:1 methanol:oil at 60 °C for 2 h [20]. The few variations in the results are probably due to the different fatty acid profiles of chicken fats and different enzyme sources.

Table 2. Optimized factors for biodiesel synthesis via enzymatic and chemical modes of transesterification of chicken fat oil.

Feedstock Oil	Catalysts/Biocatalysts	Reaction Time (Hours)	Reaction Temperature °C	$CH_3OH:Oil$ Molar Ratio	Catalyst's Concentration (%)	Biodiesel Yield (%)
WCFO	$Fe_3O_4\_PDA\_Lipase$	36	42	6:1	6	90.6
WCFO	*Aspergillus terreus* Lipase	36	35	6:1	1	78.4
WCFO	CH_3ONa	1.25	60	6:1	1	87.1
WCFO	KOH	1	60	6:1	1	84.8

2.5. ANOVA for Transesterification Data of WCFO

The influence of various reaction parameters such as linear factors, 1st order interactions and quadratic expressions on percentage biodiesel yield are described in the ANOVA table (Table 3). The terms (a)–(d) represents the quadratic models based on findings of $Fe_3O_4\_PDA\_Lipase$, *Aspergillus terreus* lipase, sodium methoxide and KOH-catalyzed transesterification of WCFO, respectively. The statistical analysis depicted that the linear term, A—reaction time, had a significant impact for models a, b and c on biodiesel yield ($p < 0.0001$, 0.0003 and 0.0003, respectively), which were <0.05, while for model d it was not significant. The linear term B–reaction temperature, showed *p* values of 0.1743, <0.0001, 0.0004 and <0.0001 for models a, b, c and d, respectively. The $Fe_3O_4\_PDA\_Lipase$ catalyzed transesterification was not affected significantly by temperature change in the selected range. Reaction temperature significantly affected the biodiesel yield for *Aspergillus terreus* lipase, which was temperature sensitive. The *p* values for the linear term C—$CH_3OH:Oil$, was <0.05 for model (a) and (c) but it was 0.1524 for model (b) and 0.7970 for model (d), which is >0.05. D—catalysts/biocatalysts concentration, was proven to have a significant effect on biodiesel yield for all the four models. A previous report on *Jatropha curcas* seed oil transesterification showed the significant impact of catalyst concentration, methanol-to-oil molar ratio, reaction temperature and reaction time on biodiesel yields [21] In case of Model (a), the 1st order interaction terms AC, AD and CD were found to be significant having *p*-values of 0.0007, 0.0001 and 0.0018, respectively, which were less than 0.05; however, for Model (b) only BD and CD were found significant. In case of Model (c), the 1st order interaction variables, i.e., AD and CD, were significant with *p*-values of 0.0065 and 0.0017 being less than 0.05; for Model (d), only AC 1st order interactions were imparting a significant impact on biodiesel yield with p-values lower than 0.05. Where the quadratic terms C^2 and D^2 were significant for Models (a) having a $p < 0.05$, for Model (b) the statistical significance was noted among the quadratic terms B^2, C^2 and D^2. In Model (c), B^2 and C^2 were significant, while in the case of Model (d), A^2 and D^2 were significantly affecting the biodiesel yield with $p < 0.05$.

Table 3. RSM-based ANOVA for transesterification of waste chicken fat oil (WCFO).

Source	Df	SS (MS) [a]	F-Value (*p* Value) [a]	SS (MS) [b]	F Value (*p* Value) [b]	SS (MS) [c]	F Value (*p* Value) [c]	SS (MS) [d]	F Value (*p* Value) [d]
Model	14	7213.87 (515.28)	70.99 (< 0.0001)	3018.69 (215.62)	63.48 (< 0.0001)	3159.24 (225.66)	41.95 (< 0.0001)	3806.96 (271.93)	41.87 (< 0.0001)
A—Reaction Time	1	310.24 (310.24)	42.74 (< 0.0001)	76.08 (76.08)	22.40 (0.0003)	114.01 (114.01)	21.19 (0.0003)	4.84 (4.84)	0.75 (0.4014)
B—Reaction Temperature	1	14.76 (14.76)	2.03 (0.1743)	1656.84 (1656.84)	487.79 (< 0.0001)	107.62 (107.62)	20.01 (0.0004)	268.40 (268.40)	41.33 (< 0.0001)
C—CH_3OH:Oil	1	188.55 (188.55)	25.97 (0.0001)	7.72 (7.72)	2.27 (0.1524)	376.29 (376.29)	69.95 (< 0.0001)	0.45 (0.45)	0.069 (0.7970)
D—Catalyst/Biocatalyst Concentration	1	3164.57 (3164.57)	435.97 (< 0.0001)	277.26 (277.26)	81.63 (< 0.0001)	460.40 (460.40)	85.59 (< 0.0001)	2557.47 (2557.47)	393.81 (< 0.0001)
AB	1	9.79 (9.79)	1.35 (0.2637)	15.05 (15.05)	4.43 (0.0525)	0.46 (0.46)	0.085 (0.7750)	0.45 (0.45)	0.070 (0.7952)
AC	1	129.39 (129.39)	17.83 (0.0007)	8.70 (8.70)	2.56 (0.1303)	16.61 (16.61)	3.09 (0.0993)	105.50 (105.50)	16.25 (0.0011)
AD	1	104.67 (104.67)	14.42 (0.0018)	0.83 (0.83)	0.25 (0.6277)	53.66 (53.66)	9.97 (0.0065)	0.13 (0.13)	0.021 (0.8880)
BC	1	3.05 (3.05)	0.42 (0.5264)	2.16 (2.16)	0.64 (0.4376)	1.27 (1.27)	0.24 (0.6346)	1.56 (1.56)	0.24 (0.6313)
BD	1	0.81 (0.81)	0.11 (0.7424)	222.82 (222.82)	65.60 (< 0.0001)	17.21 (17.21)	3.20 (0.0939)	1.17 (1.17)	0.18 (0.6774)
CD	1	193.91 (193.91)	26.71 (0.0001)	47.01 (47.01)	13.84 (0.0021)	77.88 (77.88)	14.48 (0.0017)	12.23 (12.23)	1.88 (0.1901)
A^2	1	2.49 (2.49)	0.34 (0.5670)	11.18 (11.18)	3.29 (0.0897)	5.19 (5.19)	0.096 (0.3416)	115.85 (115.85)	17.84 (0.0007)
B^2	1	1.19 (1.19)	0.16 (0.6913)	22.09 (22.09)	6.50 (0.0222)	79.02 (79.02)	14.69 (0.0016)	0.15 (0.15)	0.023 (0.8811)
C^2	1	524.04 (524.04)	72.19 (< 0.0001)	112.00 (112.00)	32.97 (< 0.0001)	1011.94 (1011.94)	188.12 (< 0.0001)	1.20 (1.20)	0.19 (0.6731)
D^2	1	96.56 (96.56)	13.30 (0.0024)	112.82 (112.82)	33.21 (< 0.0001)	2.48 (2.48)	0.46 (0.5078)	117.67 (117.67)	18.12 (0.0007)
Residual	15	108.88 (7.26)	-	50.95 (3.40)	-	80.69 (5.38)	-	97.41 (6.49)	-
Lack of Fit	10	96.75 (9.67)	3.99 (0.0701)	30.77 (5.13)	2.29 (0.1276)	58.46 (5.85)	1.31 (0.4021)	85.11 (8.51)	3.46 (0.0916)
Pure Error	5	12.13 (2.43)	-	20.18 (2.24)	-	22.24 (4.45)	-	12.30 (2.46)	-
Cor Total	29	7322.75	-	3069.64	-	3239.93	-	3904.37	-

Note: Fe_3O_4_PDA_Lipase (a), *Aspergillus terreus* lipase (b), sodium methoxide (c) and KOH (d). SS stands for sum of squares and MS is mean square.

The 3D surface plots of the significant 1st order interaction terms are presented in Figure 2. Figure 2a–c presents the significant 1st order interaction terms of Model (a). Figure 2a shows the 3D surface plot between the methanol-to-oil ratio and reaction time; it reveals that the yield increases with an increase in reaction time and methanol:oil, but further increases in the methanol-to-oil ratio resulted in a decreased biodiesel yield. The joint impact of time and concentration to increase the biodiesel yield is given in Figure 2b. Figure 2c presents the 3D plot between bio-catalyst/enzyme concentration and methanol:oil for Model (a). The plot shows that enzyme concentration directly increases the biodiesel yield but the increase of metahonl:oil after a specific limit decreases the biodiesel yield.

Figure 2d,e presents the 3D response surface plots for Model (b). Figure 2d reveals the relation between catalyst concentration and reaction temperature; an increase in temperature decreases the biodiesel yield probably due to the denaturation of free enzyme. Figure 2e presents possible impact of methanol:oil and bio-catalyst/enzyme concentration on biodiesel yield.

Figure 2f,g are the 3D plots of the significant 1st order interaction terms of Model (c). Figure 2f presents the relation between catalyst concentration and reaction time. It is observed that increase in both parameters increases the biodiesel yield. Figure 2g shows the relation between the methanol-to-oil ratio and catalyst concentration. The catalyst concentration increased the biodiesel yield but by further increasing the methanol-to-oil ratio up to certain level, however beyond optimal level, a decrease in biodiesel yield was noted.

Figure 2h presents the response of surface plot on the only significant interaction term of Model (d), which is between the reaction time and methanol-to-oil ratio. It showed that biodiesel yield increased with time but after a specific period any further increase in time was not effective.
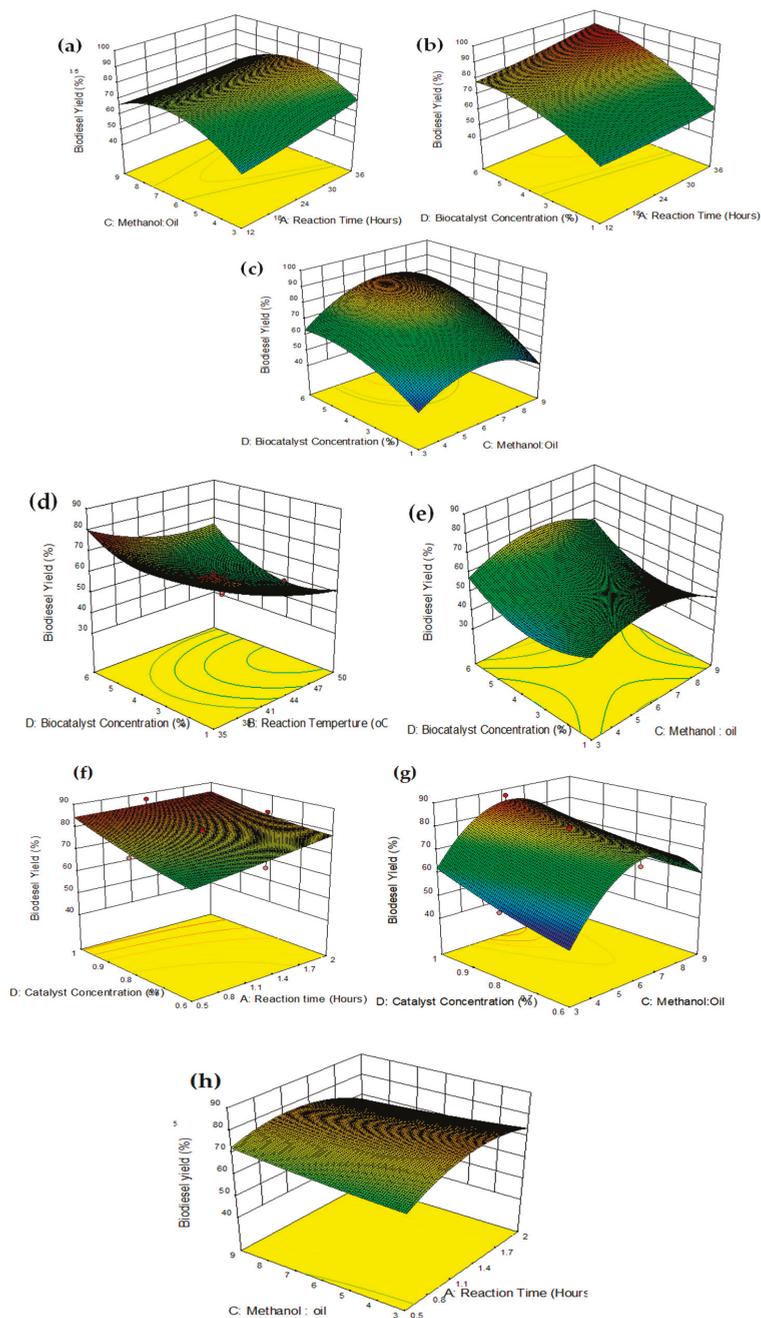
Figure 2. Response surface graphs for the significant 1st order interaction terms of Model (**a**) (A × C), (**b**) (A × D), (**c**) (C × D); Model (b), (**d**) (B × D)**,** (**e**) (C × D); Model (c), (**f**) (A × D)**,** (**g**) (C × D); Model (d), (**h**) (A × C).

2.6. FTIR Spectroscopic Analysis of Feedstock Oil, Biodiesel and Composition of Fatty Acid Methyl Esters

Asymmetric bending of the CH_3 group was observed in the region between 1425 and 1447 cm^{-1} and in the region ranging from 1188 to 1200 cm^{-1} which were basic characteristic peaks of biodiesel. While the C=O stretch vibrations observed in the region between 1700 and 1800 cm^{-1} and CH_2 asymmetric and symmetric stretching vibrations appeared at 2900–3100 cm^{-1} were present in FTIR spectra of both feedstock oil and synthesized biodiesel samples. However, signals in the 1390–1400 cm^{-1} region confirmed the O–CH_2 group and in the 1095–1101 cm^{-1} region defined the asymmetric axial stretching of O–CH_2–C for WCFO in FTIR spectra; however, these bands were absent in their respective biodiesel spectra. The above spectroscopic observations were according with the findings of a previous study [22]. The palmitic acid methyl ester (C16:0) 17.96%, stearic acid methyl ester (C18:0) 20.85%, oleic acid methyl ester (C18:1) 42.92% and linoleic acid methyl ester (C18:2) 16.54%, respectively, were the major FAMEs (Table 4). The current findings were found comparable with those reported by a previous study [16].

Table 4. The fatty acid methyl ester (ME) composition of the synthesized biodiesel.

Biodiesel Type	Palmitic Acid (C16:0) ME %	Stearic Acid (C18:0) ME %	Oleic Acid (C18:1, cis) ME %	Linoleic Acid (C18:2, cis)
WCFAO	17.96	20.85	42.92	16.54

2.7. Fuel Characteristics of WCFO Biodiesel

Fuel analysis plays a vital role in the evaluation of the manufactured biodiesel for its technical compatibility with conventional diesel. Fuel analysis of the understudy biodiesel samples was carried out in accordance with ASTM standard methods and the findings are mentioned below.

Kinematic viscosity is considered as one of the most significant fuel properties, as it is related to the resistance of flow that mainly occurs due to internal friction. If a biofuel contains greater values of kinematic viscosity, it will result in poor fuel atomization or incomplete combustion. Flash point and fuel volatility are inversely related to each other. Similarly, high values of cloud point generally result in problems such as fuel-line clogging. The kinematic viscosity (mm^2/s), flash point (°C), fire point (°C), pour point (°C) and cloud point (°C) values for WCFO biodiesel are given in Table 5. The mentioned fuel-quality parameters were found comparable with those from previous studies [19,20].

Table 5. Fuel characteristics of WCFO biodiesel.

Properties	WCFOB
Kinematic viscosity (mm^2/s) at 40 °C	4.9 ± 0.55
Flash point °C	171 ± 2.51
Fire point °C	187 ± 3.51
Pour point °C	3.0 ± 2.0
Cloud point °C	6.3 ± 2.37

3. Materials and Methods

Chemicals and reagents of analytical grade were utilized during study and were procured from Sigma-Aldrich (Saint Louis, MO, USA) and Merck (Darmstadt, Germany). Lipase from *Aspergillus terreus* was produced through fermentation at the Institute of Industrial Biotechnology, GC University, Lahore, Pakistan. The nano-biocatalyst (Fe_3O_4_PDA_Lipase) prepared and characterized in our previous work has been used as the immobilized lipase for biodiesel production from chicken fat oil [17]. The chicken fat was collected from the local market of Gujrat City, Pakistan.

3.1. Pre-Treatment of Feedstock

The collected waste chicken fat was heated at 100 °C to convert it into liquid. The liquid was then filtered to remove the solid waste. Since the chicken fat oil contains a high free fatty acid (FFA) content, alkaline catalysts are not suitable for un-treated chicken fat oil. In this case, acid esterification was used to reduce the acid value before alkaline transesterification. For this purpose, the chicken fat oil was taken in three neck flasks equipped with a thermometer and a glass condenser. The third neck was used to withdraw the sample. Oil was homogenized by heating and stirring at 600 rpm. Briefly, 50 mg concentrated sulfuric acid and 2.25 g methanol for each gram of FFA present in the oil was mixed in a beaker. The acid value of the sample was checked after specific time intervals by taking small aliquots. The process was carried out till the acid number reduced to the desired value. After completion of the process, the mixture was put in a separating funnel. Three layers were formed after some time. The top layer consisted of unreacted methanol and the lower layer was water while the middle layer was fatty acid methyl esters (FAMEs) and oil. The middle layer was collected for chemical transesterification [23]. The enzymatic transesterification, however, was done without pre-acid esterification. The collected oil was subjected to analysis for some basic parameters, including saponification value, acid value (AV) and peroxide value. The degree of unsaturation of feedstock oil was determined by iodine number. Similarly, the specific gravity was also determined along with density and the refractive index.

3.2. Experimental Design

The Central Composite Response Surface Methodology (CCRD–RSM) was used to evaluate the impact of the different conditions, namely A) reaction time, B) reaction temperature, C) CH_3OH: oil and D) the catalyst/biocatalyst's concentration on percentage biodiesel yield, for different transesterification routs using catalysts (KOH, CH_3ONa, free *Aspergillus terreus* lipase and Fe_3O_4_PDA_Lipase). The ranges of the selected parameters for the four models are presented in Table 6.

Table 6. The ranges of reaction parameters: reaction time, temperature, CH_3OH-to-oil ratio and enzyme concentration used for optimization studies of enzymatic and chemical transesterification.

Design	Catalysts/Biocatalysts	A—Reaction Time (h)	B—Reaction Temp (°C)	C—CH_3OH:oil	D—Catalyst/Biocatalyst Concentration (%) (with Respect to Substrate)
(a)	Fe_3O_4_PDA_Lipase	12–36	35–50	3:1–9:1	1–6
(b)	*Aspergillus terreus* lipase	12–36	35–50	3:1–9:1	1–6
(c)	CH_3ONa	0.5–2	40–60	3:1–9:1	0.6–1
(d)	KOH	0.5–2	40–60	3:1–9:1	0.6–1

In each case, thirty experiments were carried out as per CCRD factorial design. Chemically, this process was performed in three neck flasks equipped with a temperature regulator. A stirrer and reflux condenser were also attached with the flask. The reactions were accomplished at 500 RPM.

3.2.1. Chemical Transesterification

Briefly, 50 g of pre-treated waste chicken fat oil (WCFO) in a flat-bottom three-neck flask was subjected to pre-heating for moisture removal. The molecular weight of the understudy chicken fat oil (873.72 g/mol) was calculated as per a previously reported method [24]. The next step involved the mixing of known amounts of methanol and catalysts, respectively. The resultant mixture was transferred gently to the flask. The whole mixture present in the flask was kept for heating with stirring. According to the CCRD design, the reaction conditions were maintained. The reaction was allowed to proceed, and on completion the mixture was separated into two layers. The upper layer was taken and further processed to obtain the refined biodiesel.

3.2.2. Enzymatic Transesterification

For enzymatic biodiesel production, firstly waste chicken oil was mixed with methanol. A specific amount of immobilized lipase (Fe_3O_4_PDA_Lipase) for design (a) and free *Aspergillus terreus* lipase for design (b) was introduced to the oil/methanol solution and the reaction mixture was subjected to orbital shaking at 200 rpm with a 0.5% water content (with respect to oil), for a specific time period [25]. The CCRD was followed to set the alcohol to oil molar ratio, enzyme concentration, reaction temperature and time. After completion of the reaction, the glycerol was removed to obtain crude biodiesel, which was purified to get refined biodiesel. Magnetic nano-biocatalyst was separated from biodiesel and glycerol by magnetic decantation.

For the optimization studies, suitable statistical models based on experimental data were employed. Linear, 2FI, cubical and quadratic models were tested. The lack-of-fit test values, model significance (*p*-value), the R^2 and adjusted R^2 values provided the base to select most appropriate statistical model. Finally, the response surface graphs were utilized to check the influence of the studied reaction conditions on the yield of biodiesel.

3.3. Quantification and Characterization of Synthesized Biodiesel

For the FTIR spectroscopic study, a Carry660 FTIR spectrophotometer (Agilent Technologies, Santa Clara, CA, United States) was used and FTIR spectra were drawn over 400–4000 cm^{-1} scanning range.

The biodiesel from waste chicken oil was subjected to GC–MS analysis in order to evaluate the fatty acid methyl esters content (FAMEs). For this purpose, the GC–MS (QP 2010) instrument with a dB 5 column (Shimadzu, Japan) having diameter of 0.15 mm was used. The sample size (1.0 µL) was taken with a split ratio of 1:100, while a source of carrier gas, helium, was used having a 1.20 ml/min flow rate. The oven temperature was kept at 160.0 to 260.0 °C with a ramp rate of 4 °C per minute. The scanning of mass was done from 40.0 to 560.0 *m/z*. The detection of FAMEs was ascertained by comparing the relative retention time of each discrete FAMEs with reliable standards of FAMEs and by comparison with the NIST mass spectral library.

Fuel characteristics of the biodiesel were evaluated by some test experiments utilizing the ASTM standard procedure, i.e., cloud point (ASTM D 2500), viscosity (ASTM D 455), pour point (ASTM D 97) and flash point (ASTM D 93) [19]. The measurements were made in triplicate and the results were analyzed with the help of statistical tools.

4. Conclusions

The waste chicken fat oil was transformed into biodiesel by alkaline and enzymatic transesterifications. The reaction time, temperature, methanol:oil ratio and catalyst concentration were selected for the process optimization. Among all the catalysts and enzymes used, Fe_3O_4_PDA_Lipase-catalyzed transesterification of the studied feedstock oil was proved to be the most efficient to give maximum biodiesel yield. On the other hand, in case of chemical catalysis, CH_3ONa was proved to be better than KOH when chicken fat oil was used as the feedstock. FTIR spectroscopy and GC/MS characterization further confirmed biodiesel formation. The compositional profiles and fuel characteristics of the synthesized biodiesel showed a promising compatibility of WCFO as a potential candidate for biodiesel production for future fuel regimes.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Berchmans, H.J.; Hirata, S. Biodiesel production from crude *Jatropha curcas* L. seed oil with a high content of free fatty acids. *Bioresour. Technol.* **2008**, *99*, 1716–1721. [CrossRef] [PubMed]
2. Niculescu, R.; Clenci, A.; Iorga-Siman, V. Review on the use of diesel–biodiesel–alcohol blends in compression ignition engines. *Energies* **2019**, *12*, 1194. [CrossRef]
3. Rehan, M.; Gardy, J.; Demirbas, A.; Rashid, U.; Budzianowski, W.M.; Pant, D.; Nizami, A.S. Waste to biodiesel: A preliminary assessment for Saudi Arabia. *Bioresour. Technol.* **2018**, *250*, 17–25. [CrossRef]
4. Talebian-Kiakalaieh, A.; Amin, N.A.S.; Mazaheri, H. A review on novel processes of biodiesel production from waste cooking oil. *Appl. Energy* **2013**, *104*, 683–710. [CrossRef]
5. Jamil, F.; Al-Muhtaseb, A.H.; Al-Haj, L.; Al-Hinai, M.A.; Hellier, P.; Rashid, U. Optimization of oil extraction from waste "Date pits" for biodiesel production. *Energ. Conver. Manag.* **2016**, *117*, 264–272. [CrossRef]
6. Lotero, E.; Goodwin, J.G., Jr.; Bruce, D.A.; Suwannakarn, K.; Liu, Y.; Lopez, D.E. The catalysis of biodiesel synthesis. *Catalysis* **2006**, *19*, 41–83.
7. Gürü, M.; Koca, A.; Can, Ö.; Çınar, C.; Şahin, F. Biodiesel production from waste chicken fat based sources and evaluation with Mg based additive in a diesel engine. *Renew. Energy* **2010**, *35*, 637–643. [CrossRef]
8. Chiu, M.C.; Gioielli, L.A. Solid fat content of abdominal chicken fat, its stearins and its binary mixtures with bacon. *Food Sci. Technol.* **2002**, *22*, 151–157.
9. Gugule, S.; Fatimah, F.; Rampoh, Y. The utilization of chicken fat as alternative raw material for biodiesel synthesis. *Anim. Prod.* **2011**, *13*, 115–121.
10. Liu, K.S. Preparation of fatty acid methyl esters for gas-chromatographic analysis of lipids in biological materials. *J. Am. Oil Chem. Soc.* **1994**, *71*, 1179–1187. [CrossRef]
11. Coppini, M.; Magro, J.D.; Martello, R.; Valério, A.; Zenevicz, M.C.; Oliveira, D.D.; Oliveira, J.V. Production of methyl esters by enzymatic hydroesterification of chicken fat industrial residue. *Braz. J. Chem. Eng.* **2019**, *36*, 923–928. [CrossRef]
12. Christopher, L.P.; Kumar, H.; Zambare, V.P. Enzymatic biodiesel: Challenges and opportunities. *Appl. Energy* **2014**, *119*, 497–520. [CrossRef]
13. Dumri, K.; Hung Anh, D. Immobilization of lipase on silver nanoparticles via adhesive polydopamine for biodiesel production. *Enzym. Res.* **2014**, *2014*, 389739. [CrossRef]
14. Aryee, A.N.; Simpson, B.K.; Cue, R.I.; Phillip, L.E. Enzymatic transesterification of fats and oils from animal discards to fatty acid ethyl esters for potential fuel use. *Biomass Bioenergy* **2011**, *35*, 4149–4157. [CrossRef]
15. da Silva, J.R.P.; Nürnberg, A.J.; da Costa, F.P.; Zenevicz, M.C.; Lerin, L.A.; Zanetti, M.; Valério, A.; de Oliveira, J.V.; Ninow, J.L.; de Oliveira, D. Lipase NS40116 as catalyst for enzymatic transesterification of abdominal chicken fat as substrate. *Bioresour. Technol. Rep.* **2018**, *4*, 214–217. [CrossRef]
16. Feddern, V.; Kupski, L.; Cipolatti, E.P.; Giacobbo, G.; Mendes, G.L.; Badiale, F.E.; de Souza-Soares, L.A. Physico-chemical composition, fractionated glycerides and fatty acid profile of chicken skin fat. *Eur. J. Lipid Sci. Technol.* **2010**, *112*, 1277–1284. [CrossRef]
17. Touqeer, T.; Mumtaz, M.W.; Mukhtar, H.; Irfan, A.; Akram, S.; Shabbir, A.; Rashid, U.; Nehdi, I.A.; Choong, T.S.Y. Fe$_3$O$_4$-PDA-Lipase as surface functionalized nano biocatalyst for the production of biodiesel using waste cooking oil as feedstock: Characterization and process optimization. *Energies* **2019**, *13*, 177. [CrossRef]
18. Andrade, M.F.; Parussulo, A.L.; Netto, C.G.C.M.; Andrade, L.H.; Toma, H.E. Lipase immobilized on polydopamine-coated magnetite nanoparticles for biodiesel production from soybean oil. *Biofuel Res. J.* **2016**, *3*, 403–409. [CrossRef]
19. Alptekin, E.; Canakci, M. Optimization of pretreatment reaction for methyl ester production from chicken fat. *Fuel* **2010**, *89*, 4035–4039. [CrossRef]
20. Mata, T.M.; Cardoso, N.; Ornelas, M.; Neves, S.; Caetano, N.S. Evaluation of two purification methods of biodiesel from beef tallow, pork lard, and chicken fat. *Energy Fuels* **2011**, *25*, 4756–4762. [CrossRef]
21. Kamel, D.A.; Farag, H.A.; Amin, N.K.; Zatout, A.A.; Ali, R.M. Smart utilization of jatropha (*Jatropha curcas* Linnaeus) seeds for biodiesel production: Optimization and mechanism. *Ind. Crops. Prod.* **2018**, *111*, 407–413. [CrossRef]

22. Mumtaz, M.W.; Mukhtar, H.; Anwar, F.; Saari, N. RSM based optimization of chemical and enzymatic transesterification of palm oil: Biodiesel production and assessment of exhaust emission levels. *Sci. World J.* **2014**, *2014*. [CrossRef] [PubMed]

23. Chai, M.; Tu, Q.; Lu, M.; Yang, Y.J. Esterification pretreatment of free fatty acid in biodiesel production, from laboratory to industry. *Fuel Process. Technol.* **2014**, *125*, 106–113. [CrossRef]

24. Marulanda, V.F.; Anitescu, G.; Tavlarides, L.L. Investigations on supercritical transesterification of chicken fat for biodiesel production from low-cost lipid feedstocks. *J. Supercrit. Fluids* **2010**, *54*, 53–60. [CrossRef]

25. Katiyar, M.; Ali, A. Onepot lipase entrapment within silica particles to prepare a stable and reusable biocatalyst for transesterification. *J. Am.OilChem. Soc.* **2015**, *92*, 623–632. [CrossRef]

Challenges and Opportunities in Identifying and Characterising Keratinases for Value-Added Peptide Production

Juan Pinheiro De Oliveira Martinez [1], Guiqin Cai [1], Matthias Nachtschatt [1], Laura Navone [1], Zhanying Zhang [1], Karen Robins [1,2] and Robert Speight [1,*]

[1] Science and Engineering Faculty, Queensland University of Technology, Brisbane, QLD 4000, Australia; j5.martinez@hdr.qut.edu.au (J.P.D.O.M.); guiqin.cai@qut.edu.au (G.C.); m.nachtschatt@qut.edu.au (M.N.); laura.navone@qut.edu.au (L.N.); jan.zhang@qut.edu.au (Z.Z.); sustainbiotech@iinet.net.au (K.R.)

[2] Sustain Biotech, Sydney, NSW 2224, Australia

* Correspondence: robert.speight@qut.edu.au; Tel.: +61-7-3138-0373

Abstract: Keratins are important structural proteins produced by mammals, birds and reptiles. Keratins usually act as a protective barrier or a mechanical support. Millions of tonnes of keratin wastes and low value co-products are generated every year in the poultry, meat processing, leather and wool industries. Keratinases are proteases able to breakdown keratin providing a unique opportunity of hydrolysing keratin materials like mammalian hair, wool and feathers under mild conditions. These mild conditions ameliorate the problem of unwanted amino acid modification that usually occurs with thermochemical alternatives. Keratinase hydrolysis addresses the waste problem by producing valuable peptide mixes. Identifying keratinases is an inherent problem associated with the search for new enzymes due to the challenge of predicting protease substrate specificity. Here, we present a comprehensive review of twenty sequenced peptidases with keratinolytic activity from the serine protease and metalloprotease families. The review compares their biochemical activities and highlights the difficulties associated with the interpretation of these data. Potential applications of keratinases and keratin hydrolysates generated with these enzymes are also discussed. The review concludes with a critical discussion of the need for standardized assays and increased number of sequenced keratinases, which would allow a meaningful comparison of the biochemical traits, phylogeny and keratinase sequences. This deeper understanding would facilitate the search of the vast peptidase family sequence space for novel keratinases with industrial potential.

Keywords: keratinase; serine protease; metalloprotease; peptidase; keratin hydrolysis; keratin waste; valorisation; bioactive peptides

1. Introduction

Millions of tonnes of waste keratin are produced every year in the poultry, meat processing, leather and wool textile industries. The global poultry meat processing industry alone produces 40×10^6 tonnes of waste feathers annually [1]. With the transition away from the fossil fuel-centric economy to a sustainable circular economy, the valorisation of keratin materials addresses the waste problem and facilitates the integration of waste keratin into new value chains to enable a circular economy.

Traditionally, keratin waste has been sent to landfill or rendering, or used as fertilizer, feather meal or incinerated [2,3]. There is, however, an opportunity for livestock industries to produce higher value products from waste keratin. There are multiple thermochemical methods available to prepare hydrolysed keratin for various value-adding opportunities [4]. However, the use of peptidases with keratinolytic activity for keratin hydrolysis protects the integrity of the keratin amino acids in

most cases and allows control over the peptide size in the hydrolysate that is not readily achievable with other methods [5]. This degree of control allows the production of bespoke medical biomaterials, smart biocomposites, protein feed supplements with enhanced nutritional and bioactive properties as well as personal care products with enhanced functional and bioactive properties.

Identifying peptidases with keratinolytic activity is an inherent problem associated with the search for new enzymes. Keratinase activity however appears to be dependent on the accessibility of the keratin substrate to the enzyme [6,7]. Thermochemical or biochemical treatment of the keratin, with emphasis on the reduction of the disulphide bond and disruption of other important bonds involved in the structural stability of keratin like isopeptide, hydrogen and glycolytic bonds [6,8–10], appears to be the prerequisite for enzymatic hydrolysis. Sulphitolysis, which involves reduction of the disulphide bond in keratin, often acts synergistically with keratinases in nature [6,7]. Although destabilization of the keratin structure is a prerequisite for keratin hydrolysis, not all peptidases can hydrolyse keratin. Peptidases like trypsin, papain and pepsin cannot hydrolyse keratin as efficiently as peptidases with keratinolytic activity, even if the reduction of disulphide bond has already occurred [11]. The elucidation of the unique characteristics of peptidases with keratinolytic activity that differentiate them from the other peptidases, would be an important breakthrough in the search for new and robust keratinases for the valorisation of keratin waste.

This paper reviews twenty sequenced peptidases with keratinolytic activity from the serine protease and metalloprotease families by comparing their biochemical characteristics and will highlight the difficulties associated with the interpretation of these data.

2. Keratin: A Complex and Strong Structure

Keratins are important structural proteins produced by vertebrate epithelia that have various physiological function. Keratins can act as a protective barrier to water, against infection or cushion tissue from mechanical impact. The two main types of keratins proteins are α-keratin and β-keratins. These two types are further divided into acidic or basic, soft or hard, and have different molecular weights [4,9,12,13]. The following section describes the complexity of the keratin structure, which provides insight into the resistance of keratin to hydrolysis. This review will concentrate on hard α-keratin and β-keratin, γ-keratins and the keratin-associated proteins, which are common to mammalian hair, bristles, wool, hooves, horns and feathers.

α-Keratin has an α-helix structure, which is stabilized by hydrogen bonding and the presence of multiple cysteines forming disulphide bridges. α-Keratin is characterized by a lower sulphur content compared to other keratins and a molecular mass of 60–80 kDa [4]. Hard α-keratin is the major protein of mammalian fibres, nails, hooves and horns. In contrast, hard β-keratins are characteristic of the hard, cornified epidermis of reptiles and birds, e.g., feathers, claws and scales, and have a twisted β-sheet-like structure. They also form the major component of the fibre cuticle. The β-keratin pleated sheets consist of β-strands, which are laterally packed and can have a parallel or antiparallel orientation. The β-sheets are held together by hydrogen bonds and the planar nature of the peptide bond, which results in the stable pleated β-sheet [13]. β-Keratins have a molecular mass of 10–22 kDa. A third type of keratin, γ-keratin, is a globular protein with a high sulphur content and a molecular weight of about 15 kDa. This keratin, along with keratin-associated proteins, form the matrix between the microfibrils and microfibrils of the fibre cortex of mammalian fibres and stabilize the structure of the cortex via extensive disulphide bridge formation.

The complex structural organization of all mammalian fibres is very similar [8]. The hair fibre consists of an outermost cuticle layer, which is composed of overlapping flattened scale-like cells that form a protective sheath around the cortex [8]. The major protein of the fibre cuticle is β-keratin [4]. The cortex is composed of hard α-keratin intermediate filaments embedded in a sulphur-rich matrix. These filaments surround the medulla when present, as is the case for coarser fibres. The cell membrane complex binds the cuticle and cortical cells.

The cuticle layer is laminated and consists of the following layers—the cuticle filament-associated surface membrane, the cystine-rich exocuticular *a*-layer, the lower exocuticle and the endocuticle, which contains only low levels of sulphur-containing amino acids and constitutes the inner lining of the cuticle [8]. The outermost layer of the cuticle provides a hydrophobic barrier, which protects the fibre surface from water and chemical compounds. This cuticle filament-associated surface membrane is 2–7 nm thick and composed of highly cross-linked proteins and lipids. The major fatty acid of the cuticle surface lipids found in human and animal hair is 18-methyleicosanoic acid [14]. It is covalently linked to the protein matrix below by a thioester linkage and the protein matrix is cross-linked by isopeptide bonds [15]. An isopeptide bond results from the transglutaminase-catalysed formation of an amide bond between the amino acid side chains of the amino acid residues in the keratin protein, for example, lysine and glutamine [9].

The cortical cells are assembled as keratin intermediate filaments and have a diameter of 7–8 nm in all mammalian fibres [8]. These intermediate filaments form ordered aggregates or microfibrils and macrofibrils depending on species and function (Figure 1). The hard α-keratin intermediate filaments are assembled from tetramers, a pair of laterally aligned and antiparallel dimeric molecules. On average, keratin intermediate filaments contain eight tetramers. In the case of wool, the cortex region is composed of an orthocortex and paracortex with different intermediate filament/matrix packing. The proportion of ortho- and paracortex in the wool fibre determines the degree of crimping [13].
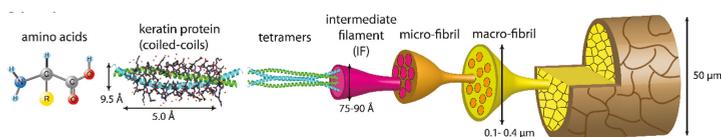
Figure 1. Structure of keratin. Adapted from work in [12] under the Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0/deed.en).

Keratin peptide heterodimers are formed when a type I (acidic) polypeptide chain and a type II (basic) polypeptide chain align in parallel. Each polypeptide chain is composed of a central α-helical region (about 46 nm in length) with non-helical head and tail domains [13]. The head and tail domains are rich in cysteine, glycine and tyrosine amino acids. Disulphide and isopeptide bonds are formed with other keratin intermediate filaments, cysteine-rich matrix proteins and keratin associated proteins, which stabilize the fibre [8–10]. The disulphide bonds along with the N-acetyl glucosamine-glycosylated serine and threonine in the head and tail domains also stabilize the heterodimers [6].

3. Thermochemical Methods of Keratin Degradation

There are multiple thermochemical methods available to prepare hydrolysed keratin for various value-adding opportunities, with specific processes chosen depending on the end-use [4]. Thermochemical methods include solubilization of keratin in organic solvents, ionic liquids or by hydrothermal methods; oxidation or reduction of the disulphide bridges; disruption of the hydrogen bonds with compounds like urea; and acid or base hydrolysis.

The composition of the final hydrolysate will depend on the method used to hydrolyse the keratin. Some of the thermochemical processes result in a hydrolysate containing a highly diverse mix of keratin-derived peptides and free amino acids and others are more specific. However, in most cases, the amino acid composition is modified. The processes and hydrolysate products will be described in more detail in the following section.

After solubilization of keratin with solvents like with *N,N*-dimethylformamide or dimethyl sulfoxide, precipitation is required with acetone and drying to produce a powder of keratin [16]. The major drawback of this method is the use of large quantities of solvents, which need to be recycled or incinerated. Solubilization can also be achieved with ionic liquids. Xie et al. used the ionic liquid, 1-butyl-3-methylimidazolium chloride for the solubilization of wool keratin, which disrupted

the hydrogen bonds in the keratin macromolecules [17]. The keratin peptides were precipitated from the resulting hydrolysate with methanol. Ionic liquids are more expensive than traditional solvents and extraction of the keratins from the ionic liquid can be difficult.

Hydrothermal treatment is usually carried out at temperatures of 80–140 °C and steam pressures of 10–15 psi. Acid or base can be added to speed up the process of solubilization [18]. Under conditions of high temperature and pressure, the thermally unstable amino acids, glutamine and asparagine are degraded [19]. If base is added to this process then lysine, methionine and tryptophan are also destroyed [20,21]. Modified amino acids, lysinoalanine and lanthionine are also formed from lysine and cystine, respectively. Heating of proteins leads to a degree of racemization of the free and bound L-amino acids [22–24].

Reduction with reducing agents like thioglycolate [4], dithiothreitol [25], 2-mercaptoethanol [26], sodium sulphite [27], bisulphites [28] or cysteine [29] combined with high concentrations of compounds like urea, thiourea or surfactants, which disrupt the hydrogen bonds stabilising the keratin structure, results in the production of kerateine [30]. Kerateine contains cysteine thiol and cysteine sulfonate in place of the disulphide bonds. Kerateine is less soluble in water and can be re-cross-linked if exposed to an oxidant [4].

The microstructure of wool keratin after treatment for 4 h at 65 °C with 2-mercaptoethanol, EDTA, high concentrations of urea and pH 9 was investigated by Cardamone [31]. Analysis of the hydrolysate revealed a defined mixture of microfibrillar and intermediate filaments. This mixture of subunits was suitable for producing self-assembling biomaterials.

Oxidation of keratin by oxidants like peracetic acid [32] or peroxycarboximidic acid [33] leads to the formation of keratose. Keratose contains sulfonic acid groups and cysteic acid instead of the disulphide bonds [4]. These keratoses are hydroscopic, water soluble and the disulphide bridges cannot spontaneously re-form under oxidative conditions. Keratoses are not as stable as kerateines.

Oxidative sulphitolysis has been patented and commercialized to produce three functional keratin protein and peptide products. These products are based on S-sulphonated keratin intermediate filaments, S-sulphonated keratin high-sulphur proteins and keratin peptides [34]. The process aims at maintaining the structural integrity of the keratin proteins. The cystine groups in the wool keratin are converted to S-sulfocysteine using sodium sulphite or sodium metabisulfite and then oxidized with cupraammonium hydroxide. The intermediate filaments and peptides can undergo crosslinking by reductive desulfonation of the cysteines in the filaments and peptides and subsequent reformation of the intermolecular disulphide bonds.

One of the disadvantages of alkaline hydrolysis of keratins is the modification or degradation of amino acids (Table 1). Alkaline hydrolysis of keratins at higher temperatures results in the degradation of the thermally unstable amino acids, asparagine, glutamine, arginine, serine, threonine and cysteine [5]. Lysinoalanine and 8-aminoalanine are formed under alkaline conditions [35,36]. Another modification that occurs is the racemization of free or bound L-amino acids to the D-enantiomers [23,37,38]. Free amino acids racemize ten times slower than bound amino acids [24]. Following, for example, prolonged treatment of wool keratin at 70 °C and pH 9–11, lanthionyl residues [31] and dehydroalanine [39] are formed from cystine. Cystine and hydroxy amino acids were destroyed if the alkaline treatment was performed in the presence of reducing agents [40].

Table 1. Amino acid modification during alkaline treatment.

Amino Acid	Degradation Products	Reference
Asparagine	Aspartate, ammonia	[41]
Glutamine	Glutamate, ammonia	[41]
Arginine	Ornithine, citrulline, 3-aminopiperidin-2-one	[42]
Serine	Glycine, alanine, oxalic acid, lactic acid, ammonia	[43]
Threonine	Glycine, alanine, α-aminobutyric acid, ammonia	[44]
Cysteine	Pyruvic acid, sodium sulfide, ammonia	[45]
* Cystine, lysine, arginine	Lanthionine, lysinoalanine, ornithinalanine; * dehydroalanine [39]	[35,36]
L-amino acids	D-amino acids	[37]

Note: * Dehydroalanine is probably formed from the cleavage of the C-S bond in cystine.

Acid hydrolysis of keratins leads to the loss of some amino acids like serine, threonine, tyrosine and cystine and the conversion of asparagine, glutamine, methionine and tryptophan into other compounds ([5,19,46] Table 2). Polypeptides, resulting from the acid hydrolysis of keratin, have a more amorphous structure than alkaline hydrolysates, because most of the hydrogen bonds are broken during this process [47]. A typical acid hydrolysis of keratin uses hydrochloric acid [48,49] or sulphuric acid [50] at high temperatures.

Zhang et al. showed that acid hydrolysis was not as effective as other treatments mentioned above [49]. Wool keratin was hydrolysed with 4M hydrochloric acid at 95 °C for 24 h, resulting in 33% solubilization of the wool keratin. Increasing the treatment time had no effect on the yield, suggesting that there is a recalcitrant portion of the keratin resistant to acid hydrolysis.

Thermochemical methods offer cheap and versatile processes for hydrolysing keratin for a variety of applications. However, the use of harsh chemicals and conditions, the lack of ability to control the processes in most cases and the often unfavourable modification of the amino acids or peptides present environmental problems and peptide mixes that would be unsuitable for some applications. Using enzymes working under mild conditions to catalyse the hydrolysis offers a favourable alternative.

Table 2. Amino acid modification during acid treatment.

Amino Acid	Degradation Products	Reference
Asparagine	Aspartate, ammonia	[19,41]
Glutamine	Glutamate, ammonia	[19,41]
Methionine	Methionine sulfoxide	[19]
Tryptophan	Oxindolylalanine, dioxindolylalanine	[46]

4. Microbial Degradation of Keratin

The first peptidases with keratinolytic activity were found in *Bacillus* sp. and *Streptomyces* sp. and belong to the serine peptidase family [51]. The ability to degrade keratin is widespread and has been identified in bacteria and fungi [4,52]. Keratin-degrading microorganisms have been isolated from many sources like skin, feathers, hair, nails, soil, geothermal hot stream and wastewater, which is reflected in the optimum pH and temperature of the keratinase activity of these microorganisms. The pH optimums of keratinases range from pH 5.5 for the fungal keratinase from *Trichophyton mentagrophytes* [53] to pH 12.5 for the keratinase from *Brevibacillus* sp. AS-S10-11 [54]. Although, temperature optimums vary from 30 °C for the keratinase from *Brevibacterium luteolum* [55] to 100 °C for the keratinase from *Fervidobacterium islandicum* AW-1 [56].

Publications from 2018 and 2019 report the isolation of diverse species of bacteria like *Streptomyces* sp. [57], *Aeromonas hydrophila* FB3 [58], *Pseudomonas putida* KT2440 [59] and *Serratia marcescens* EGD-HP20 [60,61] with keratinolytic activity. However, the number of *Bacillus* strains with keratinolytic activity prevailed over any other genus of bacteria [62–86]. Valorisation of waste feathers [5,65,66,87] and the replacement of the traditional, highly polluting hide dehairing step used in the leather industry with a more environmentally friendly enzymatic step using keratinases [2,55,79,88] were the dominant themes of these papers.

Despite the interest in the enzymatic hydrolysis of keratin, mechanisms of keratin degradation in microorganisms are not fully understood. There is evidence that microbial degradation of keratin proceeds via a consortium of enzymes (Figure 2 [6,7,89]).

Disruption of the keratin structure is an essential step in the breakdown of keratin by keratinases. Various mechanisms have been suggested for fungal systems. Disulphide bond reductases and the intracellular cysteine dioxygenase can break the structure-stabilizing disulphide bridges in keratin [6,7,90]. Cysteine dioxygenase in conjunction with aspartate aminotransferase produces the reducing agent, sulphite, from cysteine, which is secreted into the surroundings and contributes to the chemical reduction of the disulphide bond. The reduction of the disulphide bonds aids access of the endoproteases (serine protease family), exoproteases (metalloprotease family) and

oligopeptidase (metalloprotease family) to the keratin fibres or feathers. It has also been found that the membrane-bound redox system of the cell can cleave the disulphide bonds in keratin. The mechanical pressure exerted by fungal mycelia penetrating the keratin structure can also contribute to the disruption of this structure, facilitating access of the keratinase to the substrate. In nature, these mechanisms act synergistically with keratinases and speed up the degradation of keratin. Auxiliary proteins, like lytic polysaccharide monooxygenases (LPMOs), have been found associated with keratin degradation [6]. It is thought that they contribute to α- and β-keratin degradation. Until now LPMOs were thought to be associated with cellulose, chitin, hemicellulose and starch degradation only. It is possible that these enzymes hydrolyse the glycolytic bond between N-acetylglucosamine and serine and threonine in the head and tail region of the intermediate filaments, which contributes to the destabilization of the keratin structure.
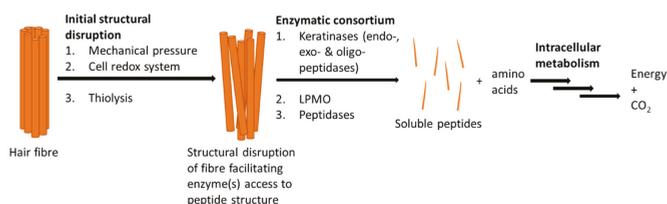
Figure 2. Possible mechanisms for microbial degradation of keratin (LPMO = lytic polysaccharide monooxygenase).

However, examples of peptidases with keratinolytic activity that do not need the assistance of disulphide reducing enzymes or agents have also been reported. Pillai et al. isolated a serine protease from *Bacillus subtilis* P13 with reductase and keratinase activities [91]. The isolated enzyme was able to decompose feathers and dehair hides.

He et al. analysed the enzyme consortium involved in the hydrolysis of feathers by a specific strain of *Bacillus subtilis* and identified four of the enzymes involved in keratin hydrolysis [74]: a serine protease with keratinase and disulphide bond-reducing activity; a peptidase T; a γ-glutamyltransferase, which generates a free cysteinyl group from glutathionine; and a cystathionine γ-synthase, which catalyses the production of L-cystathionine from homoserine ester and cysteine. The L-cystathionine is further converted to methionine and ammonia is released.

5. Characterisation and Comparison of Keratinases from S1, S8 and M4 Peptidase Families

Many articles characterising organisms capable of degrading keratin and their possible industrial applications have been published. Yet, there are few articles that report enzyme sequences and investigate the molecular and biochemical characteristics of the enzymes produced by these organisms [92,93]. The first paper that explored the molecular aspects of a keratinase produced by *Bacillus licheniformis* was published by Lin et al., 1995. Since then, more than 40 keratinases have been sequenced. To date, peptidases with keratinolytic activity from six different peptidase families have been identified: S1, S8, M4, M5, M14 and M28. Most of the characterized keratinases are produced by *Bacilli* and are members of the S8 serine peptidase family. There are currently over 127,000 peptidase sequences from the S1 (70919), S8 (38270), M4 (6403), M5 (145), M14 (11202) and M28 (904) families deposited on the MEROPS peptidases database. These 127,000 peptidase sequences represent an enormous unmined potential for the discovery of new peptidases with keratinolytic activity if the requisite properties of a peptidase with keratinolytic activity can be identified.

The S1 family sequences, when pairwise aligned, show a minimum value of 27.27% and a maximum of 97.22% identity, with an average of 61.48% for the four available sequences. The S8 family has a minimum of 13.69% and a maximum of 99.72% identity, with an average of 63.29% for the 13 available sequences, and the M4 family has 25.56% identity between the two available sequences. Although many of the characterized enzymes have been produced by the native unmodified organism [94–98], several

examples involve heterologous expression. Different organisms have been used for recombinant production, including yeast such as *Komagataella Pastoris* (*Pichia Pastoris*) [99] and bacteria such as *Escherichia coli* [100–110] and *Streptomyces lividans* [111]. Including the pre-pro-domains with the catalytic domain in heterologous systems have been shown to maintain enzyme activity and secretion [99,102,107,110] and inclusion of C-domains, when present, is important for substrate binding and recognition [105]. Replacing the native signal peptide for the *E. coli* signal peptide when expressing in *E. coli* has also led to higher levels of expression [101].

In this section the biochemical data of twenty sequenced peptidases with keratinolytic activity from the S1 and S8 peptidase families (serine proteases) and the M4 peptidase family (metalloprotease) are compared (Table 3). Difficulties associated with the interpretation of these data are also highlighted. The selection is based on the availability of sequence and biochemical data. The M5, M14 and M28 peptidase families were excluded because each family had only one biochemically characterized example with full sequence data available.

Table 3. Keratinolytic microorganisms and their keratinases from the S1, S8 and M4 keratinases selected for this study.

Organism	Strain	Keratinase Name	Accession No. [1]	Reference
S1A Peptidases				
Actinomadura viridilutea	DZ50	KERDZ	KU550701	[94]
Actinomadura keratinilytica	Cpt29	KERAK-29	ASU91959	[95]
Streptomyces fradiae	Var. k11	SFP2	AJ784940	[99]
Nocardiopsis sp.	TOA-1	NAPase	AY151208	[111]
S8A Peptidases				
Bacillus circulans	DZ100	SAPDZ	AGN91700	[100]
Bacillus licheniformis	RPk	KerRP	EU502844	[96]
Stenotrophomonas maltophilia	BBE11-1	KerSMD	KC814180	[101]
Stenotrophomonas maltophilia	BBE11-1	KerSMF	KC763971	[101]
Bacillus pumilus	A1	KerA1	ACM47735	[97]
Bacillus pumilus	CBS	SAPB	CAO03040	[102]
Bacillus pumilus	KS12	rK$_{27}$	HM219183	[103]
Bacillus tequilensis	Q7	KerQ7	AKN20219	[104]
Bacillus cereus	DCUW	Vpr	ACC94305	[105,112]
Bacillus altitudinis	RBDV1	KBALT	APZ77034	[63]
Thermoactinomyces sp.	YT06	YT06 Protease	WP_037995056	[98]
Thermoactinomyces sp.	CDF	Protease C2	ADD51544	[106]
Meiothermus taiwanensis	WR-220	rMtaKer	5WSL	[107]
Brevibacillus sp.	WF146	WF146 Protease	AAQ82911	[108]
M4 Peptidases				
Geobacillus stearothermophilus	AD-11	RecGEOker	KJ783444	[109]
Pseudomonas aeruginosa	KS-1	KerP	HM452163	[110]

Note: [1] NCBI GenBank nucleotide accession number.

5.1. S1, S8 and M4 Peptidase Families

The S1 family is the largest family of serine proteases. The active site of S1 peptidases contains the catalytic triad, His, Asp and Ser. All enzymes characterized in this family are endopeptidases. The four peptidases in Table 3 belong to the S1A family represented by chymotrypsin as the type-example. The hydrophobic amino acid at the P1 site determines the specificity of these peptidases [113,114].

The S8 family is currently the second largest serine protease family and the most widely characterized to date [114,115]. Most of the keratinases are found in the subfamily S8A including the 14 keratinases in Table 3. They are represented by subtilisin as the type-example. Their active site contains the catalytic triad of Asp, His and Ser. In general, these enzymes are endopeptidases [116], active between neutral and moderately alkaline pH and many are thermostable [117]. Most enzymes in this family are not specific, usually cleaving after a hydrophobic residue in the peptide substrate [114,117]. S1 and S8 families are examples of convergent evolution as they catalyse the same reaction but have no sequence homology. Two calcium-binding sites contribute to thermal stability in many members of these families [114,117].

Two keratinases in Table 3 belong to the M4 family. They are characterized by a catalytic zinc ion tetrahedrally coordinated in the active site by a histidine and glutamate present in a HEXXH motif, another glutamate residue and water [118]. Most members of this family are endopeptidases and active at neutral pH. The preferred cleavage site occurs at a hydrophobic residue followed by leucine, phenylalanine, isoleucine or valine. These peptidases are stabilized by Ca^{2+} [119].

Independent of their families, keratinases usually cleave aromatic and hydrophobic amino acid residues at the P1 position. Keratins are composed of 50 to 60% aromatic and hydrophobic residues, which could partially explain the keratinase specificity for keratin [120–122]. Most of these peptidases are stabilized by divalent cations like Ca^{2+} and are extracellular [119,123,124].

5.2. Problems Associated with Keratinase Assays

There are several issues with the current methods used to characterize keratinases. The assays are not standardized in the literature in terms of reaction conditions and substrates. The most common method used to measure keratinase activity is a colorimetric assay that uses the commercially available derivative of wool, keratin azure [125] or azokeratin (sulfanilic acid-azokeratin [126]). However, batch variability and the fact that the chromogenic agents are only bound to the outer portion of the substrate compromises reproducibility. Quantification of the soluble peptides generated by hydrolysis of keratin has also been used to determine the effectiveness of keratinases on keratin substrates. Common quantification methods used are Bradford [95,127], Lowry [128,129] or measurement of absorption at 280 nm [106,111] (see Table 4). Each of these methods have several limitations. The Coomassie Blue dye used in the Bradford method preferentially reacts with arginine and lysine in the peptide mix and alkaline pH and detergents interfere with the reaction. The Folin–Ciocalteu dye used in the Lowry method oxidizes the aromatic amino acids residues in the protein and is affected by reducing agents. Only tyrosine, tryptophan and cysteine absorb at 280 nm and other compounds like DNA in the solution can interfere with the measurement [130]. The simplest and probably most accurate method for quantifying keratinase activity is the measurement of weight loss when the insoluble keratin substrates like mammalian hair fibres, feathers or wool are solubilised through hydrolysis [96,127].

Table 4 describes a variety of substrates that have been used to assay keratinase activity in the literature. The substrates that were used include keratin azure (wool), keratin powder, soluble keratin, keratin (undefined), feathers and feather meal powder. It was not possible to ascertain the source and integrity of most of these keratin substrates from the papers. The pretreatment of these substrates is also an important aspect in determining the keratinase activity. Keratin powder and solubilized keratin were generally obtained from commercial sources; however, the sources and preparation were not described. Pretreatments like autoclaving and milling [103,107], or treating with solvents at high temperatures [106], are known methods for keratin powder preparation from the literature. In the case of the rK_{27} keratinase, the feather powder used in the assay was autoclaved and dried at 60 °C [103]. These preparation methods, as already described in Section 3, would compromise the keratin structure. The keratinases, KerRP [96], Ker-A1 [97] and SAPB [102] were assayed on keratins of unknown source. In the WF146 protease assay, the feather substrate was washed with ethanol prior to use in the assay, which would likely remove the protective lipid layer [108].

Co-treatment can also affect the integrity of the keratin substrates during enzymatic hydrolysis [125,131,132]. Except for KerQ7 [104], all assays in Table 4 were carried under alkaline conditions between pH 8 and pH 12.5 and temperatures ranging from 50 to 80 °C. These conditions would most likely contribute to the weakening the keratin structure. Keratinase assays with SAPDZ [100], KerQ7 [104], KERDZ [94], and KERAK-29 [95] were supplemented with the divalent cations Ca^{2+} or $Mn2^{+}$. Divalent cations are known to stabilize serine proteases [114,117]. Other keratinase studies added reducing agents, like β-mercaptoethanol (protease C2 [106], WF146 protease [108]) or dithiothreitol (SFP2 [99]) to the assay mixture. Reducing agents are known to break the disulphide bond leading to disruption of the keratin.

Table 4. Keratinase pH and temperature optimums of the selected S1, S8 and M4 keratinases with associated assay conditions. See text for further details of the assays.

Protein	* pH	* Temp (°C)	Conditions	PT	CT
S1A Peptidases					
KERDZ	11	80	10 g/L keratin azure, 50 mM bicarbonate-NaOH buffer, pH 11 mixed 1:1 with the enzyme, 30 min, 80 °C, 200 rpm (Abs$_{595nm}$).	-	2 mM CaCl$_2$
KERAK-29	10	70	1 mL of 10 g/L keratin azure, 100 mM Glycine-NaOH buffer mixed 1:1 with the enzyme, pH 10, 20 min, 70 °C (Abs$_{595nm}$).	-	5 mM MnSO$_4$
SFP2	10	60	5 mg keratin azure, 50 mM Tris-HCl, pH 8.5, 1 h, 37 °C (Abs$_{595nm}$).	-	10 mM DTT
NAPase	12.5	60	60 mg wool keratin powder, Glycine-NaOH, pH 10 or 50 mM KCl-NaOH, pH 12.5, 30 °C, 2 h (Abs$_{280nm}$).	Not specified	-
S8A Peptidases					
SAPDZ	12.5	85	10 g/L keratin azure, 100 mM KClNaOH, 250 rpm, 20 min incubation, 85 °C (Abs$_{595nm}$).	-	5 mM CaCl$_2$
KerRP	9 (11)	60 (65–70)	0.8% w/v keratin diluted 1:1 in enzyme, 1 h incubation, 60 °C (Abs$_{280nm}$).	Not specified	-
KerSMD	8	60	1% w/v soluble keratin, 50 mM Gly-NaOH, 20 min, 50 °C (Folin–Ciocalteu).	Not specified	-
KerSMF	9	60	1% w/v soluble keratin, 50 mM Gly-NaOH, 20 min, 50 °C (Folin–Ciocalteu).	Not specified	-
KerA1	9 (10)	60 (65)	0.8% w/v keratin diluted 1:1 in enzyme solution, 1h, 50 °C (Abs$_{280nm}$).	Not specified	-
SAPB	10.6	65	1% keratin w/v, on 100 mM glycine-NaOH Buffer, pH 10.6, 30 min, 55 °C. 2 mM CaCl$_2$ supplemented.	Not specified	-
rK$_{27}$	9	70	20 mg feather powder, Gly-NaOH 50 mM, 1 h (Abs$_{280nm}$).	Washed with Triton X-100 (1%), rinsed with water, autoclaved, dried in an oven at 60 °C for 1 h, milled then sieved with 2 mm pore size.	-
KerQ7	7	30	10 g/L keratin azure, 50 mM HEPES buffer, 30 min, 200 rpm (Abs$_{595nm}$).	-	1 mM CaCl$_2$
Vpr	8.5	50	2% w/v chopped feather keratin, 50 °C, 15 min, pH 7.5.	-	-
KBALT	8	85	5 mg keratin azure, 0.8 mL buffer, 15 min incubation, pH 6 to 12, 25 to 95 °C (Abs$_{595nm}$).	-	-
YT06 protease	8–9	65	1% soluble keratin, 50 mM Gly-NaOH, pH 9, 20 min (Folin–Ciocalteu).	Not specified	-
Protease C2	11	60–80	5% keratin powder, 50 mM Tris-HCl pH 8, 60 min 60 °C (Abs$_{280nm}$).	100 °C incubation in DMSO for 2 h. Protein precipitated with acetone 2:1 v/v [133]	0.5% β-ME
rMtaKer	10	65	1% feather powder on 50 mM HEPES, pH8.0, 25–95 °C. Supplemented with 10 mM CaCl$_2$, 150 mM NaCl (Ninhydrin).	Chicken feathers rinsed, air-dried, ground by ball mill.	-
WF146 protease	-	80	10 mg of feathers, 50 °C or 80 °C, 1 ml Tris-HCl 50 mM buffer, pH 8.0, 10 mM CaCl$_2$, multiple time points from 0 to 20 h (Abs$_{280nm}$).	70 Ethanol wash, rinse water, dry, cut 2–3 mm long	1% β-ME
M4 Peptidases					
RecGEOker	9	60	4 mg keratin azure, 50 mM Tris-HCl, pH 7.8, 1 h (Wool-Folin–Ciocalteu; Abs$_{595nm}$).	-	-
KerP	9	50	20 mg chicken feathers, Glycine-NaOH buffer, pH 10, 60 °C, 60 min (Abs$_{280nm}$)	-	-

Note: Source organism, accession numbers and references can be found in Table 3; * in some cases pH and temperature optimums were determined on both casein and keratin substrates. The casein optimums are in brackets; Temp = temperature; PT = pretreatment; CT = co-treatment; β-ME = β-mercaptoethanol; DTT = dithiothreitol. Quantification methods, where available, are in brackets after the assay condition description.

The challenges with the keratinase assays discussed above highlight the need for standardized assays and substrates used to test keratinases and the challenges faced in attempting to compare and analyse data from the literature when the assays are not comparable.

5.3. The Effect of Additives on Selected S1, S8 and M4 Keratinases

Various additives were tested on the selected S1, S8 and M4 keratinases-cationic, anionic and neutral detergents, oxidizing agents, reducing agents, mono- and divalent metals. Table 5 contains a summary of additives that had a positive effect on keratinase activity. A positive effect was defined as ≥ 110% activity compared to the control without additive. Some of the papers used keratin as the assay substrate, some used casein and in some cases, both were tested.

Table 5. Influence of additives on the activity of selected S1, S8 and M4 keratinases. Numbers in brackets correspond to the % activity compared to 100% in the absence of the additive.

Protein	Metal ions (%)		Detergents (%)		Reducing Agents (%)	Solvents/Others (%)
S1A Peptidases						
KERDZ	Ca^{2+} (270) Mg^{2+} (180) Fe^{2+} (145)					
KERAK-29	Ca^{2+} (150) Mg^{2+} (110) Mn^{2+} (210)		Zwittergent (114) Tween-20 (130) Triton X-100 (132) Tween-80 (145) TTAB (116) CHAPS (140)	Sulfobetaine (135) LAS (118) SDS (115) CTAB (110)	β-ME (102)	H_2O_2 (170)
SFP2	Cu^{2+} (149) Ni^{2+} (116)				DTT (278) β-ME (235)	
NAPase						Isopropanol (130)
S8A Peptidases						
SAPDZ	Ca^{2+} (450) Mg^{2+} (195) Mn^{2+} (280)	Zn^{2+} (180) Cu^{2+} (110) Co^{2+} (113)				
KerRP	* Ca^{2+}					
KerSMD	Ca^{2+} (112)				Na_2SO_3 (116)	
KerSMF	** Ca^{2+} (105)		**Tween-20 (112)**		Na_2SO_3 (115) DTT (115)	
KerA1	**Ca^{2+} (123) Mg^{2+} (199) Na^+ (135)**		**Tween 80 (113)**		**β-ME (Casein 100) (Keratin 192)**	
SAPB	Ca^{2+} (157) Mg^{2+} (112) Na^+ (118)		LAS (114) Tween 80 (119)	Tween 20 (117) SDS (119)	β-ME (192)	Urea (165) H_2O_2 (168)
rK$_{27}$	**Stability only tested**		**Triton X-100 (677) Tween-80 (242) Saponin (461) Sodium Cholate (276) SDS (186)**		**DTT (267) β-ME (323)**	**NaClO (276) H_2O_2 (275)**
KerQ7	**Ca^{2+} (417) Mg^{2+} (175) Mn^{2+} (250)**	**Ba^{2+} (121) Sn^{2+} (115)**				
KBALT	**Ca^{2+} (127) Mg^{2+} (134)**	**Zn^{2+} (129) Ba^{2+} (115)**	SDS (128)		**β-ME (102.5)**	
YT06 Protease	**Mg^{2+} (118) Mn^{2+} (196)**	**Ni^{2+} (120) Ba^{2+} (115)**	**Tween-20 (170)**		**β-ME (623)**	
M4 Peptidases						
RecGEOker	**Mg^{2+} (112) Mn^{2+} (116) Zn^{2+} 1 mM (58); 10 mM (52) Ca^{2+} 1 mM (101); 10 mM (66)**		**Triton X-100 (115) Tween 40 (180) Tween 60 (133)**	**Tween 80 (122) Triton X-305 (153)**	**DTT (139)**	

Note: Source organism, accession numbers and references can be found in Table 3; Bold = tested on a keratinous substrate; Not bold = tested on casein; * = Only tested for binding; ** included for comparison with KerSMD; underlined = denaturing detergents; DTT = dithiothreitol; β-ME = β-mercaptoethanol; LAS = linear alkylbenzene sulfonate; SDS = sodium dodecyl sulfate; TAED = tetraacetylethylenediamine; TTAB = tetradecyltrimethylammonium bromide; CHAPS = 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; CTAB = cetrimonium bromide.

Despite there being large differences in concentrations of metals additives, incubation time and temperature, in general, supplementation with Ca^{2+} showed the largest increase in activity except for KerSMF [101] and kerA1 [97]. In the case of KerSMF, Ca^{2+} had no effect on activity and in the case of kerA1, Mg^{2+} addition increased activity by 199% compared to 123% for Ca^{2+}. In general, Ca^{2+} > Mg^{2+} > Mn^{2+} had a positive effect on all the S1 and S8 keratinases (Table 5). The effect of these divalent metals on M4 metalloproteases is discrete compared to serine proteases. Only the addition of magnesium and manganese ions resulted in keratinase activity slightly above the control without additives. These divalent ions have been described to stabilize the active structure of the peptidases by binding to the metal-binding sites [100]. Other explanations for the higher activity are possible stabilization of enzyme/substrate complex [101] or formation of salt or ion bridges that maintain the enzyme conformation [97,122,128]. Furthermore, these metal ions reduce thermal denaturation [134]. Lin et al. observed that aqualysin, a thermostable peptidase from the S8 family, was only stable at high temperatures in the presence of 1 mM Ca^{2+} [135].

Several studies were carried out on the keratinase activity in the presence of metal ions (Zn^{2+}, Cu^{2+}, Co^{2+}, Ba^{2+}, Sn^{2+} and Ni^{2+}) were carried out (Table 5). The addition of the metal ions improved activity between 10% and 29% except for SAPDZ [100], where Zn^{2+} addition increased activity by 80% and Cu^{2+} addition increased activity of SFP2 [99] by 49%. Li et al. characterized SFP1, a non-keratinolytic peptidase similar to SFP2 and produced by the same organism [99]. It showed increased activity with copper ions, possibly due to the stabilization of the enzyme. Copper ions acting as a stabilizer has rarely been described in previous serine protease studies, and it is not known whether there is a copper-binding site stabilizing the enzyme [136]. In another example, peptidases were more stable in the presence of copper ions, which resulted in a reduction in both autolysis and thermal inactivation rates [137].

Detergents, in general, enable the disruption or formation of hydrophobic and hydrophilic bonds and assist in the extraction of proteins into aqueous media [138]. Detergents can act as denaturing agents on enzymes. Denaturing detergents are anionic (SDS, LAS) or cationic (CTAB, TTAB). They denature proteins by breaking protein–protein interactions. Non-denaturing detergents are non-ionic (Triton X-100, Tweens, cholate, saponin) or zwitterionic (CHAPS, sulfobetaine, zwittergent), and their action is milder and enzyme function is usually maintained. In most cases the addition of denaturing and non-denaturing detergents resulted in an increase in activity (110–150%). However, the addition of the non-ionic detergents to the assay mixture with keratin as substrate of rK_{27} had a dramatic effect on activity compared to the control without detergent [103]. Activities of 677% (Triton X-100), 242% (Tween 80), 461% (saponin) and 276% (cholate) were achieved. The addition of the anionic denaturing detergent, SDS to the assay increased the keratinase activity to 186%. The addition of the non-ionic detergents, Tween 40, Tween 60, Tween 80 and Triton X-305 to the assay mixture with keratin as substrate for the M4 keratinase, RecGEOker [109], showed increased activity to 180%, 133%, 122% and 153%, respectively. This example showed a definite trend of increasing activity with decreasing Tween 80 (monounsaturated C18 derivative) < Tween 60 (saturated C18 derivative) < Tween 40 (saturated C16 derivative). The partial solubilizing action of detergents on the insoluble keratin substrate might explain why both denaturing and non-denaturing detergents have a positive effect on keratin hydrolysis. There are insufficient examples to confirm this Tween effect on keratinases in general.

The reduction of disulphide bonds, destabilizes keratins and acts synergistically with keratin hydrolysis in nature [6,8–10]. Sodium sulphite, dithiothreitol (DTT) and β-mercaptoethanol were tested on some of the keratinases in Table 5. The reducing agents had a positive effect on all S8 keratinases tested with keratin as the substrate. The increase in activity ranged from 115% for Na_2SO_3 (KerSMF [101]) to 623% for β-mercaptoethanol (YT06 protease [98]) except in the case of KBALT [63], where β-mercaptoethanol had no effect on the activity. β-Mercaptoethanol doubled the activity of SAPB [102] when tested with casein as substrate. DTT also increased the activity of the M4 keratinase, RecGEOker (139% [109]), when tested with keratin as substrate. None of the S1 enzymes were tested

with keratin and reducing agents. However, the two assays with casein and reducing agent showed on one hand, no effect from β-mercaptoethanol on KERAK-29, [95] and on the other hand, a considerable effect on SFP2 (DTT, 278%; β-mercaptoethanol, 235% [99]). It should be noted that where disulphide bonds present in the enzyme are essential for function the inclusion of reducing agents may negatively affect activity.

Chaotropic agents are comparable to detergents, breaking non-covalent interactions and allowing protein denaturation [139–141]. Urea and isopropanol are chaotropic agents (Table 5). The activity of SAPB [102] was increased to 165% in the presence of urea compared to the control and the activity of NAPase [111] was increased to 130% in the presence of isopropanol [111].

The effect of the oxidizing agents, H_2O_2 and sodium hypochlorite, was also studied on three S8 and S1 keratinases, SAPB [102], rK$_{27}$ [103] and KERAK-29 [95]. Activity was significantly increased in all cases (Table 5).

In most cases the effect of additives like divalent cations, detergents, reducing agents, chaotropic agents and oxidizing agents have a positive effect on keratinase activity. Nearly all compounds capable of disrupting the integrity of the keratin structure without inactivating the keratinase appear to have a positive effect on keratinase activity. The effect of compounds disrupting the keratin structure was, in some cases like rK$_{27}$ [103], significant.

5.4. Substrate Specificity

Table 6 summarizes the substrate specificity data of the selected keratinases from the literature. In general, a variety of keratins and other proteins like gelatin, casein and albumin were tested. To compare the selected keratinases, the values in Table 6 have been normalized using the activity of designated keratin substrates (keratin azure, keratin, feather or wool) as 100% activity.

KerQ7 [104] was the only keratinase in Table 6 tested on multiple types of keratins. KerQ7 showed a preference for the β-keratin-rich feather meal and feathers. The activity on feather meal was only 16% higher than feathers. The activity on rabbit hair, goat hair and bovine hair was 88%, 74% and 50% of the activity on feathers, respectively, whereas activity on wool was only 12%. These substrates are rich in α-keratins [4,9,12,13]. Substrate fibre thickness and fibre surface area may also contribute to the variations in enzyme activity. Nonetheless, the keratinase activity toward various substrates is likely to be multifactorial. KerSMD and KerSMF, from *Stenotrophomonas maltophilia*, showed less activity towards feather powder and wool than keratin azure [101]. KerSMD and KerSMF had similar activity on feather powder (54% and 71%, respectively) and wool (59% and 78%, respectively). However, KerSMD showed an activity of 1589% towards soluble keratin compared to an activity of 126% for KerSMF on the same substrate.

No trends were observable on non-keratin substrates. For example, SAPDZ [100] showed 81% activity on gelatin compared to keratin, whereas the activity of kerA1 [97] and SAPB [102] on gelatin was 22% and 146% compared to keratin, respectively. The same inconsistencies can be seen with casein. The activities of SAPB [102] and KerSMD [101] on casein are 153% and 2800%, respectively, compared to keratin azure, whereas KerSMF [101] has only slightly lower activity on casein (91%) compared to keratin azure.

Keratinases are known for their activity on "hard-to-degrade" proteinogenic substrates. Most of the characterized keratinases in the literature are also capable of degrading collagen, which is an example of another complex and hard-to-degrade substrate [142]. A study in 2008 characterized the first keratinase without collagenase activity [143]. Only three enzymes from the S8 family were tested on collagen or azocoll (azocollagen). Vpr [105] presented collagenase activity (129%) and while SAPDZ [100] did not. Protease C2 [106] showed a surprisingly high activity (24000%) on azocoll compared to keratin azure (100%). KERDZ [94], from the S1 peptidase family, had no activity on collagen, whereas RecGEOker [109], belonging to the M4 metallopeptidase family, was able to hydrolyse collagen. The differences in activity between substrates may be attributed to the specific peptide sequences in the substrates and the sequence specificity of the enzymes.

Table 6. Substrate specificity of the selected S1, S8 and M4 keratinases. Numbers in brackets correspond to the % activity relative to other substrates.

Protein	Keratins (%)	Natural Proteins (%)		Modified Protein (%)	Esters and Others (%)	
S1A Peptidases						
KERDZ	Keratin (100) [2]	Gelatin (90) Casein (79) Albumin (75) Elastin (50)	Myoglobin (41) Hemoglobin (20) Collagen type 1/2 (0)	Azocasein (80) Azoalbumin (70)	BAEE (91) TAME (100) BCEE (95) BTEE (0) ATEE (0)	
SFP2	Keratin (100)[2]	Casein (111)				
S8A Peptidases						
SAPDZ	Keratin (100)	Gelatin (81) Casein (95) Albumin (72)	Hemoglobin (66) Collagen type 1/2 (0)	Azocasein (91) Keratin Azure (100) [1]	BAEE (0) BCEE (0)	BTEE (100) ATEE (95)
KerSMD	Feather powder (54) Soluble keratin (1589) Wool (59)	Casein (2800)		Keratin Azure (100) [1]		
KerSMF	Feather powder (71) Soluble keratin (126) Wool (78)	Casein (92)		Keratin Azure (100) [1]		
KerA1	Keratin (100) [2]	Gelatin (22) Casein (222)	Elastin (54) BSA (97) Egg albumin (4)	Azocasein (177) Azokeratin (92)		
SAPB	Keratin (100)[2]	Gelatin (146) Casein (153)	Bovine serum albumin (80) Egg albumin (18) Gluten (30)	Azocasein (123) Azokeratin (96)	BTEE (109) ATEE (115)	
rK₂₇	Powdered chicken feather > haemoglobin > meat protein > hoof keratin > fibrin > elastin > gelatine > casein > BSA > azocasein > keratin azure					
KerQ7	Rabbit hair (88) Wool (12) Goat hair (74) Feather meal (116) Bovine hair (50) Feather (100) [3]					
Vpr	Feather meal (50) Keratin (100) [2]	Gelatin (147) Casein (156)	Fibrin (145) Collagen (129)			
Protease C2	Bovine hair (274) Feather (439)	Albumin (8571) Elastin (11)		Keratin azure (100) [1] Azocasein (102857) Azocoll (24000)		
M4 Peptidases						
RecGEOker	Wool (100) [5]	Gelatin (92) Casein (95)	Albumin (37) Collagen type 1 (98)			

Note: Source organism, accession numbers and references can be found in Table 3; Activity normalized to the following substrates—[1] keratin azure, [2] keratin; [3] feather, [4] wool; BAEE = N-α-benzoyl-L-arginine ethyl ester; TAME = N-α-p-tosyl-L-arginine methyl; BCEE = benzoyl-citrulline ethyl ester; BTEE = N-benzoyl-L-tyrosine ethyl ester; ATEE = N-acetyl-L-tyrosine ethyl ester; DTNB = 5,5′-dithiobis-(2-nitrobenzoic acid); Azocoll = commercially available azocollagen.

Some enzymes also showed esterase activity, which may be of importance for facilitating enzyme access to the substrate. Fatty acids of the cuticle surface are linked via a thioester to the protein matrix below in keratin fibres and feathers [14]. Only three enzymes in Table 6 have been characterized on ester substrates. The two S8 family peptidases—SAPDZ [100] and KerRP [96]—appear to have similar ester substrate affinity with both showing activity against N-benzoyl-L-tyrosine ethyl ester (BTEE) and N-acetyl-L-tyrosine ethyl ester (ATEE). In contrast, KERDZ (S1 family) had no activity towards these substrates but was active towards N-α-benzoyl-L-arginine ethyl ester (BAEE), N-α-p-tosyl-L-arginine methyl (TAME) and benzoyl-citrulline ethyl ester (BCEE) [94].

In general, the peptidases from S1, S8 and M4 families (Table 6) present varied substrate specificities. There are limited examples in the S1 and M4 families to detect trends but even within the S8 peptidases examples there were no obvious substrate preferences.

6. Potential Applications of Keratinases

New keratinases with improved properties for commercialization and the keratin hydrolysates they produce represent an opportunity for adding value to keratin waste.

Commercial keratinases are sold for a variety of applications (Table 7) such as the degradation of infectious prions, as supplements for animal feed to improve its nutritional value, removal of corns and calluses from skin, treatment of acne and nail fungi and, they are also incorporated into cosmetic skin peeling and depilatory creams [6,52,144]. Other applications include the use in cleaning products for unblocking drainpipes and septic tanks.

Table 7. Some examples of commercial keratinases.

Trade Name	Source	EC Number	Substrate or Function	Supplier
Versazyme [1,3]	*Bacillus licheniformis*	3.4.21.62/ S8 family	Improving nutritional value of poultry feed & prions degradation	Bioresource Int'l, Inc.
Ronozyme ProAct [2]	*Nocardiopsis prasina*	3.4.21.-/serine protease	Improving nutritional value animal feed	DSM/Novozymes
Cibenza DP100 [2]	*Bacillus licheniformis* PWD-1	-	Improving nutritional value animal feed	Novus International
Pure Keratinase 100 [3]	*Bacillus licheniformis* PWD-1	-	Prion degradation from medical & dental instruments	Proteus Biotech
BioGuard Plus [3]	Proprietary blend of microorganisms – incl. keratinase producer	-	Cleaning drainpipes, septic tanks & digesters	RuShay Inc.
Keratoclean sensitive PB [3]	*Bacillus licheniformis* (PB333 keratinase)	-	Treatment acne, dead skin removal, promotes cell renewal	Proteus Biotech
Keratoclean Hydra PB [3]	*Bacillus licheniformis*	-	Removal of corns & call uses, acne, Hirsutism, peeling	Proteus Biotech
FixaFungus [3]	-	-	Treatment of toenail fungal infections	Proteus Biotech

Source: [1] [6], [2] [127], [3] [52].

There are also a number of promising applications of keratinases that have not been commercialized to date: dag or manure balls removal from cattle hides and tails [145]; extraction of glucocorticoids from chicken feathers to monitor the stress level in poultry breeding and production programmes [146] extraction of chicken feather cholesterol as a precursor to bile salts that can be used to produce bio-emulsifiers and biosurfactants in the cosmetic industry [18]; selective hydrolysis of wool from wool/polyester or mixed textiles to facilitate textile recycling [147]; and dehairing of hides in the leather industry [64,65,82].

The use of keratinases for the processing of keratinous waste might be advantageous for high value products. The use of enzymes instead of thermochemical methods for keratin hydrolysis reduces chemical modification arising from harsh chemical hydrolysis and might allow a degree of control of the peptide composition that is produced. Keratin hydrolysates are widely used in protein feed supplements [18]. Feather waste, for example, is hydrolysed with saturated steam under high pressure (sometimes with the addition of lime) to produce feather meal, which is used as a feed supplement [148]. These conditions lead to the loss or modification of some of the amino acids, which impacts the nutritional value and digestibility of the feather meal. Hydrolysis with keratinases might offer an alternative, which reduces the energy requirements of the process and enhances the nutritional value of the supplement.

Keratin-derived bioactive peptides have been reported in the literature. These peptides have a range of activities like antimicrobial [149], antihypertensive [150], anti-inflammatory [151–154], antioxidant [149,150,155], inhibition of early stage amyloid aggregation [156], antidiabetic [157] or anti-aging [158–160] depending on the keratin source and the method of preparation. Producing protein feed supplements with antioxidant or anti-inflammatory properties as well as skin and hair

products with antioxidant, anti-inflammatory, antimicrobial or anti-aging properties would most likely increase the value of these products.

Keratin peptides and subunits can spontaneously self-assemble [161]. This property can be exploited to form biomaterials like hydrogels, films, sponges, scaffolds and nanofibres for tissue engineering, wound healing, fibroblast cultivation and treatment of burns [161–164]. The production of smart biocomposites is also of interest. An example is the production of transparent plastic film containing citric acid from wool hydrolysate [165]. The plastic has excellent biocidal activity and could be used as a functional packaging for food.

The examples described above demonstrate the commercial potential of keratinases and the large number of opportunities they offer for adding value to keratin waste by producing bioactive protein feed supplements, personal care products and biomaterials from keratin hydrolysates.

7. Discovery and Future Research

The standout problem with the characterization of keratinases, demonstrated by the analysis of the assay conditions in this review, is lack of standardization of the keratinase assay combined with the small number of sequenced peptidases with keratinolytic activity that have been biochemically characterized. Both of these issues hamper the identification and comparison of true keratinases. Current experimental conditions vary in temperature, pH, buffer types and concentration, additives, substrates and their pretreatment biasing possible conclusions. It is unclear whether some proteases are keratinases or whether pretreatment or co-treatment influences their keratinolytic activity to some degree.

The uncertainty in defining keratinases and highly variable characterization of keratinases in the literature increases the challenge of finding new keratinases based on literature data or from sequence databases. However, the discovery of new keratinases is critical for expanding the opportunities for waste keratin valorisation. It would be desirable to identify new keratinases with high activities and specificities enabling control over cleavage sites, peptide molecular weights and amino acid side chain modifications.

Standardized experiments combined with phylogenetic studies and sequence analyses are needed. Standardized experiments, which avoid pre- or co-treatments, would determine the true protease activity on keratin substrates and reduce possible experimental biases. An in-depth phylogenetic analysis would help to clarify the position of keratinases within the phylogenetic trees of the peptidase families in which they are found and may help focus the search for new peptidases with keratinolytic activity. A comprehensive sequence analyses, aimed at the identification of conserved sites between peptidases with keratinolytic activity, as well as the presence of specific domains that possibly contribute to their ability to hydrolyse keratin, may assist in the development of algorithms to search the vast sequence space of the peptidase families.

References

1. Tesfaye, T.; Sithole, B.; Ramjugernath, D. Valorisation of chicken feathers: A review on recycling and recovery route—current status and future prospects. *Clean Technol. Environ. Policy* **2017**, *19*, 2363–2368. [CrossRef]

2. Thyagarajan, D.; Barathi, M.; Sakthivadivu, R. Scope of poultry waste utilization. *IOSR-JAVS* **2013**, *6*, 29–35.
3. Gooding, C.H.; Meeker, D.L. Review: Comparison of 3 alternatives for large-scale processing of animal carcasses and meat by-products. *PAS* **2016**, *32*, 259–270. [CrossRef]
4. Sinkiewicz, I.; Staroszczyk, H.; Sliwinska, A. Solubilization of keratins and functional properties of their isolates and hydrolysates. *J. Food Biochem.* **2018**, *42*, e12494. [CrossRef]
5. Chojnacka, K.; Gorecka, H.; Michalak, I.M.; Gorecki, H. A review: Valorization of keratinous materials. *Waste Biomass Valoriz.* **2011**, *2*, 3017–3021. [CrossRef]
6. Lange, L.; Huang, Y.; Busk, P.K. Microbial decomposition of keratin in nature—A new hypothesis of industrial relevance. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 2083–2096. [CrossRef]
7. Peng, Z.; Zhang, J.; Du, G.; Chen, J. Keratin waste recycling based on microbial degradation: Mechanisms and prospects. *ACS Sustain. Chem. Eng.* **2019**, *7*, 9727–9736. [CrossRef]
8. Jones, L.N. Hair structure anatomy and comparative anatomy. *Clin. Dermatol.* **2001**, *19*, 95–103. [CrossRef]
9. Bragulla, H.H.; Homberg, D.G. Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. *J. Anat.* **2009**, *214*, 516–559. [CrossRef]
10. Plowman, J.E. Proteomic database of wool components. *J. Chromatog. B.* **2003**, *787*, 63–76. [CrossRef]
11. Da Silva, R.R. Different processes for keratin degradation: The ways for the biotechnological application of keratinases. *J. Agric. Food Chem.* **2018**, *66*, 9377–9378. [CrossRef] [PubMed]
12. Yang, F.; Zhang, Y.; Rheinstädter, M.C. The structure of people's hair. *PeerJ* **2014**, *2*, e619. [CrossRef] [PubMed]
13. Wang, B.; Yang, W.; McKittrick, J.; Meyers, M.A. Keratin: structure, mechanical properties, occurrence in biological organisms, and efforts of bioinspiration. *Prog. Mater. Sci.* **2016**, *76*, 229–318. [CrossRef]
14. Ganske, F.; Meyer, H.H.; Deutz, H.; Bornscheuer, U. Enzyme-catalysed hydrolysis of 18-methyleicosanoic acid-cysteine thioester. *Eur. J. Lipid Sci. Technol.* **2003**, *105*, 627–632. [CrossRef]
15. Koepke, V.; Nilssen, B. Wool surface properties and their influence on dye uptake—A microscopical study. *J. Text. Inst.* **1960**, *51*, T1398–T1413. [CrossRef]
16. Wolski, T. *Modified Keratin Proteins, Their Physiochemical Properties, Analysis and Application*; Medical Academy: Lublin, Poland, 1985.
17. Xie, H.; Li, S.; Zhang, S. Ionic liquids as novel solvents for the dissolution of and blending of wool keratin fibers. *Green Chem.* **2005**, *7*, 606–608. [CrossRef]
18. Ningthoujam, D.S.; Tamreihao, K.; Mukherjee, S.; Khunjamayum, R.; Devi, L.J.; Asem, R.S. Keratinaceous Wastes and Their Valorization Through Keratinolytic Microorganisms. In *Keratin*; Blumenberg, M., Ed.; IntechOpen: London, UK, 2018; Volume 1, pp. 129–148.
19. Wu, G. *Amino Acids: Biochemistry and Nutrition*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2013.
20. Papadopoulis, M.C. Effect of processing on high-protein feedstuffs: a review. *Biol. Wastes* **1989**, *29*, 123–138. [CrossRef]
21. Latshaw, J.D.; Musharaf, N.; Retrum, R. Processing of feather meal to maximize its nutritional value for poultry. *Animal Feed Sci. Technol.* **1994**, *47*, 179–188. [CrossRef]
22. Schwass, D.E.; Finley, J.W. Heat and alkaline damage to proteins: Racemization and lysinoalanine formation. *J. Ag. Food Chem.* **1984**, *32*, 1377–1382. [CrossRef]
23. Liardon, R.; Hurrell, R.F. Amino acid racemization in heated and alkali-treated proteins. *J. Ag. Food Chem.* **1983**, *31*, 432–437. [CrossRef]
24. Liardon, R.; Ledermann, S. Racemization kinetics of free and protein-bound amino acids under moderate alkaline treatment. *J. Agric. Food Chem.* **1986**, *34*, 557–565. [CrossRef]
25. Vasconcelos, A.; Freddi, G.; Cavaco-Paulo, A. Biodegradable materials based on silk fibroin and keratin. *Biomacromolecules* **2008**, *9*, 1299–1305. [CrossRef] [PubMed]
26. Yamauchi, K.; Yamauchi, A.; Kusonoki, T.; Kohda, A.; Konishi, Y. Preparation of stable aqueous solution of keratins, and physiochemical and degradational properties of films. *J. Biomed. Mater. Res. Part Res.* **1996**, *31*, 439–444. [CrossRef]
27. Wang, S.; Taraballi, F.; Tan, L.P.; Ng, K.W. Human keratin hydrogels support fibroblast attachment and proliferation in vitro. *Cell Tissue Res.* **2012**, *347*, 795–802. [CrossRef]
28. Tonin, C.; Aluigi, A.; Vineis, C.; Varesano, A.; Montarsolo, A.; Ferrero, F. Thermal and structural characterization of poly (ethylene-oxide)/keratin blend films. *J. Therm. Anal. Calorim.* **2007**, *89*, 601–608. [CrossRef]
29. Wang, K.; Li, R.; Ma, J.H.; Jian, Y.K.; Che, J.N. Extracting keratin from wool using cysteine. *Green Chem.* **2016**, *18*, 476–481. [CrossRef]

30. Torchinski, Y.M. *Sulfur in Proteins*; Pergamon Press: Oxford, UK, 1981.

31. Cardamone, J.M. Investigating the microstructure of keratin extracted from wool: Peptide sequence (MALDI-TOF/TOF) and protein conformation (FTIR). *J. Mol. Struct.* **2010**, *969*, 97–105. [CrossRef]

32. De Guzman, R.C.; Merrill, M.R.; Richter, J.R.; Hamzi, R.I.; Greengauz-Roberts, O.K.; Van Dyke, M.E. Mechanical and biological properties of keratose biomaterials. *Biomaterials* **2011**, *32*, 8205–8217. [CrossRef]

33. Cardamone, J.M.; Nunez, A.; Garcia, R.A.; Aldema-Ramos, M. Characterizing wool keratin. *Lett. Mater. Sci.* **2009**, *2009*, 147175. [CrossRef]

34. Kelly, R.J.; Worth, G.H.; Roddick-Lanzilotta, A.D.; Rankin, D.A.; Ellis, G.D.; Mesman, P.J.R.; Summers, C.G.; Singleton, D.J. The Production of Soluble Keratin Derivatives. U.S. 7148327B2, 2001.

35. Cornfield, M.C.; Robson, A. The amino acid composition of wool. *Biochem. J.* **1955**, *59*, 62–68.

36. Mellet, P. The influence of alkali treatment on native and denatured proteins. *Tex. Res. J.* **1968**, *38*, 977–983. [CrossRef]

37. Masters, P.M.; Friedman, M. Racemization of amino acids in alkali-treated food proteins. *J. Agric. Food Chem.* **1979**, *27*, 507–511. [CrossRef]

38. Provansal, M.M.P.; Cug, J.L.A.; Cheftal, J.C. Chemical and nutritional modifications of sunfower proteins due to alkaline processing. Formation of amino acid cross-links and isomerization of lysine residues. *J. Agric. Food Chem.* **1975**, *23*, 938–943. [CrossRef]

39. Gaidau, C.; Epure, D.-G.; Enascuta, C.E.; Carsote, C.; Sendrea, C.; Proietti, N.C.W.; Gu, H. Wool keratin total solubilisation for recovery and reintegration—An ecological approach. *J. Clean. Prod.* **2019**, *236*, 117586. [CrossRef]

40. Koleva, M.; Danalev, D.; Ivanova, D.; Vezenkov, L.; Vassiliev, N. Synthesis of two peptide mimetics as markers for chemical changes of wool's keratin during skin unhairing process and comparison the wool quality obtaianed by ecological methods for skin unhairing. *Bulg. Chem. Commun.* **2009**, *41*, 160–164.

41. Banga, A.K. *Therapeutic Peptides and Proteins: Formulation, Process and Development Systems*; CRC Press: Boca Raton, FL, USA, 1997.

42. Murray, K.; Rasmussen, P.; Neustaedter, J.; Luck, J.M. The hydrolysis of arginine. *J. Biol. Chem.* **1965**, *240*, 705–709.

43. Daft, F.S.; Coghill, R.D. The alkaline decomposition of serine. *J. Biol. Chem.* **1931**, *90*, 341–350.

44. Hunt, S. The non-protein amino acids. In *Chemistry and Biochemistry of the Amino Acids*; Barrett, G.C., Ed.; Springer: Dordrecht, Germany, 1985.

45. Stapleton, I.; Swan, J. Amino acids and peptides. VI. Studies on cystine and αα'-dimethylcystine in relation to the alkaline degradation of protein disulphides. *Aust. J. Chem.* **1960**, *13*, 416–425. [CrossRef]

46. Ohta, T.; Suzuki, S.; Todo, M.; Kurechi, T. The decomposition of tryptophan in acid solutions: specific effect of hydrochloric acid. *Chem. Pharm. Bull.* **1981**, *29*, 1767–1771. [CrossRef]

47. Tung, W.S.; Daoud, W.A. Photocatalytic self-cleaning keratins: A feasibility study. *Acta Biomater.* **2009**, *5*, 50–56. [CrossRef]

48. Kurbanoglu, E.B.; Kurbanoglu, N.I. A new process for the utilization of ram horn waste. *J. Biosci. Bioeng.* **2002**, *94*, 202–206. [CrossRef]

49. Zhang, J.; Li, Y.; Li, J.; Zhao, Z.; Liu, X.; Li, Z.; Han, Y.; Hu, J. Isolation and characterization of biofunctional keratin particles extracted from wool waste. *Powder Technol.* **2013**, *246*, 356–362. [CrossRef]

50. Kurbanoglu, E.B.; Kurbanoglu, N.I. Ram horn hydrolysate as enhancer of xanthan production in batch culture of *Xanthomonas campestris* EBK-4 isolate. *Process Biochem.* **2007**, *42*, 1146–1149. [CrossRef]

51. Brandelli, A. Bacterial keratinases: useful enzymes for bioprocessing agroindustrial wastes and beyond. *Food Bioproc. Technol.* **2008**, *1*, 105–116. [CrossRef]

52. Sharma, R.; Devi, S. Versatility and commercial status of microbial keratinases: A review. *Rev. Environ. Sci. Biotechnol.* **2018**, *17*, 19–45. [CrossRef]

53. Muhsin, T.H.; Aubaid, A.H. Partial purification and some biochemical characteristics of exocellular keratinase from *Trichophyton mentagrophytes* var erinacei. *Mycopathologia* **2001**, *150*, 121–125. [CrossRef]

54. Rai, S.K.; Mukherjee, A.K. Optimization of the production of an oxidant and detergent-stable alkaline beta-keratinase from *Brevibacillus* sp. strain AS-S10-11: application of enzyme in laundry detergent formulations and in the leather industry. *Biochem. Eng. J.* **2011**, *54*, 47–56. [CrossRef]

55. Thankaswamy, S.R.; Sundaramoorthy, S.; Palanivel, S.; Ramudu, K.N. Improved microbial degradation of animal hair waste from leather industry using *Brevibacterium luteolum* (MTCC 5982). *J. Clean. Prod.* **2018**, *189*, 701–708. [CrossRef]

56. Nam, G.-W.; Lee, D.-W.; Lee, H.-S.; Lee, N.-J.; Kim, B.-C.; Choe, E.-A.; Hwang, J.-K.; Suhartono, M.T.; Pyun, Y.-R. Native feather-degradation by *Fervidobacterium islandicum* AW-1, a newly isolated keratinase-producing thermophilic anaerobe. *Arch. Microbiol.* **2002**, *178*, 538–547. [CrossRef]

57. Kshteri, P.; Roy, S.S.; Sharma, S.K.; Singh, T.S.; Ansari, M.A.; Sailo, B.; Singh, S. Feather degrading, phytostimulating, and biocontrol potential of actinobacteria from North Easatern Himalayan Region. *J. Basic Microbiol.* **2018**, *58*, 730–738. [CrossRef]

58. Mamangkey, J.; Suryanto, D.; Munir, E.; Mustopa, A.Z. Isolation, molecular identification and verification of gene encoding bacterial keratinase from crocodile (crocodylus porosus) feces. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *305*, 012085. [CrossRef]

59. Pernicova, I.; Enev, V.; Marova, I.; Obruca, S. Interconnection oof waste chicken feather biodegradation and keratinase and *mcl*-PHA prooduction employing *Pseudomonas putida* KT2440. *Appl. Food Biotechnol.* **2019**, *6*, 83–90.

60. Fuke, P.; Gujar, V.V.; Khardenavis, A. Genome annotation and vlidation of keratin hydrolyzing proteolytic enzymes from *Serratia marcescens* EGD-HP20. *Appl. Biochem. Biotechnol.* **2018**, *18*, 970–986. [CrossRef] [PubMed]

61. Fuke, P.; Pal, R.R.; Khardenavis, A.A.; Purohit, H. *In silico* characterization of broad range proteases produced by *Serratia marcescens* EGD-HP20. *J. Basic Microbiol.* **2018**, *58*, 492–500. [CrossRef]

62. Bhari, R.; Kaur, M.; Singh, R.S. Thermostable and halotolerant keratinae from *Bacillus aerius* NSMk2 with remarkable dehairing and laundry applications. *J. Basic Microbiol.* **2019**, *59*, 555–558. [CrossRef]

63. Pawar, V.A.; Prajapati, A.S.; Akhani, R.C.; Patel, D.H.; Subramanian, R.B. Molecular and biochemical characteriztion of a thermostable keratinase from *Bacillus altitudinis* RBDV1. *3 Biotech* **2018**, *8*, 107–113. [CrossRef]

64. Cao, S.; Li, D.; Ma, X.X.Q.; Song, J.; Lu, F.; Li, Y. A novel unhairing enzyme produced by heterolgous expression of keratinase gene (*kerT*) in *Bacillus subtilis*. *World J. Microbiol. Biotechnol.* **2019**, *35*, 122–131. [CrossRef]

65. Kumar, M.; Bhatia, D.; Khatak, S.; Kumar, R.; Sharma, A.; Malik, D.K. Optimization and purification of keratinase from *Bacillus anthracis* with dehairing application. *J. Pure Appl. Microbiol.* **2019**, *13*, 585–590. [CrossRef]

66. Kalaikumari, S.S.; Vennila, T.; Monika, V.; Chandraraj, K.; Gunasekaran, P. Bioutilization of poultry feather for keratinase production and its application in leather industry. *J. Clean. Prod.* **2019**, *208*, 44–45. [CrossRef]

67. Hamiche, S.; Mechri, S.; Khelouia, L.; Annane, R.; El Hattab, M.; Badis, A.; Jaouadi, B. Purification and biochemical characterization of two keratinases from *Bacillus amyloliquefaciens* S13 isolated from marine brown algae *Zonaria tournefortii* with potential keratin-biodegradation and hide-unhairing activities. *Int. J. Biol. Macromol.* **2019**, *122*, 758–769. [CrossRef]

68. Devi, C.S.; Shankar, R.S.; Kumar, S.; Mohanasrinivasan, B. Production of keratinase from a newly isolates feather degrading *Bacillus cereus* VITSDVM4 from poultry waste. *Natl. Acad. Sci. Lett.* **2018**, *41*, 307–311. [CrossRef]

69. Arokiyaraj, S.; Varghese, R.; Ahmed, B.A.; Duraipandiyan, V.; Al-Dhabi, N.A. Optimizing the fermentation conditions and enhanced production of keratinase from *Bacillus cereus* isolated from halophilic environment. *Saudi J. Biol. Sci.* **2019**, *26*, 378–381. [CrossRef] [PubMed]

70. Abdel-Fattah, A.M.; El-Gamal, M.S.; Ismail, S.A.; Emran, M.A. Biodegradation of feather waste by keratinase produced from newly isolated *Bacillus licheniformis* ALW1. *JGEB* **2018**, *16*, 311–318. [CrossRef]

71. Cavello, I.; Urbieta, M.S.; Segretin, A.B.; Giaveno, A.; Cavalitto, S.; Donati, E.R. Assessment of keratinase and other hydrolytc enzymes in thermophilic bacteria isolated from geothermal areas in Patagonia Argentina. *Geomicrobiol. J.* **2018**, *35*, 156–165. [CrossRef]

72. Adetunji, C.O.; Adejumo, I.O. Efficacy of crude and immobilized enzymes from *Bacillus licheniformis* for production of biodegraded feather meal and their assessment. *Environ. Technol. Inno.* **2018**, *11*, 116–124. [CrossRef]

73. Hashem, A.M.; Abdel-Fattah, A.M.; Ismail, S.A.; El-Gamal, M.S.; Esawy, M.A.; Emran, M.A. Optimization, characterization and thermodynamic studies on *B. licheniformis* ALW1 keratinase. *Egypt J. Chem.* **2018**, *61*, 591–607. [CrossRef]

74. He, Z.; Sun, R.; Tang, Z.; Bu, T.; Wu, Q.; Li, C.; Chen, H. Biodegradation of feather waste keratin by keratin-degrading strain *Bacillus subtilis* 8. *J. Microbiol. Biotechnol.* **2018**, *28*, 314–322. [CrossRef]

75. Oluwaseun, A.C.; Phazang, P.; Sarin, N.B. Production of ecofriendly biofertilizers produced from crude and immobilized enzymes from *Bacillus subtilis* CH008 and their effect on the growth of *Solanum lycopersicum*. *Plant Arch.* **2018**, *18*, 1455–1462.

76. Nagarajan, S.; Eswaran, P.; Masilamani, R.P.; Natarajan, H. Chicken feather compost to promote the plant growth activity by using keratinolytic bacteria. *Waste Biomass Valor.* **2018**, *9*, 531–538. [CrossRef]

77. Imtiaz, A.; Rehman, A. *Bacillus subtilis* BML5 isolated from soil contaminated with poultry waste has keratinolytic activity. *Pakistan J. Zool.* **2018**, *50*, 143–148. [CrossRef]

78. De Paiva, D.P.; De Oliveira, S.S.A.; Mazotto, A.M.; Vermehlo, A.B.; De Oliveira, S.S. Keratinolytic activity of *Bacillus subtilis* LFB-FIOCRUZ 1266 enhanced by whole-cell mutagenesis. *3 Biotech* **2019**, *9*, 2–13. [CrossRef] [PubMed]

79. Sharma, S.; Prasad, R.K.; Chatterjee, S.; Sharma, A.; Variable, M.G.; Yadav, K.K. Characterization of Bacillus species with keratinase and cellulase properties isolated from feather dumping and cockroach gut. *Proc. Natl. Acad. Sci. India. Sect. B Biol. Sci* **2019**, *89*, 1079–1086. [CrossRef]

80. Suharti, D.R.T.; Nilamsari, N.R. Isolation and characterization of newly keratinase producing *Bacillus* sp. N1 from tofu liquid waste. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *230*, 012088. [CrossRef]

81. Jin, M.; Chen, C.; He, X.; Zeng, R. Characterization of an extreme alkaline-stable keratinase from the draft genome of feather-degrading *Bacillus* sp. JM7 from deep-sea. *Acta Oceanol. Sin.* **2019**, *38*, 87–95. [CrossRef]

82. Tian, J.; Xu, Z.; Long, X.; Tian, Y.; Shi, B. High-expression keratinase by *Bacillus subtilis* SCK6 for enzymatic dehairing of goatskins. *Int. J. Biol. Macromol.* **2019**, *135*, 119–126. [CrossRef]

83. Koentjoro, M.P.; Prasetyo, E.N.; Rahmatullah, A.M. Optimization of keratinase production by Bacillus SLII-I bacteria in chicken feather waste medium. *ARPN J. Engin. Appl. Sci.* **2018**, *13*, 482–488.

84. Nurkhasanah, U.; Suharti. Preliminary study on keratinase fermentation by *Bacillus* sp. MD24 under solid state fermentation. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *276*, 012016. [CrossRef]

85. Suharti, S.; Riesmi, M.T.; Hidayati, A.; Zuhriyah, U.F.; Wonorahardjo, S.; Susanti, E. Enzymatic dehairing of goat skin using keratinase from *Bacillus* sp. MD24, a newly isolated soil bacterium. *Pertanika J. Trop. Agric. Sci.* **2018**, *41*, 1449–1461.

86. Ashokkumar, M.; Irudayaraj, G.; Yellapu, N.; Manonmani, A.M. Molecular characterization of *bmyC* gene of the mosquito pupicidal bacterial, *Bacillus amyloliquefaciens* (VCRC B483) and in silico analysis of bacillomycin D snthetase C protein. *World J. Microbiol. Biotechnol.* **2018**, *34*, 116–126. [CrossRef]

87. Yusuf, I.; Ahmad, S.A.; Pang, L.Y.; Yasid, N.A.; Shukor, M.Y. Effective production of keratinase by gellum gum-immobilised *Alcaligenes* sp. AQ05-001 using heavy metal-free and polluted feather wastes as substrates. *3 Biotech* **2019**, *9*, 32–43. [CrossRef]

88. Sultana, N.; Saha, P. Studies on potential application of crude keratinase enzyme from *Stenotrophomonas* sp. for dehairing in leather processing industry. *J. Environ. Biol.* **2018**, *39*, 324–330. [CrossRef]

89. Kang, D.; Herschend, J.; Al-Soud, W.A.; Mortensen, M.S.; Gonzalo, M.; Jacquiod, S.; Sorensen, S.J. Enrichment and characterization of an environmental microbial consortium displaying efficient keratinolytic acitivty. *Bioresour. Technol.* **2018**, *270*, 303–310. [CrossRef] [PubMed]

90. Mercer, D.K.; Stewart, C.S. Keratin hydrolysis by dermophytes. *Med. Mycol. J.* **2019**, *57*, 13–22. [CrossRef]

91. Pillai, P.; Mandge, S.; Archana, G. Statistical optimization of production and tannery application of a keratinolytic serine protease from *Bacillus subtilis* P13. *Process Biochem.* **2011**, *46*, 1110–1117. [CrossRef]

92. Purchase, D. *Microbial Keratinases: Characteristics, Biotechnological Applications and Potential*; CABI: Wallingford, UK, 2016.

93. Gupta, R.; Sharma, R.; Beg, Q.K. Revisiting microbial keratinases: Next generation proteases for sustainable biotechnology. *Crit. Revs. Biotechnol.* **2013**, *33*, 216–228. [CrossRef]

94. Elhoul, M.B.; Jaouadi, N.Z.; Rekik, H.; Benmrad, M.O.; Mechri, S.; Moujehed, E.; Kourdali, S.; Hattab, M.E.; Badis, A.; Bejar, S.; et al. Biochemical and molecular characterization of new keratinoytic protease from *Actinomadura viridilutea* DZ50. *Int. J. Biol. Macromol.* **2016**, *92*, 299–315. [CrossRef] [PubMed]

95. Habbeche, A.; Saoudi, B.; Jaouadi, B.; Haberra, S.; Kerouaz, B.; Boudelaa, M.; Badis, A.; Ladjama, A. Purification and biochemical characterization of a detergent-stable keratinase from a newly thermophilic actinomycete *Actinomadura keratinilytica* strain Cpt29 isolated from poultry compost. *J. Biosci. Bioeng.* **2014**, *117*, 413–421. [CrossRef]

96. Fakhfakh, N.; Kanoun, S.; Manni, L.; Nasri, M. Production and biochemical and molecular characterization of a keratinolytic serine protease from chicken feather-degrading *Bacillus licheniformis* RPk. *Can. J. Microbiol.* **2009**, *55*, 427–436. [CrossRef]

97. Fakhfakh-Zouari, N.; Hmidet, N.; Haddar, A.; Kanoun, S.; Nasri, M. A novel serine metalloprotease from a newly isolated *Bacillus pumilus* A1 grown on chicken feather meal: biochemical and molecular characterization. *Appl. Biochem. Biotechnol.* **2010**, *162*, 329–344. [CrossRef]

98. Wang, L.; Qian, Y.; Cao, Y.; Huang, Y.; Chang, Z.; Huang, H. Production and characterization of keratinolytic proteases by a chicken feather-degrading thermophilic strain, *Thermoactinomyces* sp. YT06. *J. Microbiol. Biotechnol.* **2017**, *27*, 2190–2198. [CrossRef]

99. Li, J.; Shi, P.-J.; Han, X.-Y.; Meng, K.; Yang, P.-L.; Wang, Y.-R.; Luo, H.-Y.; Wu, N.-F.; Yao, B.; Fan, Y.-L. Functional expression of the keratinolytic serine protease gene *sfp2* from *Streptomyces fradiae* var. k11 in *Pichia pastoris*. *Protein Expres. Purif.* **2007**, *54*, 79–86. [CrossRef] [PubMed]

100. Benkiar, A.; Nadia, Z.J.; Badis, A.; Rebzani, F.; Soraya, B.T.; Rekik, H.; Naili, B.; Ferradji, F.Z.; Bejar, S.; Jaouadi, B. Biochemical and molecular characterization of a thermo- and detergent-stable alkaline serin keratinolytic protease from *Bacillus circulans* strain DZ100 for detergent formulations and feather-biodegradation process. *Int. Biodeterior. Biodegradation* **2013**, *83*, 129–138. [CrossRef]

101. Fang, Z.; Zhang, J.; Liu, B.; Jiang, L.; Du, G.; Chen, J. Cloning, heterologous expression and characterization of two keratinases from *Stenotrophomonas maltophilia* BBE11-1. *Proc. Biochem.* **2014**, *49*, 647–654. [CrossRef]

102. Jaouadi, B.; Ellouz-Chaabouni, S.; Rhimi, M.; Bejar, S. Biochemical and molecular characterization of a detergent-stable serine alkaline protease from Bacillus pumilus CBS with high catalytic efficiency. *Biochimie* **2008**, *90*, 1291–1305. [CrossRef]

103. Rajput, R.; Sharma, R.; Gupta, R. Cloning and characterization of a thermostable detergent-compatible recombinant keratinase from *Bacillus pumilus* KS12. *IUBMB* **2011**, *58*, 109–118.

104. Jaouadi, N.Z.; Rekik, H.; Elhoul, M.B.; Rahem, F.Z.; Hila, C.G.; Aicha, H.S.B.; Badis, A.; Toumi, A.; Bejar, S.; Jaouadi, B. A novel keratinase from Bacillus tequilensis strain Q7 with promising potential for the leather bating process. *Int. J. Biol. Macromol.* **2015**, *79*, 952–964. [CrossRef]

105. Ghosh, A.; Chakrabarti, K.; Chattopadhyay, D. Cloning of feather-degrading minor extracellular protease from Bacillus cereus DCUW: dissection of the structural domains. *Microbiology* **2009**, *155*, 2049–2057. [CrossRef]

106. Wang, L.; Cheng, G.; Ren, Y.; Dai, Z.; Zhao, Z.-S.; Liu, F.; Li, S.; Wei, Y.; Xiong, J.; Tang, X.-F.; et al. Degradation of intact chicken feathers by *Thermoactinomyces* sp. CDF and characterization of its keratinolytic protease. *Appl Microbiol. Biotechnol.* **2015**, *99*, 3949–3959. [CrossRef]

107. Wu, W.-L.; Chen, M.-Y.; Tu, I.-F.; Lin, Y.-C.; Kumar, N.E.; Chen, M.-Y.; Ho, M.-C.; Wu, S.-H. The discovery of novel heat-stable keratinases from *Meiothermus taiwanensis* WR-220 and other extremophiles. *Sci. Rep.* **2017**, *7*, 4658–4669. [CrossRef]

108. Liang, X.; Bian, Y.; Tang, X.-F.; Xiao, G.; Tang, B. Enhancement of keratinolytic activity of a thermophilic subtilase by improving its autolysis resistance and thermostability under reducing conditions. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 999–1006. [CrossRef]

109. Gegeckas, A.; Gudiukaite, R.; Debski, J.; Citavicius, D. Keratinous waste decomposition and peptide production by keratinase from *Geobacillus stearothermophilus* AD-11. *Int. J. Biol. Macromol.* **2015**, *75*, 158–165. [CrossRef] [PubMed]

110. Sharma, R.; Murty, N.A.R.; Gupta, R. Molecular characterization of N-terminal pro-sequence of keratinase ker P from *Pseudomonas aeruginosa*: identification of region with chaperone activity. *Appl. Biochem. Biotechnol.* **2011**, *165*, 892–901. [CrossRef] [PubMed]

111. Mitsuiki, S.; Ichikawa, M.; Oka, T.; Sakai, M.; Moriyama, Y.; Sameshima, Y.; Goto, M.; Furukawa, K. Molecular characterization of a keratinolytic enzyme from an alkaliphilic *Nocardiopsis* sp. TOA-1. *Enzyme Microb. Technol.* **2004**, *34*, 482–489. [CrossRef]

112. Ghosh, A.; Chakrabarti, K. Degradtion of raw feather by a novel high molecular weight extracellular protease from newly isolated *Bacillus cereus* DCUW. *J. Ind. Microbiol. Biotechnol.* **2008**, *35*, 825–834. [CrossRef]

113. Tripathi, L.P.; Sowdhamini, R. Genome-wide survey of prokaryotic serine proteases: analysis of distribution and domain architectures of five serine protease families in prokaryotes. *BMC Genomics* **2008**, *9*, 549–576. [CrossRef]

114. Rawlings, N.D.; Waller, M.; Barrett, A.J.; Bateman, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **2014**, *42*, D503–D509. [CrossRef]

115. Siezen, R.J. Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. *Protein Eng.* **1991**, *4*, 719–737. [CrossRef]

116. Page, M.J.; Di Cera, E. Serine peptidases: classification, structure and function. *Cell Mol. Life Sci.* **2008**, *65*, 1220–1236. [CrossRef]

117. Donlon, J.; Polaina, J.; MacCabe, A.P. (Eds.) *Industrial Enzymes*; Springer: Dordrecht, Germany, 2007.

118. Jongeneel, C.V.; Bouvier, J.; Bairoch, A. A unique signature identifies a family of zinc-dependent metallopeptidases. *FEBS Lett.* **1989**, *242*, 211–214. [CrossRef]

119. De Kreij, A.; Venema, G.; Van den Burg, B. Substrate specificity in the highly heterogeneous M4 peptidase family is determined by a small subset of amino acids. *JBC* **2000**, *275*, 31115–31120. [CrossRef]

120. Gregg, K.A. A comparison of genomic coding sequences for feather and scale keratins: Structural and evolutionary implications. *EMBO J.* **1984**, *3*, 175–178. [CrossRef] [PubMed]

121. Gradisar, H.; Friedrich, J.; Krizaj, I.; Jerala, R. Similarities and specificities of fungal keratinolytic proteases: comparison of keratinases of *Paecilomyces marquandii* and *Doratomyces microsporus* to some known proteases. *Appl. Environ. Microbiol.* **2005**, *71*, 3420–3426. [CrossRef]

122. Brandelli, A.; Daroit, D.J.; Riffel, A. Biochemical features of microbial keratinases and their production and applications. *Appl Microbiol. Biotechnol.* **2010**, *85*, 1735–1750. [CrossRef] [PubMed]

123. Di Cera, E. Serine proteases. *IUBMB Life* **2009**, *61*, 510–515. [CrossRef] [PubMed]

124. Wu, J.W.; Chen, X.L. Extracellular metalloproteases from bacteria. *Appl. Microbiol. Biotechnol.* **2011**, *92*, 253–262. [CrossRef]

125. Wainwright, M. A new method for determining the microbial degradation of keratin in soils. *Experientia* **1982**, *38*, 243–244. [CrossRef]

126. Riffel, A.; Lucas, F.; Heeb, P.; Brandelli, A. Characterization of a new keratinolytic bacterium that completely degrades native feather keratin. *Arch. Microbiol.* **2003**, *179*, 258–265. [CrossRef]

127. Navone, L.; Speight, R. Understanding the dynamics of keratin weakening and hydrolysis by proteases. *PLoS ONE* **2018**, *13*, e0202608. [CrossRef]

128. Farag, A.M.; Hassan, M.A. Purification, characterization and immobilization of a keratinase from *Aspergillus orizae*. *Enzyme Microb. Technol.* **2004**, *34*, 85–93. [CrossRef]

129. Grazziotin, A.; Pimentel, F.A.; De Jong, E.V.; Brandelli, A. Nutritional improvement of feather protein by treatment with microbial keratinase. *Animal Feed Sci. Technol.* **2006**, *126*, 135–144. [CrossRef]

130. Gill, S.C.; Von Hippel, P.H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **1989**, *182*, 319–326. [CrossRef]

131. Daroit, D.J.; Corrêa, A.P.F.; Brandelli, A. Keratinolytic potential of a novel *Bacillus* sp. P45 isolated from the Amazon basin fish *Piaractus mesopotamicus*. *Int. Biodeter. Biodegr.* **2009**, *63*, 358–363. [CrossRef]

132. Pereira, J.Q.; Lopes, F.C.; Petry, M.V.; Da Costa Medina, L.F.; Brandelli, A. Isolation of three novel Antarctic psychrotolerant feather-degrading bacteria and partial purification of keratinolytic enzyme from Lysobacter sp. A03. *Int. Biodeter. Biodegr.* **2014**, *88*, 1–7. [CrossRef]

133. Wawrzkiewicz, K.; Wolski, T.; Lobarzewski, J. Screening the keratinoltic activity of dermatophytes in vitro. *Mycopathologia* **1991**, *114*, 1–8. [CrossRef] [PubMed]

134. Bressollier, P.; Letourneau, F.; Urdaci, M.; Verneuil, B. Purification and characterization of a keratinolytic serine proteinase from *Streptomyces albidoflavus*. *Appl. Environ. Microbiol.* **1999**, *65*, 2570–2576. [CrossRef] [PubMed]

135. Lin, S.J.; Yoshimura, E.; Sakai, H.; Wakagi, T.; Matsuzawa, H. Weakly bound calcium ions involved in the thermostability of aqualysin I, a heat-stable subtilisin-type protease of *Thermus aquaticus* YT-1. *BBA-Protein Struct. M.* **1999**, *133*, 132–138. [CrossRef]

136. Meng, K.; Li, J.; Cao, Y.; Shi, P.; Wu, B.; Han, X.; Bai, Y.; Wu, N.-F.; Yao, B. Gene cloning and heterologous expression of a serine protease from *Streptomyces fradiae* var.k11. *Can. J. Microbiol.* **2007**, *53*, 186–195. [CrossRef]

137. Öztürk, N.Ç.; Kazan, D.; Denizci, A.A.; Erarslan, A. The influence of copper on alkaline protease stability toward autolysis and thermal inactivation. *Eng. Life Sci.* **2012**, *12*, 662–671. [CrossRef]

138. Walker, J.M. *The Proteins Protocols Handbook*, 2nd ed.; Humana Press: Totowa, NJ, USA, 1996.

139. Fershi, A.R. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; Freeman W. H. & Co.: New York, NY, USA, 1999.

140. Kunugi, S.; Tanaka, N. Cold denaturation of proteins under high pressure. *Biochim. Biophys. Acta* **2002**, *1595*, 309–314. [CrossRef]

141. Salvi, G.; De Los Rios, P.; Vendruscolo, M. Effective interactions between chaotropic agents and proteins. *Proteins Struct. Funct. Bioinf.* **2005**, *61*, 429–499. [CrossRef]

142. Suzuki, Y.; Tsujimoto, Y.; Matsui, H.; Watanabe, K. Decomposition of extremely hard-to-degrade animal proteins by thermophilic bacteria. *J. Biosci. Bioeng.* **2006**, *102*, 73–81. [CrossRef] [PubMed]

143. Macedo, A.J.; Beys da Silva, W.O.; Termignoni, C. Properties of a non collagen-degrading *Bacillus subtilis* keratinase. *Can. J. Microbiol.* **2008**, *54*, 180–188. [CrossRef] [PubMed]

144. Gupta, R.; Rajput, R.; Sharma, R.; Gupta, N. Biotechnological applications and prospective market of microbial keratinases. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 9931–9940. [CrossRef] [PubMed]

145. Navone, L.; Speight, R.E. Enzyme systems for effective dag removal from cattle hides. *Anim. Prod. Sci.* **2019**, *59*, 1387–1398. [CrossRef]

146. Alba, A.C.; Strauch, T.A.; Keisler, D.H.; Wells, K.D.; Kesler, D.C. Using a keratinase to degrade chicken feathers for improved extraction of glucocorticoids. *Gen. Comp. Endocr.* **2019**, *270*, 35–40. [CrossRef]

147. Navone, L.; Moffitt, K.; Hansen, K.-A.; Blinco, J.; Payne, A.; Speight, R. Closing the textile loop: enzymatic fibre spearation and recycling of wool/polyester fabric blends. *J. Waste Manag.* **2020**, *102*, 149–160. [CrossRef]

148. El Boushy, A.R.Y.; Van der Poel, F.B. Poultry by-Products. In *Handbook of Poultry Feed from Waste*; Springer: Dordrecht, Germany, 2000; pp. 90–152. [CrossRef]

149. Sundaram, M.; Legadavi, R.; Banu, N.A.; Gayathri, V.; Palanisammy, A. A study of antibacterial activity of keratin nanoparticles from chicken feather waste against *Staphylococcus aureus* (Bovine mastitis bacteria) and its antioxidant activity. *Eur. J. Biotechnol. Biosci.* **2015**, *6*, 1–5.

150. Ohba, R.; Deguchi, T.; Kishikawa, M.; Arsyad, F.; Morimura, S.; Kida, K. Physiological function of enzymatic hydrolysates of collagen or keratin contained in livestock or fish waste. *Food Sci. Technol. Res.* **2003**, *9*, 91–93. [CrossRef]

151. Li, L.; Wang, W.; Shi, B.-H.; Li, J.-C.; Zhang, R.-Z.; Liu, S.-T.; Chen, R.-M.; Gao, W.-H.; Chen, G.-R.; Zheng, Y.-Q.; et al. *Anti-Inflammatory Activity of Antelope Horn Keratin and its Tryptic Hydrolysate*; Nutrition Press Inc.: Trumbull, CT, USA, 1999.

152. Kelly, R.; Ellis, G.; Macdonald, R.; McPherson, R.; Middlewood, P.; Nuthall, M.; Rao, G.-F.; Roddick-Lanzilotta, A.; Sigurjonsson, G.; Singleton, D. Keratin and Soluble Derivatives Thereof for a Nutraceutical and to Reduce Oxidative Stress and to Reduce Inflammation and to Promote Skin Health. U.S. Patent 0065506, 22 March 2007.

153. Kelly, R.J.; Ellis, G.D.; Macdonald, R.J.; McPherson, R.A.; Middlewood, P.G.; Nuthall, M.G.; Rao, G.-F.; Roddick-Lanzilotta, A.D.; Sigurjonsson, G.F.; Singleton, D.J. Nutraceutical Composition Comprising Soluble Keratin or Derivatives Thereof. U.S. Patent 7579317, 25 August 2009.

154. Cutler, P. *Protein Purification Protocols*; Humana Press: Totowa, NJ, USA, 2004.

155. Zeng, W.-C.; Zhang, W.-C.; Zhang, W.-H.; Shi, B. Antioxidant activity and characterization of bioactive polypeptides from bovine hair. *Funct. Polym.* **2013**, *73*, 573–578. [CrossRef]

156. Jones, L.N.; Sinclair, R.D.; Ecroyd, H.; Lui, Y.; Bennett, L.E. Bioprospecting keratinous material. *Int. J. Trichol.* **2010**, *2*, 47–49. [CrossRef]

157. Fontoura, R.; Daroit, D.J.; Correa, A.P.F.; Meira, S.M.M.; Mosquera, M.; Brandelli, A. Production of feather hydrolysates with antioxidant, angiotensin-1 converting enzyme- and dipeptidyl peptidase-IV-inhibitory activities. *New Biotechnol.* **2014**, *31*, 506–513. [CrossRef] [PubMed]

158. Yeo, I.; Lee, Y.-J.; Song, K.; Jin, H.-S.; Lee, J.-E.; Kim, D.; Lee, D.-W.; Kang, N.J. Low molecular weight keratins with anti-skin aging activity produced by anaerobic digestion of poultry feathers with Fervidobacterium islandicum AW-1. *J. Biotechnol.* **2018**, *271*, 17–25. [CrossRef]

159. Jin, H.-S.; Song, K.; Baek, J.-H.; Lee, J.-E.; Kim, D.J.; Nam, G.-W.; Kang, N.J.; Lee, D.-L. Identification of matrix metalloproteinasae-1-suppressive peptides in feather hydrolysates. *J. Agric. Food Chem.* **2018**, *66*, 12719–12729. [CrossRef] [PubMed]

160. Jin, H.-S.; Park, S.Y.; Kim, J.-Y.; Lee, J.-E.; Lee, H.-S.; Kang, N.J.; Lee, D.-W. Fluorescence-based quantification of bioactive keratin peptides from feathers for optimising large-scale anaerobic fermentation and purification. *Biotechnol. Bioproc. Eng.* **2019**, *2*, 240–249. [CrossRef]

161. Rouse, J.G.; Van Dyke, M.E. A review of keratin-based biomaterials for biomedical applications. *Materials* **2010**, *3*, 999–1014. [CrossRef]

162. Tachibana, A.; Furuta, Y.; Takesima, H.; Tanabe, T.; Yamauchi, K. Fabrication of wool keratin sponge scaffolds for long-term cell cultivation. *J. Biotechnol.* **2002**, *93*, 165–170. [CrossRef]

163. Tang, L.; Ollague, S.J.; Kelly, R.; Kirsner, R.S.; Li, J. Wool-derived keratin stimulates uman keratinocyte migration and types IV and VIIcollagen expression. *Exp. Dermatol.* **2012**, *21*, 458–460. [CrossRef]

164. Loan, F.; Marsh, C.; Cassidy, S.; Simcock, J. Keratin-based products for effective wound care management in superficial and partial thickeness burns injuries. *Burns* **2016**, *4*, 541–547. [CrossRef]

165. Sanchez Ramirez, D.O.; Carletto, R.A.; Tonetti, C.; Giachet, F.T.; Varesano, A.; Vineis, C. Wool keratin film plasticized by citric acid for food packaging. *Food Pack. Shelf Life* **2017**, *12*, 100–110. [CrossRef]

Article

Characterisation of the First Archaeal Mannonate Dehydratase from *Thermoplasma acidophilum* and Its Potential Role in the Catabolism of D-Mannose

Dominik Kopp [1], Robert Willows [1,2] and Anwar Sunna [1,2,*

[1]   Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia;
      dominik.kopp@hdr.mq.edu.au (D.K.); robert.willows@mq.edu.au (R.W.)

[2]   Biomolecular Discovery and Design Research Centre, Macquarie University,
      Sydney, New South Wales 2109, Australia

*   Correspondence: anwar.sunna@mq.edu.au; Tel.: +61-2-9850-4220

Abstract: Mannonate dehydratases catalyse the dehydration reaction from mannonate to 2-keto-3-deoxygluconate as part of the hexuronic acid metabolism in bacteria. Bacterial mannonate dehydratases present in this gene cluster usually belong to the xylose isomerase-like superfamily, which have been the focus of structural, biochemical and physiological studies. Mannonate dehydratases from archaea have not been studied in detail. Here, we identified and characterised the first archaeal mannonate dehydratase (TaManD) from the thermoacidophilic archaeon *Thermoplasma acidophilum*. The recombinant TaManD enzyme was optimally active at 65 °C and showed high specificity towards D-mannonate and its lactone, D-mannono-1,4-lactone. The gene encoding for TaManD is located adjacent to a previously studied mannose-specific aldohexose dehydrogenase (AldT) in the genome of *T. acidophilum*. Using nuclear magnetic resonance (NMR) spectroscopy, we showed that the mannose-specific AldT produces the substrates for TaManD, demonstrating the possibility for an oxidative metabolism of mannose in *T. acidophilum*. Among previously studied mannonate dehydratases, TaManD showed closest homology to enzymes belonging to the xylose isomerase-like superfamily. Genetic analysis revealed that closely related mannonate dehydratases among archaea are not located in a hexuronate gene cluster like in bacteria, but next to putative aldohexose dehydrogenases, implying a different physiological role of mannonate dehydratases in those archaeal species.

Keywords: mannonate dehydratase; mannose metabolism; *Thermoplasma acidophilum*; mannono-1,4-lactone; 2-keto-3-deoxygluconate; aldohexose dehydrogenase

1. Introduction

Mannonate dehydratases (EC 4.2.1.8) catalyse the conversion of mannonate to 2-keto-3-deoxygluconate (KDG) and have been studied as part of the hexuronic acid gene cluster in several organisms, such as *Escherichia coli*, *Bacillus stearothermophilus*, *Bacillus subtilis* and *Erwinia chrysanthemi* [1–4]. The hexuronate gene cluster encodes enzymes involved in the metabolism of glucuronate and galacturonate [5]. Glucuronate is a common sugar acid present in glucuronoxylan, a constituent of plant cell walls which can serve as the only carbon source for growth of some bacteria [3,6,7]. Glucuronate is also present in the mucus layer of mammals, providing a carbon source for anaerobic gut bacteria, such as *E. coli* [8,9]. As part of the hexuronate metabolism, mannonate dehydratase converts mannonate to KDG (Figure 1A). In *E. coli*, and some species of *Erwinia*, KDG is phosphorylated and cleaved into pyruvate and 3-phosphoglycerate, which can be further metabolised in the tricarboxylic acid cycle, or the Entner-Doudoroff pathway [4,5].

Mannonate dehydratases are represented in different enzyme families, such as the xylose isomerase-like superfamily or the enolase superfamily. Mannonate dehydratases encoded in the bacterial hexuronate gene clusters usually belong to the xylose isomerase-like superfamily. Crystal structures have been solved for xylose isomerase-like mannonate dehydratases from *Streptococcus suis*, *E. coli* K12 and *Enterococcus faecalis* in native form and in complex with Mn^{2+} ions [10,11]. His311 and Tyr325 in the binding pocket were identified as crucial for the activity of the mannonate dehydratases from Gram-positive bacteria, such as *S. suis*. However, in Gram-negative bacteria (e.g., *E. coli* K12), an additional inserted sequence in the binding pocket rendered the dehydratase less active [11]. Members of the enolase superfamily show a conserved structure and reaction mechanism, but differ in their physiological functions. Within their conserved barrel structure, mannonate dehydratases of this family share conserved ligand-binding sites for Mg^{2+}, which are essential for the stabilisation of the enediolate intermediate [12]. Among mannonate dehydratases of the enolase superfamily, some representatives with diverse functions have been found, which are involved in the catabolism of sugar acids other than glucuronate or galacturonate [13]. Several crystal structures have been solved for mannonate dehydratases from the enolase superfamily, including enzymes from *Chromohalobacter salexigens* and *Novosphingobium aromaticivorans* [12,14]. However, to the best of our knowledge, no archaeal mannonate dehydratase has been investigated so far.

Here, we used genomic analysis to identify the first functional mannonate dehydratase from the thermoacidophilic archaeon *Thermoplasma acidophilum* (TaManD). The archaeal mannonate dehydratase was recombinantly-expressed in *E. coli*, purified and functionally characterised. TaManD showed high amino acid sequence identity to bacterial mannonate dehydratases from the xylose isomerase-like superfamily. In the genome of *T. acidophilum*, the gene encoding for TaManD is located adjacent to an aldohexose dehydrogenase (AldT), which has been shown previously to catalyse the oxidation of mannose to mannonate with high specificity [15,16]. However, the physiological significance of the oxidation and its products were not investigated further. We identified the products of an AldT-mediated oxidation of D-mannose using NMR spectroscopy and confirmed that TaManD is able to convert the products to KDG. This demonstrates that in principle, a mannose metabolism based on AldT and ManD is possible in *T. acidophilum* (Figure 1B).
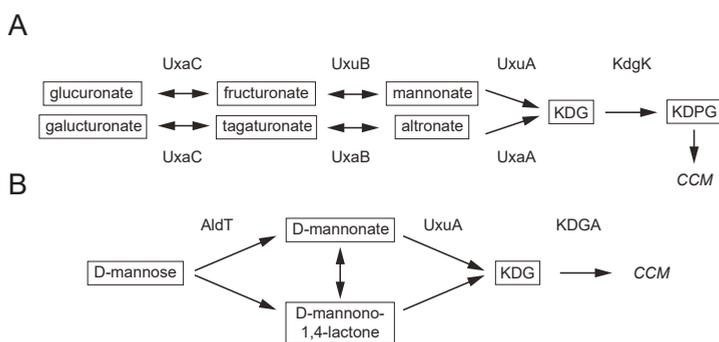
Figure 1. Hexuronate metabolism in *E. coli* and possible mannose catabolism in *T. acidophilum*. (**A**) Role of *uxuA* mannnonate dehdyratase in dissimilation of hexuronates in *E. coli* adapted from Peekhaus and Conway [8]. (**B**) Role of *uxuA* mannonate dehydratase in a possible mannose metabolism in *T. acidophilum*. UxaC: hexuronate isomerase, UxuB: mannonate oxidoreductase, UxaB: altronate oxidoreductase, UxuA: mannonate dehydratase, UxaA: altronate dehydratase, KdgK: 2-keto-3-deoxygluconate kinase, KDG: 2-keto-3-deoxygluconate, KDPG: 2-keto-3-deoxy-6-phosphogluconate, AldT: aldose dehydrogenase, KDGA: 2-keto-3-deoxygluconate aldolase, CCM: central carbon metabolism.

2. Results and Discussion

2.1. Screening for a Functional Mannonate Dehydratase in T. acidophilum

Based on the previously identified mannose-specific AldT, we searched the genomic neighbourhood of its gene locus (Ta0754) for enzymes which could display activity towards mannonate. In the proximity of Ta0754, several genes are located that encode for hypothetical dehydratases, suggesting their potential to convert mannonate to KDG (Table S1). We identified the protein product of the Ta0753 gene as a functional mannonate dehydratase from *T. acidophilum* by heterologous expression in *E. coli*, and compared its properties and genomic context to other previously studied mannonate dehydratases.

Among previously characterised mannonate dehydratases, TaManD shares highest protein sequence identity (31.5%) with enzymes in the xylose isomerase-like superfamily, such as the mannonate dehydratase from *S. suis* [10]. Key amino acid residues for substrate binding (His311 and Tyr325) and binding of the cofactor Mn^{2+} (Asp310, Cys237, His199 and His266) in *S. suis* are conserved in the amino acid sequence of TaManD. A much higher amino acid sequence identity (62.9%) is shared between TaManD and putative mannonate dehydratases from closely related archaeal species, *Ferroplasma acidarmanus* and *Ferroplasma acidiphilum*, suggesting an archaeal clade of mannonate dehydratases.

Phylogenetic analysis showed that mannonate dehydratases from *T. acidophilum*, *F. acidarmanus* and *F. acidiphilum* are more closely related to bacterial mannonate dehydratases from the xylose isomerase-like superfamily than to bacterial or archaeal mannonate dehydratases of the enolase superfamily (Figure 2). Mannonate dehydratases of the enolase superfamily have a substantially different structure from those in the xylose isomerase-like superfamily, and therefore, are only distantly related to TaManD. Despite fulfilling a similar function, the dihydroxy-acid dehydratase from *S. solfataricus*, which belongs to the IlvD/EDD superfamily, is rather unrelated to the xylose isomerase-like and enolase superfamilies [17].

For bacterial xylose isomerase-like mannonate dehydratases, a physiological function of the enzyme in the catabolism of hexuronates has been demonstrated [18,19]. Although TaManD is annotated as *uxuA* mannonate dehydratase, which implies a role in the metabolism of hexuronates, it is not present in a classical hexuronate gene cluster known from bacteria such as *E. coli*, *B. subtilis* or *B. stearothermophilus* (Figure 3) [2,3]. In contrast, the gene encoding for TaManD is located adjacent to AldT in the genome of *T. acidophilum*. Similarly, the two closely related putative mannonate dehydratases in *F. acidarmanus* and *F. acidiphilum* are also located adjacent to putative aldohexose dehydrogenases. Therefore, a different physiological role can be proposed for these mannonate dehydratases. In the following, we functionally characterise TaManD and show that AldT is able to produce the substrate needed for a subsequent conversion to KDG mediated by TaManD.
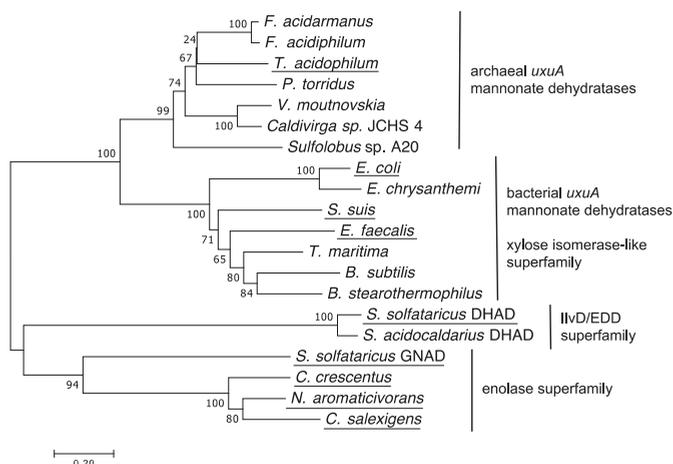
Figure 2. Phylogenetic relationship of different putative and confirmed dehydratases belonging to different enzyme families. Characterised enzymes are underlined. Evolutionary analyses were conducted in MEGA 7. 20 protein sequences were aligned using MUSCLE. The phylogenetic tree was inferred using the neighbour-joining method. The scale bar indicates an evolutionary distance of 0.20 nucleotide per position in the sequence. The number next to the nodes represent bootstrap confidence values estimated from 500 replicates. Protein sequences were retrieved from the Uniprot database and their entry codes are as follows. *Ferroplasma acidarmanus*: S0AL33, *Ferroplasma acidiphilum*: A0A1V0N416, *Thermoplasma acidophilum*: Q9HK52, *Picrophilus torridus*: Q6L2R9, *Vulcanisaeta moutnovskia*: F0QYL3, *Caldivirga* sp. JCHS 4: A0A101XEY7, *Sulfolobus* sp. A20: A0A1C8ZTN0, *Escherichia coli* K12: P24215, *Erwinia chrysanthemi/Dickeya dadantii*: E0SEP1, *Streptococcus suis*: A0A142UME2, *Enterococcus faecalis*: Q82ZC9, *Thermotoga maritima*: Q9WXS4, *Bacillus subtilis*: O34346, *Bacillus stearothermophilus*: A0A087LHB3, *Sulfolobus solfataricus* DHAD: Q97UB2, *Sulfolobus acidocaldarius* DHAD: Q4J860, *Sulfolobus solfataricus* GNAD: Q97U27, *Caulobacter crescentus/Caulobacter vibroides*: Q9AAR4, *Novosphingobium aromaticivorans*: A4XF23, *Chromohalobacter salexigens*: Q1QT89.
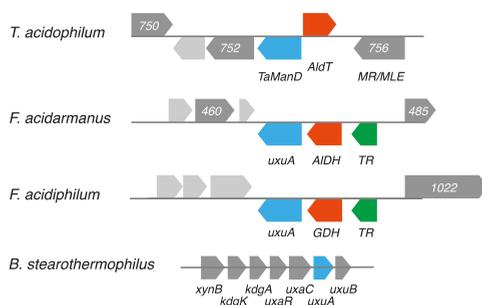
Figure 3. Schematic maps of uxuA mannonate dehydratases and their genomic neighbourhoods in *T. acidophilum*, *F. acidarmanus*, *F. acidiphilum* and *B. stearothermophilus*. Genes are annotated in the NCBI database as follows: In blue, *TaManD/uxuA*: *T. acidophilum* mannonate dehydratase/uxuA D-mannonate dehydratases; in orange, *AldT/AlDH/GDH*: aldohexose dehydrogenase/aldose dehydrogenase/glucose-1-dehydrogenase; in green, TR: transcriptional regulator; in light grey, annotated as hypothetical protein. Note: 750: FAA hydrolase family protein, 752: L-rhamnonate dehydratase, 756: mandelate racemase/muconate lactonizing enzyme family protein, 460: transposase, 485: acetyl-coenzyme A synthetase, 1022: acetate-CoA ligase, *xynB*: β-xylosidase, *kdgK*: 2-keto-3-deoxygluconate kinase, *kdgA*: 2-keto-3-deoxy-6-phosphogluconate aldolase, *uxaR*: regulatory protein, *uxaC*: hexuronate isomerase, *uxuB*: mannonate oxidoreductase.

2.2. Substrate Conformation for TaManD Activity

Before acquiring enzyme kinetics and studying the characteristics of TaManD, we investigated which substrate conformation is encountered by TaManD under physiological conditions. In aqueous solutions, free sugar acids coexist with their lactone in an equilibrium, which is defined by the stability of the lactone, the temperature and the pH of the solution. Many lactones hydrolyse spontaneously in water, although several lactones, including D-arabinonolactone [20], L-rhamnonolactone [21] and D-xylonolactone [22], have been reported to be hydrolysed by lactonases. Mannonate is able to form a lactone by covalent bond formation between carbon 1 and carbon 5 (δ-lactone) or carbon 1 and carbon 4 (γ-lactone) (Figure 4A). The equilibrium of the two different envelope forms has been studied by NMR spectroscopy and it was found that the γ-lactone is strongly favoured over the δ-lactone [23]. However, these studies do not describe equilibria between lactone and free acid form in a physiological buffer. In a cellular environment mannonate can either be produced by a mannonate oxidoreductase (UxuB) as part of the catabolism of hexuronic acids or by an aldohexose dehydrogenase, such as AldT, in a hypothetical oxidative mannose catabolism (Figure 1). In order to study the conformation of the substrate for TaManD in a physiological environment, we first acquired decoupled ^{13}C NMR spectra for D-mannono-1,4-lactone in physiological buffer at pH 7, in its lactone form (incubated with HCl) and after hydrolysis to the free sugar acid (incubated with NaOH) (Figure 4B). The spectra obtained indicated that the lactone and the free sugar acid can be distinguished by their chemical shifts. In the free sugar acid form, carbon 6 yields a chemical shift at 63.51 ppm, whereas in D-mannono-1,4-lactone, carbon 6 displays a chemical shift at 63.02 ppm. D-mannono-1,4-lactone in buffer shows both chemical shifts, indicating that D-mannono-1,4-lactone and D-mannonate are present in an equilibrium at pH 7.
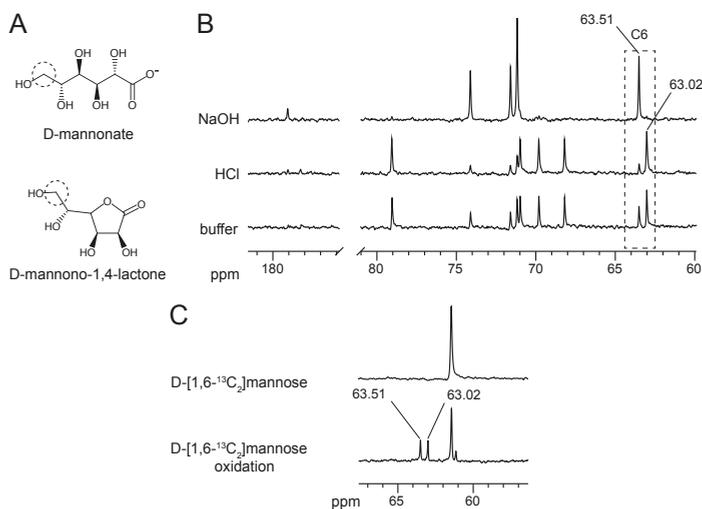
Figure 4. Analysis of D-mannono-1,4-lactone under different conditions and products of AldT-mediated oxidation of isotope labelled D-mannose. (**A**) Chemical structures of the free sugar acid, D-mannonate, and its lactone, D-mannono-1,4-lactone. Carbon 6, which was used to differentiate the two different substrate forms in NMR spectra, is circled for both structures. (**B**) Decoupled 1D ^{13}C NMR spectra of D-mannono-1,4-lactone after incubation in 1 M NaOH, 1 M HCl and in 0.1 M sodium phosphate (NaP) buffer pH 7. Chemical shift for carbon 6 is shown by a dashed rectangle. Characteristic chemical shifts for the lactone (63.02 ppm) and for the free sugar acid (63.51 ppm) are indicated. (**C**) Decoupled 1D ^{13}C NMR spectra before (**top**) and after (**bottom**) oxidation of labelled D-mannose with AldT. Chemical shifts of carbon 6 for lactone and free sugar acid are indicated.

Next, we oxidised isotope-labelled D-[1,6-^{13}C$_2$]mannose using recombinantly-expressed AldT and acquired 1D ^{13}C NMR spectra of the reaction product. The chemical shifts for carbon 6 of mannono-1,4-lactone and D-mannonate were used to analyse the reaction products of the mannose oxidation with AldT, as these were close together but well resolved and will have similar relaxation rates in both forms. The spectrum after partial AldT oxidation of D-[1,6-^{13}C$_2$]mannose (Figure 4C) shows that equal quantities of D-mannono-1,4-lactone and D-mannonate are produced for the subsequent dehydration by TaManD.

2.3. Expression and Purification of TaManD

For the further biochemical characterisation, TaManD was expressed together with a tobacco etch virus (TEV) sequence, which allowed the proteolytic cleavage of the His-tag after the first metal affinity purification. In order to study the effect of bivalent metal cofactors on enzyme activity, cleavage of the His-tag from the recombinant enzyme was necessary, since cations (Ni^{2+}, Mg^{2+}) might influence enzyme activity assays of a His-tagged enzyme. After His-tag cleavage, the recombinant enzyme showed a single band on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) with an apparent mass of 38 kDa (Figure S1). Peptide mass fingerprinting was performed with liquid chromatography-electrospray ionisation-tandem mass spectrometry (LC ESI MS/MS) and the National Center for Biotechnology Information (NCBI) database searches using the mascot software confirmed the identity of the purified protein as uxuA mannonate dehydratase from *T. acidophilum*. In addition, the apparent native molecular mass of the protein was estimated from size exclusion chromatography to be 225 kDa, indicating a hexameric structure of the protein.

2.4. Enzyme Characterisation

Initial tests showed that TaManD lost almost its complete activity after purification, suggesting that essential cofactors for activity were removed during the purification process. Accordingly, the effect of several additives was tested for their influence on the activity of TaManD with D-mannono-1,4-lactone. TaManD showed strongly enhanced activity in the presence of β-mercaptoethanol and the metal ions Mn^{2+}, Ni^{2+}, Mg^{2+}, Ca^{2+} and Co^{2+} (Figure 5). The highest activation (100%) was observed in the presence of β-mercaptoethanol, which was significantly higher than the activating bivalent cation Ni^{2+} (71.5% of maximum enzyme activity, p-value: 0.0022). Mn^{2+}, Mg^{2+} and Co^{2+} also activated TaManD and resulted in 50–60% increase in maximum activity when compared to the TaManD control (without any additive). No significant difference in activity was observed in the presence of Fe^{2+}, Zn^{2+}, Cu^{2+}, EDTA, DTT and glutathione.
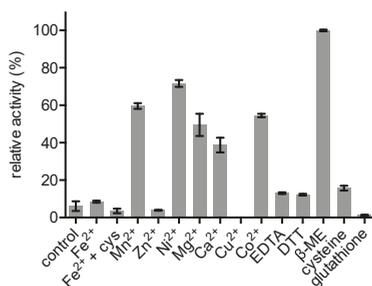
Figure 5. Effect of metal ions, chelating and reducing agents on the activity of TaManD. Purified enzyme (0.6 µg) was pre-incubated for 1 h with 1 mM of each additive in 50 mM HEPES pH 7. Activity was then determined in reactions containing 10 mM D-mannono-1,4-lactone in 50 mM HEPES pH 7 incubated for 1 h at 55 °C before analysis with the semicarbazide assay. Activity is expressed in relation to the maximum enzyme activity. Fe^{2+} + cys: Fe^{2+} was prepared with 1 mM cysteine. DTT: dithioerythritol, EDTA: ethylendiaminetetraacetate, β-ME: β-mercaptoethanol.

TaManD was activated by similar bivalent ions and reducing factor, as previously studied bacterial mannonate dehydratases from the xylose isomerase-like superfamily. However, none of the bacterial mannonate dehydratases described so far have been shown to have an equally strong activation with both metal ions, Co^{2+} and Ni^{2+}, as observed with TaManD. Early investigations of mannonate dehydratases focused on *E. coli* and showed different strengths in activation after incubation with bivalent cations [1,24–26]. In summary, most studies showed that β-mercaptoethanol, Fe^{2+} and Mn^{2+} resulted in highest enzyme activity (90–100% of maximum enzyme activity). Incubation with other bivalent cations resulted in less activity compared to the enzyme's maximum activity (Co^{2+}: 27–80%, Ni^{2+}: 5–40% and Zn^{2+}: 5–25%, depending on the study). More recently, crystal structures of xylose isomerase-like mannonate dehydratases from *S. suis* and *E. coli* have been solved and revealed the presence of primarily Mn^{2+} and lower amounts of Mg^{2+}, Ni^{2+} and Zn^{2+} in their binding sites [10,11].

T. acidophilum was originally isolated from a hot and acidic environment and accordingly this archaeon displays optimal temperature and pH for growth at 59 °C and pH 1–2, respectively [27]. The purified TaManD was active between 35 °C and 70 °C, with an optimal temperature for activity at 65 °C (Figure 6A). Thermostability of the enzyme in the absence of substrate was studied at temperatures between 55 °C and 95 °C (Figure 6B). The enzyme retains its full activity at 55 °C for at least an hour, whereas complete inactivation was observed at 75 °C, 85 °C and 95 °C within 90 min, 60 min and 15 min, respectively. The difference in the relatively high optimal temperature of the enzyme to the comparably low thermostability indicated that the enzyme is more prone to inactivation in the absence of substrate. Other enzymes obtained from this organism have been shown to display optimal temperatures in the range of 55 °C to 70 °C [15,28–30].
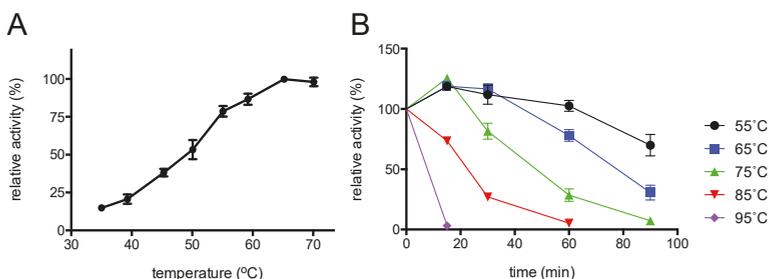
Figure 6. Effect of temperature on TaManD activity. (**A**) Optimum temperature of TaManD. (**B**) Thermostability of purified TaManD enzyme (1 µg) incubated over 90 min at various temperatures.

Following the identification of the crucial metal cofactors required for TaManD activity and determination of its optimal reaction temperature, we acquired enzyme kinetic data for the purified enzyme. Considering the fact that under physiological conditions an equilibrium of both lactone and free sugar acid exists, the kinetic data was acquired in reactions with substrate provided in these two different forms. The substrate was either hydrolysed with NaOH to obtain free sugar acid, D-mannonate, or the lactone was prepared in buffer, equaling an equilibrium of lactone and free sugar acid. TaManD displayed a higher activity in reactions with D-mannonate compared to reactions with D-mannono-1,4-lactone prepared in buffer at pH 7 (Table 1). The maximal velocity of the reaction, V_{max}, with D-mannonate as substrate was slightly higher (24–27%) compared to reactions with the lactone prepared in buffer (p-value: 0.06), whereas no difference in the affinity between the two different forms of the substrate could be observed (Table 1, Figure S2).

The pH optimum of TaManD was dependent on the form of the substrate. TaManD displayed highest activity between pH 5 and 7 with mannonate, while with D-mannono-1,4-lactone in buffer, the enzyme displayed overall lower activities with an apparent optimum at pH 7 (Figure 7). Unlike previously studied enzymes from *T. acidophilum* (e.g., AldT), TaManD did not retain maximum activity above pH 7 with both substrates [15].

Table 1. TaManD kinetic data with D-mannonate (prepared by hydrolysis with NaOH) and with D-mannono-1,4-lactone prepared in buffer (pH 7). Non-linear fitting was performed (Figure S2).

Substrate	K_m (mM)	V_{max} (U/mg)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
D-mannonate	5.37 ± 0.90	2.39 ± 0.11	1.64	0.30
D-mannono-1,4-lactone in buffer	4.90 ± 0.53	1.90 ± 0.06	1.33	0.26
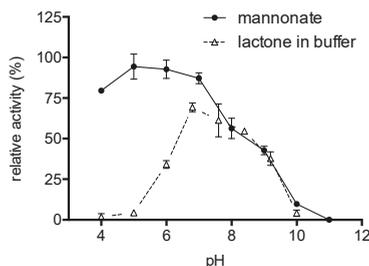
Figure 7. The pH optima of reactions with D-mannono-1,4-lactone prepared in buffer pH 7 (lactone in buffer) and with D-mannonate (prepared by hydrolysis with NaOH). Activity values are expressed as relative activity (%) normalised to the highest overall activity observed in the assay.

Based on the observations from experiments with different forms of substrate, we assume that D-mannonate is either the only form, or at least more accessible to TaManD than its lactone. At low pH, TaManD shows low activity with D-mannono-1,4-lactone, supposedly because the lactone does not hydrolyse to the free acid. In contrast, if D-mannono-1,4-lactone was hydrolysed to D-mannonate and used as substrate at low pH, high enzyme activity should be observed. Similarly, the difference in V_{max} for D-mannonate and D-mannono-1,4-lactone in buffer can be explained by slow hydrolysis of the lactone to the free acid. Reactions with hydrolysed D-mannonate occur faster compared to the lactone in buffer, since more substrate is readily available.

2.5. Product Identification and Substrate Specificity

In order to identify the product of the TaManD reaction with mannonate, reactions containing the substrate and Co^{2+} were analysed at several time points by high-performance liquid chromatography (HPLC) (Figure 8A). Although a complete separation of the product KDG and mannonate could not be achieved on different organic acid columns, an increase of KDG was observed. In order to test for substrate promiscuity, 11 different sugar acids (Figure 8B) were incubated with purified TaManD and tested for the formation of 2-keto-3-deoxy analogues using the semicarbazide assay. D-mannonate and D-mannono-1,4-lactone were the only substrates that showed a positive reaction after incubation for 16 h. This is in contrast to other dehydratases, including members of the enolase superfamily and the IlvD/EDD superfamily, such as the L-fuconate dehydratase from *Xanthomonas campestris* [31] or the dihydroxy-acid dehydratase from *Sulfolobus solfataricus* [17,32], which have been shown to be active with a multitude of different sugar acids.
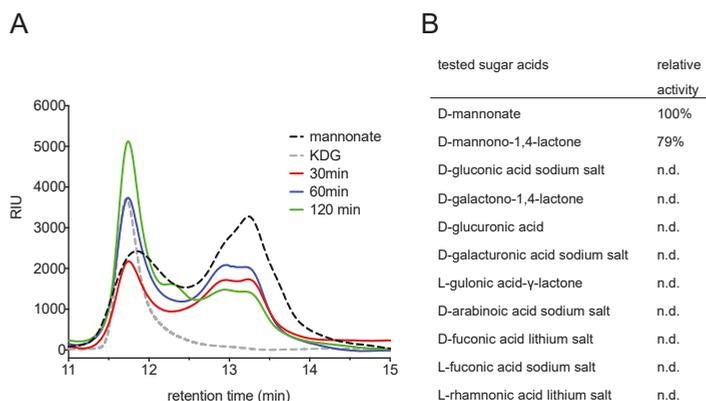
tested sugar acids	relative activity
D-mannonate	100%
D-mannono-1,4-lactone	79%
D-gluconic acid sodium salt	n.d.
D-galactono-1,4-lactone	n.d.
D-glucuronic acid	n.d.
D-galacturonic acid sodium salt	n.d.
L-gulonic acid-γ-lactone	n.d.
D-arabinoic acid sodium salt	n.d.
D-fuconic acid lithium salt	n.d.
L-fuconic acid sodium salt	n.d.
L-rhamnonic acid lithium salt	n.d.

Figure 8. Substrate specificity and product analysis of reactions with TaManD. (**A**) Reactions with TaManD and 10 mM mannonate analysed by HPLC equipped with a refractive index detector. Standards: 10 mM mannonate (dashed black line), 2.5 mM KDG (dashed grey line). Reactions: stopped after 30 min (red line), stopped after 60 min (blue line), stopped after 120 min (green line). RIU: refractive index units. (**B**) List of sugar acids that were tested as substrates for TaManD. Activity is expressed relative (%) to the maximal activity obtained in the assay; n.d.: no signal detected after 16 h of incubation. Lactones except for D-mannono-1,4-lactone were hydrolysed by incubation in 1 M NaOH for 1 h at room temperature and then diluted in 50 mM HEPES pH 7 to obtain free sugar acids. All sugar acids were tested at a concentration of 10 mM in duplicate experiments.

3. Materials and Methods

3.1. Cloning, Expression and Purification of Enzymes

All plasmids were constructed by amplification of genes from *T. acidophilum* DSM 1728 genomic DNA (DSMZ, Braunschweig, Germany) using PCR, restriction and ligation into vectors according to standard protocols [33]. Primer pairs and restriction enzymes used for the construction of each expression plasmid are summarised in Table S2. *E. coli* strain α-select (Bioline, Sydney, Australia) was used in all initial cloning procedures. The expression plasmid pProEX HTa-Ta0753 was then used to transform BL21-CodonPlus (DE3)-RIL competent cells (Agilent Technologies, Santa Clara, CA, USA). Plasmid pETDuet-1-AldT was used to transform BL21 (DE3) cells (NEB, Ipswitch, Burlington, MA, USA). BL21-CodonPlus (DE3)-RIL strains containing the respective expression plasmid were grown in 500 mL Luria–Bertani medium (LB) supplemented with carbenicillin (100 µg/ml) and chloramphenicol (35 µg/ml) at 37 °C to an OD_{600} of 0.6 before induction was performed with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 20 °C for 16–18 h. The BL21 (DE3) cells carrying the expression plasmid pETDuet-1-AldT were grown in 500 mL LB medium containing carbenicillin (100 µg/ml) to an OD_{600} of 0.5 at 37 °C before induction was performed with 0.4 mM IPTG at 37 °C for 4 h. All cultures were harvested by centrifugation at $5000\times g$ for 20 min, resuspended in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7 and lysed by three passages through a French pressure cell. Soluble fractions of the lysates were obtained by centrifugation at $15,000\times g$ for 20 min. His-tagged enzymes were purified by subjecting the soluble extracts to a 5 mL nickel-nitrilotriacetic acid (Ni-NTA) affinity column (GE Healthcare) equilibrated with buffer containing 50 mM sodium phosphate (NaP), 300 mM NaCl and 20 mM imidazole. The proteins were eluted isocratically with a final concentration of 400 mM imidazole. The buffer of the eluates was exchanged to 50 mM HEPES pH 7 by three rounds of centrifugation in Amicon Ultra centrifugal filters (Millipore) with a 30 kDa molecular weight cut-off.

In order to obtain TaManD without a N-terminal His-tag, the enzyme obtained after pProEX HTa-Ta0753 expression and first His-tag purification was incubated with a TEV Protease (Promega,

Madison, WI, USA). Digestion was performed according to the manufacturer's protocol with addition of 10% glycerol (v/v, final concentration) and incubation for 1.5 h at 30 °C followed by incubation at 4 °C overnight. Removal of the His-tag and the TEV Protease was achieved by a second Ni-NTA purification using a 1 mL ZetaSep Ni-NTA affinity column (emp Biotech, Berlin, Germany). The flow-through contained the pure protein and cleavage of the His-tag was verified by SDS-PAGE.

3.2. Protein Analysis

Purified proteins were separated by SDS-PAGE using 4–15% Tris-glycine polyacrylamide gels (Bio-Rad Laboratories, Hercules, USA) and visualised with Coomassie Brilliant Blue. For the identification of TaManD, the single protein band at 38 kDa was excised from the gel and analysed by LC ESI MS/MS at the Australian Proteome Analysis Facility (APAF, Macquarie University), as described elsewhere [34]. Protein identification was performed using the mascot software (2.4.1, Matrixscience, Boston, MA, USA). The native molecular mass of the recombinant protein was determined by size exclusion chromatography on an ÄKTA pure FPLC system using a Superdex Increase 200 10/300 GL column (GE Healthcare) equilibrated with 10 mM NaP buffer, 140 mM NaCl, pH 7. A calibration curve was obtained with molecular weight markers thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa) and ovalbumin (43 kDa).

3.3. Nuclear Magentic Resonance (NMR) Analysis

D-Mannono-1,4-lactone (>95%) was purchased from TCI Chemicals Co. (Tokyo, Japan). For NMR spectra of the sugar acid in different substrate forms, 1 M D-Mannono-1,4-lactone solutions were prepared in 1 M NaOH, 1 M HCl and in ultrapure water. All solutions were incubated for 1 h at room temperature, before each solution was diluted to 100 mM using ultrapure water. Then, 300 μL of each 100 mM solution (equivalent to 5 mg sugar acid) was freeze-dried and resuspended in 500 μL 0.1 M NaP buffer, pH 7. The pH of all three samples was evaluated prior to NMR analysis. The substrate incubated in HCl displayed pH 6, while the substrate in water or NaOH displayed pH 7. After transfer of each sample to 5 mm NMR tubes, 10% D_2O and 40 μM 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid (TMSP) chemical shift standard was added. The 1D decoupled ^{13}C NMR spectra of non-isotopically labelled samples were acquired on a 500 MHz Bruker Avance III HD NMR equipped with a BBFO probe at 50 °C (323 K), using power-gated proton decoupling with a 90 ° pulse with 512 scans and a 3 s relaxation delay between scans. Oxidation of D-[1,6-$^{13}C_2$]mannose was performed in reactions containing 5 mM of the substrate, 5 mM NAD^+, 20 mM NaP buffer at pH 7 and 13.5 μg purified AldT. Prior to the addition of purified enzyme and after 30 min of reaction at 50 °C, 1D power-gated proton decoupled ^{13}C NMR spectra were acquired using a 90 ° pulse with 4 scans and a 3 s relaxation delay between scans. Visualisation of all spectra was performed with iNMR 6.0 (http://www.inmr.net). After Fourier transformation and automatic phase correction, an exponential visual weighting factor of 1.5 and a smoothing factor of 10 were applied to all spectra.

3.4. Enzyme Activity

D-mannono-1,4-lactone was either used after preparation in 50 mM HEPES at pH 7 (indicated as "lactone in buffer") or after hydrolysis to its free acid form (indicated as "mannonate" or "free sugar acid"). Hydrolysis was performed according to Lamble et al. by preparing 1 M stock solutions of D-mannono-1,4-lactone in 1 M NaOH and incubation at room temperature for 1 h before dilutions were prepared in 50 mM HEPES pH 7 [35]. For temperature optimum, pH optimum and thermostability, activity was quantified using the thiobarbituric acid (TBA) assay, according to a modification by Buchanan et al. [36,37]. Unless stated otherwise, reactions (60 μL) contained 0.5–1.5 μg pure enzyme, 10 mM mannonate, 50 mM HEPES pH 7 (adjusted at 55 °C) and 1 mM $CoSO_4$. Reactions were stopped by incubation on ice or if needed by addition of 6 μL 12.5% trichloracetic acid (TCA). Next, 50 μL of the reaction mixture was oxidised using 125 μL of 25 mM periodic acid in 0.25 M H_2SO_4 at room temperature for 20 min. The oxidation was stopped by addition of 250 μL of 2% (w/v) sodium arsenite

in 0.5 M HCl. Finally, 1 mL of 0.3% TBA in water was added to the samples and the reaction mixture was boiled for 10 min. The formation of 2-keto-3-deoxy sugar acid was determined by reading A_{549nm} using 96-well microtiter plates in a Spectrostar Nano plate reader (BMG Labtech, Ortenberg, Germany) and quantified using the molar extinction coefficient of 67.8×10^3 $M^{-1} \times cm^{-1}$.

The effect of pH on enzyme activity was analysed in duplicate reactions with 1 μg enzyme and 120 mM universal buffer [38]. Reactions were incubated for 2 h at 55 °C before they were analysed with the TBA Assay. In reactions with D-mannono-1,4-lactone in buffer, pH decreased during the course of the experiment, which was accounted for in the analysis by measurement of the pH after the reaction. The optimal temperature for enzyme activity was determined in duplicate reactions containing 1 μg enzyme. Duplicate reactions were performed for 1 h at different temperatures before being analysed with the TBA Assay. For thermostability assays, 120 mM universal buffer was used and purified enzyme (1 μg) was incubated in duplicates at temperatures between 55 °C and 95 °C in the absence of substrate. Samples were removed at different times and the residual activity was measured in reactions with 10 mM mannonate for 45 min at 55 °C.

The semicarbazide assay was used according to Wichelecki et al. [12] to determine the effect of metal ions, chelators and reducing agents, and the acquisition of Michalis-Menten kinetics (K_m, V_{max}). Each reaction (60 μL) was incubated with 240 μL semicarbazide reagent at room temperature for 1 h. A_{250nm} was read in a UV transparent microtiter plate (Thermo Scientific, Waltham, MA, USA) using a Spectrostar Nano plate reader. Product formation was quantified using a standard curve of 2-keto-3-deoxygluconate (Sigma-Aldrich, St. Louis, MO, USA) prepared in the same assay buffer. K_m, V_{max} values were estimated using non-linear fitting in Prism 6 (6.0c, GraphPad software, San Diego, CA, USA). For the effect of metals, chelators and reducing agents, pure enzyme (0.6 μg) was incubated with 1 mM of each additive in 50 mM HEPES pH 7 for 1 h at room temperature. Activity was measured in duplicate reactions containing 10 mM D-mannono-1,4-lactone in 50 mM HEPES pH 7, incubated for 1 h at 55 °C and then analysed with the semicarbazide assay. Kinetic data for the determination of K_m and V_{max} were performed in duplicate reactions (60 μL), containing 1 μg purified TaManD, 50 mM HEPES pH 7, 1 mM $CoSO_4$ and different substrate concentrations (1–50 mM). Linear increase of reaction product was assured over 1 h reaction time. Reactions were performed for 45 min at 55 °C before product formation was determined using the semicarbazide assay and the KDG standard curve. Substrate specificity was analysed using 11 different sugar acids (Figure 8B). Reactions were performed in duplicate and contained 0.5 μg enzyme, 50 mM HEPES pH 7, 1 mM $CoSO_4$ and 10 mM of each sugar acid. Reactions were incubated for 16 h at 55 °C before they were analysed using the semicarbazide assay.

3.5. High Performance Liquid Chromatography (HPLC) Analysis

Reactions catalysed by TaManD and standards of mannonate and KDG were analysed using an Agilent 1290 HPLC system connected to a refractive index detector (RID) G1362A (Agilent Technologies, Santa Clara, CA, USA). Samples were analysed on an organic acid column (Agilent HiPlex H^+) with 10 mM H_2SO_4 as a mobile phase at a flow rate of 0.6 ml/min. The column was heated to 80 °C and the RID was set to 55 °C. Reactions were stopped by addition of 7 μL 12.5% TCA to a 60 μL reaction. After short centrifugation, 10 μL of the supernatant was used for HPLC analysis.

3.6. Statistical Analysis

Statistical analysis was performed for enzyme kinetics and effect of different metals and additives on enzyme activity, by two-tailed unpaired t-tests using Prism 6.

4. Conclusions

In this study, we present the first purification and characterisation of a functional archaeal mannonate dehydratase. The gene encoding for the mannonate dehydratase was found adjacent to a previously described aldohexose dehydrogenase (AldT) gene in the genome of *T. acidophilum* [15].

Previously, it has been shown that AldT from *T. acidophilum* selectively oxidises mannose, but the physiological function behind that oxidation was not investigated. Using NMR spectroscopy, we were able to show that mannonate and mannono-1,4-lactone are produced via oxidation of D-mannose by AldT. Kinetic assays confirmed that TaManD was able to convert both the sugar acid and its lactone to the central intermediate KDG at neutral pH, without further help of a lactonase. This resembles the second step of many oxidative pathways studied in archaea, including those for sugars like glucose, galactose, rhamnose, arabinose and xylose [39]. The amino acid sequence of TaManD and those of the putative mannonate dehydratases from *F. acidarmanus* and *F. acidiphilum* share high amino acid sequence identity. Although the gene annotations for those mannonate dehydratases indicate a role in the hexuronate metabolism (*uxuA*), they are all located adjacent to (putative) aldohexose dehydrogenases. It remains to be seen whether a non-phosphorylative pathway starting from mannose, similar to the non-phosphorylative Entner-Doudoroff pathway from glucose and galactose, exists in thermophilic archaea like *T. acidophilum*, *F. acidarmanus* and *F. acidiphilum*.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4344/9/3/234/s1, Figure S1: Purification of the Ta0753 gene product (TaManD) after expression in *E. coli*, Figure S2: TaManD kinetic data Table S1: Overview of genes, potentially encoding for mannonate dehydratases located in the genomic neighbourhood of AldT (Ta0754), Table S2: Oligonucleotides used in this study.

References

1. Robert-Baudouy, J.; Jimeno-Abendano, J.; Stoeber, F. D-Mannonate and D-altronate dehydratases of *Escherichia coli* K12. *Methods Enzymol.* **1982**, *90*, 288–294. [PubMed]

2. Shulami, S.; Gat, O.; Sonenshein, A.L.; Shoham, Y. The glucuronic acid utilization gene cluster from *Bacillus stearothermophilus* T-6. *J. Bacteriol.* **1999**, *181*, 3695–3704. [PubMed]

3. Mekjian, K.R.; Bryan, E.M.; Beall, B.W.; Moran, C.P. Regulation of hexuronate utilization in *Bacillus subtilis*. *J. Bacteriol.* **1999**, *181*, 426–433. [PubMed]

4. Hugouvieux-Cotte-Pattat, N.; Robert-Baudouy, J. Hexuronate catabolism in *Erwinia Chrysanthemi*. *J. Bacteriol.* **1987**, *169*, 1223–1231. [CrossRef] [PubMed]

5. Mandrand-Berthelot, M.-A.; Condemine, G.; Hugouvieux-Cotte-Pattat, N. Catabolism of hexuronides, hexuronates, aldonates, and aldarates. *EcoSal Plus* **2004**, *1*, 1–21. [CrossRef] [PubMed]

6. Reis, D.; Vian, B.; Roland, J.C. Cellulose-glucuronoxylans and plant cell wall structure. *Micron* **1994**, *25*, 171–187. [CrossRef]

7. Lawford, H.G.; Rousseau, J.D. Fermentation of biomass-derived glucuronic acid by *pet* expressing recombinants of *E. coli* B. *Appl. Biochem. Biotechnol.* **1997**, *63–65*, 221–241. [CrossRef]

8. Peekhaus, N.; Conway, T. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. *J. Bacteriol.* **1998**, *180*, 3495–3502. [PubMed]

9. Chang, D.-E.; Smalley, D.J.; Tucker, D.L.; Leatham, M.P.; Norris, W.E.; Stevenson, S.J.; Anderson, A.B.; Grissom, J.E.; Laux, D.C.; Cohen, P.S.; et al. Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7427–7432. [CrossRef] [PubMed]

10. Zhang, Q.; Gao, F.; Peng, H.; Cheng, H.; Liu, Y.; Tang, J.; Thompson, J.; Wei, G.; Zhang, J.; Du, Y.; et al. Crystal structures of *Streptococcus suis* mannonate dehydratase (ManD) and its complex with substrate: Genetic and biochemical evidence for a catalytic mechanism. *J. Bacteriol.* **2009**, *191*, 5832–5837. [CrossRef] [PubMed]

11. Qiu, X.; Tao, Y.; Zhu, Y.; Yuan, Y.; Zhang, Y.; Liu, H.; Gao, Y.; Teng, M.; Niu, L. Structural insights into decreased enzymatic activity induced by an insert sequence in mannonate dehydratase from Gram negative bacterium. *J. Struct. Biol.* **2012**, *180*, 327–334. [CrossRef] [PubMed]

12. Wichelecki, D.J.; Balthazor, B.M.; Chau, A.C.; Vetting, M.W.; Fedorov, A.A.; Fedorov, E.V.; Lukk, T.; Patskovsky, Y.V.; Stead, M.B.; Hillerich, B.S.; et al. Discovery of function in the enolase superfamily: D-mannonate and D-gluconate dehydratases in the D-mannonate dehydratase subgroup. *Biochemistry* **2014**, *53*, 2722–2731. [CrossRef] [PubMed]

13. Wichelecki, D.J.; Alyxa, J.; Vendiola, F.; Jones, A.M.; Al-obaidi, N.; Almo, S.C.; Gerlt, J.A. Investigating the physiological roles of low-efficiency D-mannonate and D-gluconate dehydratases in the enolase superfamily: pathways for the catabolism of L-gulonate and L-idonate. *Biochemistry* **2014**, *53*, 5692–5699. [CrossRef] [PubMed]

14. Rakus, J.F.; Fedorov, A.A.; Fedorov, E.V.; Glasner, M.E.; Vick, J.E.; Babbitt, P.C.; Almo, S.C.; Gerlt, J.A. Evolution of enzymatic activities in the enolase superfamily: D-mannonate dehydratase from *Novosphingobium aromaticivorans*. *Biochemistry* **2007**, *46*, 12896–12908. [CrossRef] [PubMed]

15. Nishiya, Y.; Tamura, N.; Tamura, T. Analysis of bacterial glucose dehydrogenase homologs from thermoacidophilic archaeon *Thermoplasma acidophilum*: finding and characterization of aldohexose dehydrogenase. *Biosci. Biotechnol. Biochem.* **2014**, *68*, 2451–2456. [CrossRef] [PubMed]

16. Yasutake, Y.; Nishiya, Y.; Tamura, N.; Tamura, T. Structural Insights into unique substrate selectivity of *Thermoplasma acidophilum* D-aldohexose dehydrogenase. *J. Mol. Biol.* **2007**, *367*, 1034–1046. [CrossRef] [PubMed]

17. Kim, S.; Lee, S.B. Catalytic promiscuity in dihydroxy-acid dehydratase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *J. Biochem.* **2006**, *139*, 591–596. [CrossRef] [PubMed]

18. Ashwell, G. [21] Enzymes of glucuronic and galacturonic acid metabolism in bacteria. In *Methods in enzymology*; Academic Press: Cambridge, MA, USA, 1962; Volume 5, pp. 190–208, ISBN 0076-6879.

19. Portalier, R.; Robert-Baudouy, J.; Stoeber, F. Regulation of *Escherichia coli* K-12 hexuronate system genes: *exu* regulon. *J. Bacteriol.* **1980**, *143*, 1095–1107. [PubMed]

20. Palleroni, N.J.; Duodoroff, M. Metabolism of Carbohydrate by *Pseudomonas Saccharophila* III. Oxidation of Arabinose. *J. Bacteriol.* **1957**, *74*, 180–185. [PubMed]

21. Watanabe, S.; Saimura, M.; Makino, K. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. *J. Biol. Chem.* **2008**, *283*, 20372–20382. [CrossRef] [PubMed]

22. Stephens, C.; Christen, B.; Fuchs, T.; Sundaram, V.; Watanabe, K.; Jenal, U. Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. *J. Bacteriol.* **2007**, *189*, 2181–2185. [CrossRef] [PubMed]

23. Wałaszek, Z.; Horton, D. Conformational studies on aldonolactones by NMR spectroscopy. Conformations of d-glucono-, D-mannono, D-gulono-, D-galactono-1,4-lactone in solution. *Carbohydr. Res.* **1982**, *105*, 131–143. [CrossRef]

24. Dreyer, J. The role of iron in the activation of mannonic and altronic acid hydratases, two Fe-requiring hydro-lyases. *Eur. J. Biochem* **1987**, *166*, 623–630. [CrossRef] [PubMed]

25. Robert-Baudouy, J.M.; Jimeno-Abendano, J.; Stoeber, F.R. Individualité des hydrolyases mannonique et altronique chez *E. coli* K-12. *Biochimie* **1975**, *57*, 1–8. [CrossRef]

26. Robert-Baudouy, J.M.; Stoeber, F.R. Purification et propriétés de la D-mannonate hydrolyase *d'Escherichia coli*. *Biochim. Biophys. Acta -Enzymology* **1973**, *309*, 473–485. [CrossRef]

27. Darland, G.; Brock, T.D.; Samsonoff, W.; Conti, S.F. A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science* **1970**, *170*, 1416–1418. [CrossRef] [PubMed]

28. Jung, J.H.; Lee, S.B. Identification and characterization of *Thermoplasma acidophilum* 2-keto-3-deoxy-D-gluconate kinase: A new class of sugar kinases. *Biotechnol. Bioprocess Eng.* **2005**, *10*, 535–539. [CrossRef]

29. Kim, S.M.; Paek, K.H.; Lee, S.B. Characterization of NADP$^+$-specific L-rhamnose dehydrogenase from the thermoacidophilic Archaeon *Thermoplasma acidophilum*. *Extremophiles* **2012**, *16*, 447–454. [CrossRef] [PubMed]

30. Reher, M.; Schönheit, P. Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. *FEBS Lett.* **2006**, *580*, 1198–1204. [PubMed]

31. Yew, W.S.; Fedorov, A.A.; Fedorov, E.V.; Rakus, J.F.; Pierce, R.W.; Almo, S.C.; Gerlt, J.A. Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. *Biochemistry* **2006**, *45*, 14582–14597. [CrossRef] [PubMed]

32. Carsten, J.M.; Schmidt, A.; Sieber, V. Characterization of recombinantly expressed dihydroxy-acid dehydratase from *Sulfobus solfataricus*—A key enzyme for the conversion of carbohydrates into chemicals. *J. Biotechnol.* **2015**, *211*, 31–41. [CrossRef] [PubMed]

33. Sambrook, J.; Russell, D.W.; Russell, D.W. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2001; Volume 49, ISBN 978-1-936113-42-2.

34. Atack, J.M.; Srikhanta, Y.N.; Fox, K.L.; Jurcisek, J.A.; Brockman, K.L.; Clark, T.A.; Boitano, M.; Power, P.M.; Jen, F.E.C.; McEwan, A.G.; et al. A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable Haemophilus influenzae. *Nat. Commun.* **2015**, *6*, 1–12. [CrossRef] [PubMed]

35. Lamble, H.J.; Milburn, C.C.; Taylor, G.L.; Hough, D.W.; Danson, M.J. Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *FEBS Lett.* **2004**, *576*, 133–136. [CrossRef] [PubMed]

36. Buchanan, C.L.; Connaris, H.; Danson, M.J.; Reeve, C.D.; Hough, D.W. An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. *Biochem. J.* **1999**, *343*, 563–570. [CrossRef] [PubMed]

37. Skoza, L.; Mohos, S. Stable thiobarbituric acid chromophore with dimethyl sulphoxide. Application to sialic acid assay in analytical de-O-acetylation. *Biochem. J.* **1976**, *159*, 457–462. [CrossRef] [PubMed]

38. Britton, H.T.S.; Robinson, R.A. CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. *J. Chem. Soc.* **1931**, 1456–1462. [CrossRef]

39. Bräsen, C.; Esser, D.; Rauch, B.; Siebers, B. Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 89–175. [CrossRef] [PubMed]

Comparative Analysis and Biochemical Characterization of Two Endo-β-1,3-Glucanases from the Thermophilic Bacterium *Fervidobacterium* sp.

Christin Burkhardt [1], Christian Schäfers [1], Jörg Claren [2], Georg Schirrmacher [2] and Garabed Antranikian [1,*]

[1] Institute of Technical Microbiology, Hamburg University of Technology (TUHH), Kasernenstr. 12, 21073 Hamburg, Germany; christin.burkhardt@tuhh.de (C.B.); christian.schaefers@tuhh.de (C.S.)

[2] Group Biotechnology, Clariant Produkte (Deutschland) GmbH, Semmelweisstr. 1, 82152 Planegg, Germany; joerg.claren@clariant.com (J.C.); georg.schirrmacher@eitfood.eu (G.S.)

[*] Correspondence: antranikian@tuhh.de; Tel.: +49-4042-878-3117

Abstract: Laminarinases exhibit potential in a wide range of industrial applications including the production of biofuels and pharmaceuticals. In this study, we present the genetic and biochemical characteristics of FLamA and FLamB, two laminarinases derived from a metagenomic sample from a hot spring in the Azores. Sequence comparison revealed that both genes had high similarities to genes from *Fervidobacterium nodosum* Rt17-B1. The two proteins showed sequence similarities of 62% to each other and belong to the glycoside hydrolase (GH) family 16. For biochemical characterization, both laminarinases were heterologously produced in *Escherichia coli* and purified to homogeneity. FLamA and FLamB exhibited similar properties and both showed highest activity towards laminarin at 90 °C and pH 6.5. The two enzymes were thermostable but differed in their half-life at 80 °C with 5 h and 1 h for FLamA and FLamB, respectively. In contrast to other laminarinases, both enzymes prefer β-1,3-glucans and mixed-linked glucans as substrates. However, FLamA and FLamB differ in their catalytic efficiency towards laminarin. Structure predictions were made and showed minor differences particularly in a kink adjacent to the active site cleft. The high specific activities and resistance to elevated temperatures and various additives make both enzymes suitable candidates for application in biomass conversion.

1. Introduction

β-1,3-Glucans are non-cellulosic carbohydrates that are widespread in nature. They can be found in the cell walls of fungi (pachyman), in reproductive structures of plants (callose) or as exopolysaccharide from bacteria (curdlan) [1]. β-1,3-Glycosidic linkages are also present in mixed-linked β-glucans from cereals (e.g., barley) or lichens [2]. Moreover, β-1,3-glucans are one of the most abundant carbohydrates in marine ecosystems [3]. In micro- and macroalgae, β-1,3-glucans are structurally diverse and serve as storage glucans. In brown algae, the β-1,3-glucan is named laminarin and represents up to 25% of the dry weight, depending on species, season, and growing conditions [4]. Due to the significantly higher production yields than terrestrial biomass, a high carbohydrate content, and the lack of hemicellulose and lignin, macroalgae biomass is a promising feedstock for new biorefinery concepts [5,6]. For industrial utilization of this feedstock, robust and efficient enzymes like laminarinases are required [7]. Furthermore, laminarinases could be applied for yeast extract production, as a biocontrol agent against fungal plant pathogens [8], and for partial hydrolysis of β-1,3-glucans for the production of antiviral and antitumor therapeutics [9].

For the complete enzymatic hydrolysis of β-1,3-glucans endo-acting β-1,3-glucanases (EC 3.2.1.39) and β-1,3(4)-glucanases (EC 3.2.1.6), both known as laminarinases, exo-β-1,3-glucosidases (EC 3.2.1.58) are required. These enzymes and their corresponding substrates have in common that they both are widely distributed among plants, fungi, and bacteria from many different habitats. On the basis of amino acid sequence similarities, all endo-acting laminarinases from plants can be assigned to the glycoside hydrolase family GH 17, whereas most of the bacterial laminarinases belong to GH 16 [10]. According to the Carbohydrate-Active Enzymes database (CAZy), around 40 β-1,3-glucanases from bacteria are already characterized (August 2019). Nevertheless, laminarinases with high stability at varying conditions and temperatures are desired for industrial applications [11]. To obtain highly stable and efficient enzymes for industrial application, thermophilic organisms represent an excellent resource. Moreover, these enzymes allow reactions at elevated process temperatures, which do not only reduce microbial contamination, but also increase the solubility and diffusion rates of the catalysts for complex polymeric substrates [12].

In this study, an environmental sample from an Azorean hot spring (São Miguel, Portugal) was used as a source for novel laminarinase-encoding genes. Recently, this environmental sample has been proven to be an excellent source for unique thermo-active enzymes [13]. By sequence-based screening, two putative genes were identified with sequence similarities to parts of the genomic sequence of the thermophilic bacterium *Fervidobacterium nodosum* Rt17-B1, which was completely sequenced in 2007 [14]. *Fervidobacterium nodosum* which was isolated from a hot spring in New Zealand, is able to ferment a wide range of carbohydrates [15] and is considered as a good source for novel carbohydrate degrading enzymes. So far, only one highly active and thermostable cellulase from this bacterium has been characterized [16].

Here, we describe the recombinant production and purification of the two thermoactive laminarinases FLamA and FLamB originating from *Fervidobacterium* sp. To estimate their potential relevance for industrial applications, both recombinant enzymes were characterized in detail.

2. Results

2.1. Sequence Analysis of FLamA and FLamB

By sequence-based screening, two new putative laminarinase encoding genes were identified. Metagenomic DNA from a hot spring of the Azores, which was known to contain genomic DNA of a *Fervidobacterium* strain, was used as template DNA for PCR. The primers for the amplification of the two genes encoding for FLamA and FLamB were based on two putative laminarinase genes of the complete genome sequence of *Fervidobacterium nodosum* Rt17-B1 (GenBank: CP000771). The amplified DNA fragment for *flamA* showed 99% sequence similarity to putative laminarinase genes from the genomes of *F. nodosum* Rt17-B1 and *F. pennivorans* DSM 9078 (GenBank: NC_017095). The corresponding amino acid sequence is annotated as a multispecies endo-β-1,3-glucanase found in various *Fervidobacterium* species (GenBank: WP_011994743). In comparison to that, the amplified DNA fragment for *flamB* showed 95% and 67% sequence similarity to GenBank sequences of two other putative laminarinase genes from *F. nodosum* Rt17-B1 and *F. pennivorans* DSM 9078, respectively. Due to the 99% similarity of the amino acid sequence to next hits in the database, *flamB* was annotated in GenBank with the accession number LT882624.

In the genome of *F. nodosum* Rt17-B1, the highly similar genes of *flamA* and *flamB* are embedded in two different operons. By comparing these operons, we identified putative open reading frames for proteins with the same predicted functions (Figure 1). Among both gene clusters, the related proteins revealed high sequence similarities of 62–89% on amino acid level (100% query coverage). *FlamA* and *flamB* showed 62% sequence similarity to each other. Adjacent to the putative laminarinase encoding genes, we identified genes for putative proteins involved in ABC transporter system and genes for putative alanine racemases and GH 3 proteins. Moreover, a truncated gene encoding a

Transposase_20 and a repeat region nearby one of the predicted operons indicated that the two related operons originate from a gene duplication event of a DNA cassette around 20 kb in size.
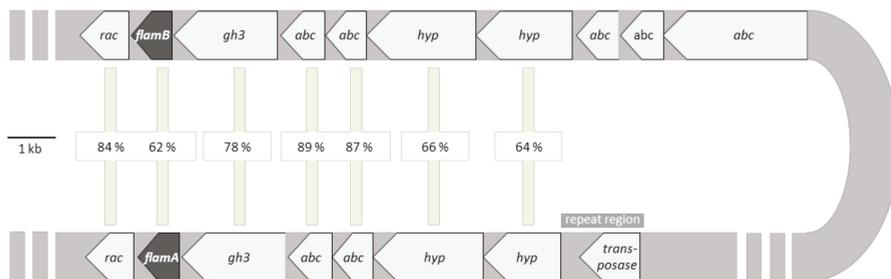

Figure 1. Predicted structural organization of the two operons containing *flamA* and *flamB* based on the highly similar genes in the genome of *Fervidobacterium nodosum*. Sequence identities of similar genes are shown in the center. Predicted genes: *rac*—putative gene for alanine racemase; *flamA* and *flamB*—genes for β-1,3-glucanases; *gh3*—putative genes for glycoside hydrolase of family GH 3; *abc*—putative genes for ABC transporter proteins; *transposase*—truncated gene of a Transposase_20 with a repeat region; *hyp*—hypothetical genes.

Structure predictions for FLamA and FLamB revealed a sandwich-like β-jelly roll fold, which is characteristic for laminarinases of class GH 16 (Figure S1). The structural prediction of both proteins was based on the catalytic residue of *Thermotoga maritima*, with 58% and 60% sequence identity to *flamA* and *flamB*, respectively. The comparison of both predictions demonstrated the high structural similarities of FLamA and FLamB with only small differences in a loop adjacent to the cleft of the active site (highlighted in color).

The sequences of both proteins showed similarities to those of laminarinases from other thermophilic bacteria. Indeed, FLamA and FLamB were similar to catalytic domains of experimentally verified β-1,3-glucanases from *T. maritima*, *T. petrophila,* and *T. neapolitana* with a 54–55% and 58–59% identity, respectively. Moreover, FLamA and FLamB were 55% and 58% similar to the catalytic domain of a β-1,3-glucanase from the hyperthermophilic Archaeon *Pyrococcus furiosus*. The phylogenetic relationship among both enzymes and other biochemically characterized GH 16 family members of bacteria is shown in Figure 2. Concerning the predicted tree based on the homologous region of the catalytic domain, FLamA and FLamB form a solid clade with the β-1,3-glucanases of the Archaeon *P. furiosus* and members of the same eubacterial order Thermotogales, which includes some of the most extremely thermophilic species currently known.

2.2. Recombinant Production of FLamA and FLamB

Both genes were expressed in *E. coli* C43(DE3) at 37 °C over 4 h of induction. The obtained proteins harboring a C-terminal 6xHis affinity tag were purified via affinity chromatography and size exclusion chromatography. Both FLamA and FLamB were purified to homogeneity with a factor of 178 and 593 and a final yield of 30% and 18%, respectively (Table S1). The SDS-PAGE revealed a molecular weight of approximately 30 kDa for both proteins, which was slightly smaller than the predicted molecular size of 34.9 kDa and 34.1 kDa for FLamA and FLamB, respectively (Figure 3). Domain prediction revealed that FLamA and FLamB consisted, in each case, of one single GH 16 domain without any further known structural elements.
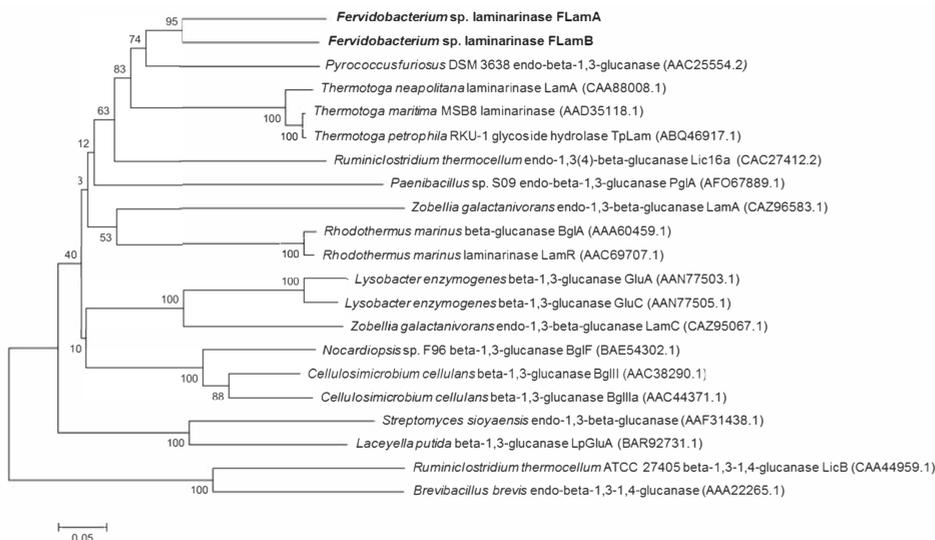
Figure 2. Phylogenetic tree of biochemically characterized GH 16 enzymes. For construction of the tree, sequences of the catalytic domains were used. GenBank accession numbers are indicated in brackets. Bootstrap values are designated on each branch of the tree.
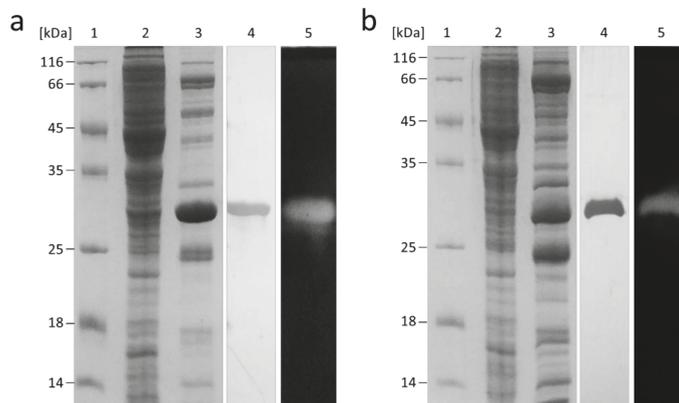
Figure 3. SDS-PAGE analysis of the His-tagged β-1,3-glucanases FLamA (**a**) and FLamB (**b**). *Line 1*, molecular weight marker; *line 2*, crude extract; *line 3*, eluate after the Ni-NTA affinity chromatography; *line 4*, purified enzymes after size exclusion chromatography; *line 5*, zymogram for activity staining of the purified enzyme.

2.3. Substrate Specificity of FLamA and FLamB

The substrate specificity of FLamA and FLamB was tested towards a number of complex carbohydrates in a range from 40 to 100 °C (Figure 4, Table 1). Significant differences between FLamA and FLamB were detected in the specific activities towards β-1,3-glucans. Highest activities of both enzymes were measured at 90 °C towards laminarin, whereby FLamB showed a higher specific activity than FLamA, with 876 U/mg and 609 U/mg, respectively. Compared to that, highest activities towards amorphous curdlan were revealed at 70 °C with 78% and 94% relative activity. For undissolved curdlan, the highest activity was determined at 80 °C, with even lower relative activities of 44% and 50% for FLamA and FLamB, respectively.
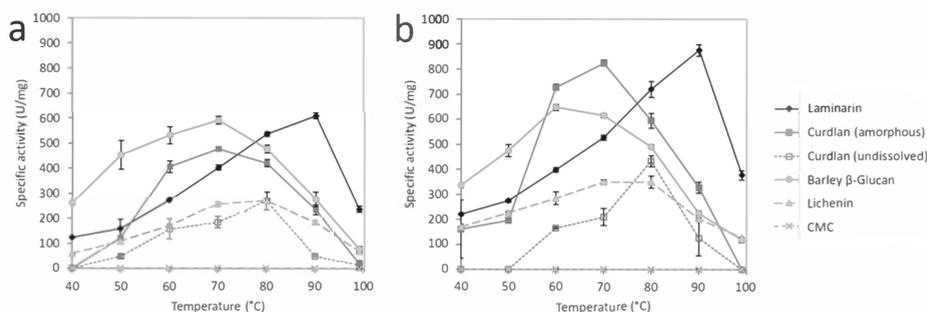
Figure 4. The temperature profiles of recombinant FLamA (**a**) and FLamB (**b**) depending on substrates laminarin, amorphous curdlan, unsolved curdlan, barley β-glucan, lichenin, and carboxymethyl cellulose (CMC).

Table 1. Substrate specificity of FLamA and FLamB at the optimal temperatures for each substrate.

Substrate	FLamA			FLamB		
	T (°C)	Specific Activity (U/mg)	Relative Activity (%)	T (°C)	Specific Activity (U/mg)	Relative Activity (%)
Laminarin	90	609 ± 12	100	90	876 ± 23	100
Curdlan *	70	478 ± 07	78	70	825 ± 13	94
Curdlan	80	270 ± 37	44	80	434 ± 22	50
Barley β-glucan	70	592 ± 16	97	60	648 ± 12	74
Lichenin	80	271 ± 5	45	70	350 ± 8	40
CMC [+]	-	0	0	-	0	0

* Amorphous curdlan, diluted in NaOH and neutralized; [+] between 40–99 °C, no activity was detectable.

In comparison to β-1,3-glucans, the specific activities towards mixed-linked glucans did not differ substantially between FLamA and FLamB. The specific activities of FLamA and FLamB towards barley β-glucan from barley were in the same range with 592 U/mg and 648 U/mg, respectively. This was determined at optimal temperatures of 70 °C (FLamA) and 60 °C (FLamB). For both enzymes, 44% and 52% residual activities towards barley β-glucan were detected at 40 °C. Moreover, both enzymes showed, at 40 °C, higher activities towards mixed-linked β-glucan (262 U/mg and 336 U/mg) than towards laminarin (124 U/mg and 220 U/mg). The specific activities towards lichenin at optimal conditions (80 °C and 70 °C) were with 271 U/mg and 350 U/mg for FLamA and FLamB, lower than towards the other tested substrates. Hydrolysis of the β-1,4-glucan CMC was not observed.

2.4. Degradation Pattern and Enzyme Kinetics

The hydrolysis products of laminarin and barley β-glucan were investigated by HPLC analysis. The FLamA and FLamB produced similar degradation patterns. Hydrolysis of laminarin (Figure 5a) mainly results in laminaribiose, glucose-mannitol units (DP2), and glucose (DP1) and to a lesser extent laminaritriose (DP3) and higher oligosaccharides. The degradation pattern of barley β-glucan was different (Figure 5b). Laminaritriose was the major product and to a lesser extent DP1 and DP2 were detected. The produced oligosaccharides indicated that FLamA and FLamB are β-1,3-glucanases with endo-acting mode. Moreover, the shifted product variation with barley β-glucan suggested that both enzymes hydrolyzed the β-1,3-glycosidic linkages in mixed-linked glucans and to a lesser extent the β-1,4-building blocks.
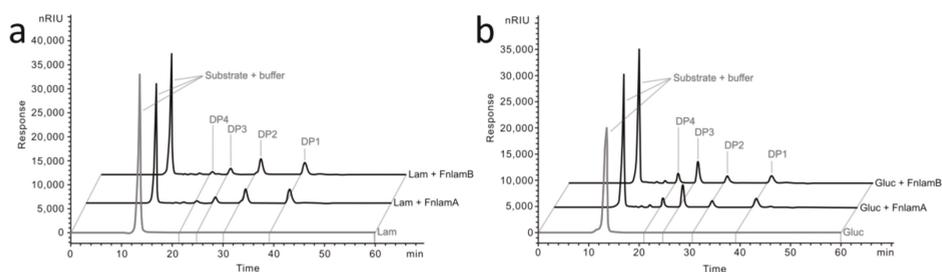
Figure 5. HPLC analysis of products from polysaccharide degradation. Hydrolysis products from laminarin (**a**) and barley β-glucan (**b**) after 18 h of incubation at 70 °C. DP: degree of polymerization. DP1: glucose. Lam: laminarin. Gluc: barley β-glucan.

The kinetic parameters were determined in the presence of laminarin. At optimal conditions (90 °C, pH 6.5) the K_m value of FLamA was 2.01 mg ml^{-1} and with that higher than that of FLamB (1.64 mg ml^{-1}). Moreover, the catalytic efficiency values (k_{cat}/K_m) were 494.3 ml s^{-1} mg^{-1} and 806.3 ml s^{-1} mg^{-1} for FLamA and FLamB, respectively. The results indicated that FLamB has a 61% higher catalytic efficiency on laminarin than FLamA, which is in agreement with the determined substrate specificities.

2.5. Effects of pH and Temperature

The effect of pH was analyzed in a range of pH 2 to 12. Measurements revealed similar pH spectra for FLamA and FLamB (Figure 6). Both enzymes showed activity in a range between pH 5 and 9. Optimal activity was detected at pH 6.5. Concerning the stability of FLamA, there was a decrease by approximately 8% in activity over the entire pH range after 24 h (Figure S3). In the same pH range, FLamB showed at least 82% residual activity.
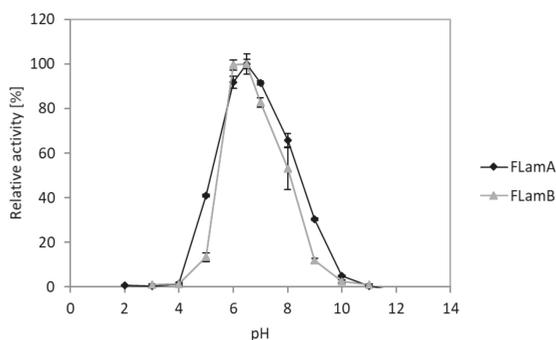
Figure 6. Influence of pH on the activity of recombinant FLamA and FLamB towards laminarin.

To examine the temperature stability of the two thermoactive enzymes, the residual activities after heat incubation at 70, 80, and 90 °C were measured (Figure 7). FLamA and FLamB exhibited more than 80% and 40% residual activity over 24 h at 70 °C, respectively. At 80 °C, the enzymes had a half-life time of 5 h (FLamA) and 1 h (FLamB). By incubating the enzymes at 90 °C after 30 min, no residual activity was detected. FLamA was, in the tested temperature range, more stable than FLamB.
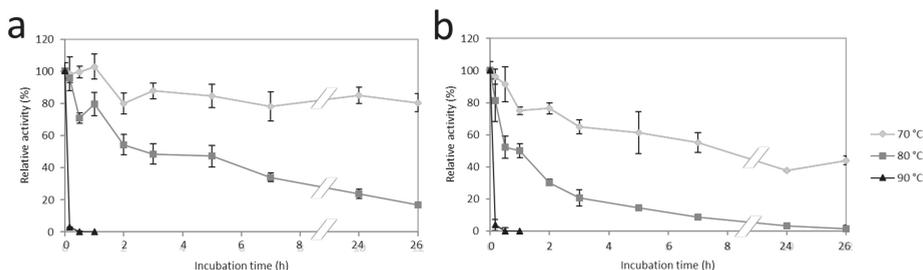
Figure 7. Influence of temperature on the stability of FLamA (**a**) and FLamB (**b**).

2.6. Effects of Metal Ions and Chemical Additives

The effect of various additives on the activities of FLamA and FLamB was investigated. For both enzymes, the examined additives had similar influences. Among the tested metal ions (5 mM), Ag^+, Al^{3+}, Cr^{3+}, Cu^{2+}, Zn^{2+}, and Fe^{2+} completely inhibited FLamA and FLamB (Table 2a). Both Co^{2+} and Ni^{2+} significantly reduced the activity of the enzymes. The same effect was observed by the addition of 5 mM SDS, cetyltrimethylammonium bromide (CTAB), and Pefabloc (Table 2b). In contrast, the reducing agents dithiothreitol (DTT) and β-mercaptoethanol had positive effects on enzyme activities.

Table 2. Influence of metal ions (**a**) and different reagents (**b**) on laminarinase activity of FLamA and FLamB.

| a	Relative Residual Activity (%) *		b	Relative Residual Activity (%) *	
Metal ion	**FLamA**	**FLamB**	**Reagent**	**FLamA**	**FLamB**
AgNO₃	0.65 ± 0.00	3.29 ± 0.58	**CHAPS**	78.95 ± 3.64	80.11 ± 2.36
AlCl₃	4.98 ± 2.31	4.73 ± 0.27	**SDS**	29.10 ± 1.08	36.70 ± 6.61
CaCl₂	89.85 ± 8.46	87.79 ± 1.60	**Triton X-100**	89.42 ± 4.03	89.77 ± 0.61
CoCl₂	9.31 ± 0.33	16.64 ± 1.30	**Tween 20**	91.83 ± 1.67	91.18 ± 2.20
CrCl₃	2.50 ± 0.04	1.32 ± 4.80	**Tween 80**	95.41 ± 3.85	92.52 ± 1.60
CuCl₂	0.66 ± 0.40	0.84 ± 4.18	**Guanidine-HCl**	91.14 ± 7.34	84.70 ± 0.68
FeCl₂	1.05 ± 0.17	3.29 ± 1.14	**Urea**	100.14 ± 1.41	93.30 ± 1.06
KCl	93.78 ± 1.00	91.08 ± 1.39	**DTT**	119.04 ± 1.16	124.61 ± 1.88
MgCl₂	92.48 ± 2.42	95.15 ± 1.57	**β-Mercaptoethanol**	125.03 ± 2.88	124.12 ± 3.42
NaCl	92.41 ± 1.53	86.42 ± 1.17	**EDTA**	109.32 ± 3.01	95.28 ± 2.20
NiCl₂	4.67 ± 0.16	7.06 ± 0.10	**Na-Iodoacetate**	105.34 ± 1.19	92.45 ± 2.57
RbCl	90.79 ± 1.76	97.47 ± 1.01	**Pefabloc**	37.79 ± 5.67	76.94 ± 12.14
SrCl₂	93.15 ± 1.84	96.20 ± 2.63	**CTAB**	24.85 ± 1.69	30.68 ± 2.60
ZnCl₂	1.85 ± 0.20	3.67 ± 0.29	**Na-Azide**	102.37 ± 2.14	103.81 ± 6.97

* Laminarinase activity without additives was defined as 100%.

3. Discussion

The two genes *flamA* and *flamB* revealed high sequence similarities of 62% (100% query coverage) with each other. Two highly similar genes within the genome of *Fervidobacterium nodosum* were located in two different operons. The pattern of the operons and adjacent genes as well as the presence of a truncated transposase gene suggest that these structures were generated by a duplication of an around 20 kb genome section. Moreover, it has been shown that *flamA* and *flamB* encode for active laminarinases possessing nearly the same biochemical characteristics, indicating that both genes form a multigene family.

There are other bacterial species which express more than one gene encoding for laminarinases. Nevertheless, these genes contain various modular architectures and the proteins possess different domains and functions like in the marine bacterium *Zobellia galactanivorans* [17]. In *Lysobacter enzymogenes,* two laminarinases are present which have high sequence identities in part of the catalytic

domain (86%) but differentiate in an additional carbohydrate binding module (CBM) in one of the proteins [18]. Similar relationships were also observed for related laminarinases from *Streptomyces sioyaensis* [19]. So far, however, only members of the family Fervibacteriaceae are known to harbor two similar genes, such as *flamA* and *flamB*, which encode for laminarinases with similar sequence and function.

It has been shown that gene duplication probably generates no benefit by enhanced gene dosage of laminarinases or adjacent genes, such as those for the ABC transporter system. The increase of gene dosages may cause higher energy consumption leading to significant reduced fitness of the organism [20,21]. Therefore, it is more likely that gene duplication leads to redundancy and, thus, facilitates mutations in one of the copies and possibly results in new benefiting gene functions [22]. A similar operon composition occurs in the closely related *F. pennivorans*, where two genes exist with 100% and 70% sequence identity to *flamA* and *flamB*, respectively. This demonstrates that *flamA* is highly conserved in both species, whereas differences in *flamB* were probably generated by point mutations. This may lead to beneficial properties, like obviously higher catalytic activities of FLamB towards β-1,3-glucans. If there is no selective advantage, duplicated genes normally become inactivated by mutations and these pseudogenes will disappear from the genome [23].

Phylogenetic analysis of the catalytic residues of FLamA, FLamB, and other characterized laminarinases revealed a solid clade with close a relationship between enzymes from the bacterial order Thermotogales, including FLamA and FLamB, and the Archaeon *Pyrococcus furiosus* (Figure 2). Thermotogales are assumed to have interchanged large numbers of genes with Archaea and Firmicutes by horizontal gene transfer (HGT) [14]. An extraordinarily high number of insertion sequence elements for example in the genome of *F. nodosum* demonstrate the great influence of foreign genes. The high sequence similarity of FLamA, FLamB, and the laminarinases of the bacterium *Thermotoga* to one from the Archaeon *P. furiosus* suggests that an ancestral gene was exchanged by HGT.

Regarding the temperature optima of these enzymes, FLamA and FLamB are both thermoactive enzymes showing highest activity at 90 °C. This optimum is similar to that of the other examined laminarinases from the phylum Thermotogae [24,25] which are to our knowledge the ones with the highest temperature optima among reported β-1,3-glucanases from eubacteria and plants. However, LamA from the Archaeon *P. furiosus* showed among all characterized laminarinases activity at the highest temperatures of 100–105 °C (Table 3) [26]. Both FLamA and FLamB showed high storage stabilities at 4 °C in a broad pH range of 3 to 11 which enables an easy handling of both laminarinases. Their stability at 70 to 90 °C is also comparable to those of the other laminarinases operating at high temperatures (Table 3). All in all, FLamA and FLamB have significantly high temperature stabilities, which are a decisive criterion for industrial application.

The investigation of the substrate specificities of FLamA and FLamB revealed different temperature optima for the tested substrates (Figure 4). Schwarz et al. [27] observed related characteristics of a laminarinase of *Clostridium thermocellum*. The optimal temperature towards curdlan is lower in consideration to the fact that curdlan forms at temperatures above 80 °C irreversible high-set gels [28], which might be less accessible for enzymes. Nevertheless, the reasons for the different temperature optima can only be speculated.

Table 3. Comparison of biochemically characterized laminarinases from thermophilic bacteria and Archaea.

Organism and Enzyme	T_{opt} (°C)	Thermal Stability			pH_{opt}	Activity (U/mg) [+]				Reference
		T (°C)	t (h)	A * (%)		Lam	Curd	Glu	Lich	
Caldicellulosiruptor sp. F32, Lam16A-GH	75	65	42	72	6.5	172	ND	2961	ND	[29]
Fervidobacterium sp., FLamA	90	80	5	50	6.5	609	270	592	271	Present study
Fervidobacterium sp., FLamB	90	80	1	50	6.5	876	434	648	350	Present study
Laceyella putida, LpGluA	80	75	0.5	45	4.2	48%	100%	ND	ND	[30]
Nocardiopsis sp. F96, BglF	70	ND	ND	ND	9.0	100%	159%	ND	815%	[31]
Pyrococcus furiosus, LamA	100/105	80	80	100	6–6.5	922	ND	99	95	[26]
Rhodothermus marinus 21, BglA	85	80	16	100	7.0	542	ND	1568	1445	[32]
Rhodothermus marinus ITI278, LamR	88	90	0.45	50	5.5	656	ND	2199	3111	[33]
Ruminiclostridium thermocellum, CelC	65	70	10	30	6.5	86	ND	504	245	[27]
Ruminiclostridium thermocellum, Lic16A	70	70	0.17	50	6.0	340	29	268	2404	[34]
Thermotoga maritima, TmβG	80	ND	ND	ND	5.0	Efficient on β-1,3-glucans				[35]
Thermotoga neapolitana, LamA	95	95	0.5	82	6.3	3100	ND	ND	90	[25]
Thermotoga petrophila, TpLam	91	80	16	60	6.2	48	ND	41	21	[24]

* Residual activity detected after incubation at the given temperature and time. [+] If available, specific activity (U/mg); otherwise relative activities (%) towards laminarin (Lam), Curdlan (Curd), barley β-glucan (Gluc) or lichenin (Lich).

So far, published thermostable β-1,3-linkage hydrolyzing enzymes generally possessed strong preferences either for β-1,3-glucans or mixed-linked glucans (Table 3). Both FLamA and FLamB take an intermediate position between those groups by degrading approximately both types of substrates with activity in the same range. Additionally, FLamA and FLamB exhibit higher substrate affinities towards mixed-linked glucans in comparison to similar enzymes from *Thermotoga neapolitana* [25] and *P. furiosus* [26]. Ilari et al. [36] investigated a deletion mutant and were able to show that the loss of a five amino acid kink at the entrance of the catalytic cleft leads to higher activity towards mixed-linked glucans. This effect was also observed in other laminarinases missing these residues [34,37]. Even in FLamA and FLamB, these amino acid residues are missing (Figure S2). Thus, differences in the kink possibly explain the altered substrate preferences between those enzymes.

The described substrate spectra and hydrolysis products suggest that FLamA and FLamB are able to cleave β-1,3 glycosidic bonds in an endo-acting mode. Nevertheless, the hydrolysis of β-1,4 glycosidic bonds in mixed-linked glucans is also possible. Similar product patterns of laminarin and barley β-glucan were observed with a laminarinase form *Caldicellulosiruptor* sp., when the authors proved that the enzyme was able to degrade β-1,4 glyosidic bonds adjacent to 3-*O*-substituted glucopyranose units [32].

For industrial application, detection of inhibitory effects and comparison with other enzymes the influences of different metal ions and additives were tested. For FLamA and FLamB similar effects were observed. The inhibitory effects of metal ions observed in this study are well known for many glucosidases, probably due to the redox effects on the amino acids [38]. The SDS had a negative influence on the activity of the enzymes by disturbing hydrophobic interactions and generating protein denaturation. Moreover, the detergent CTAB blocked the catalytic cleft of a related laminarinase by hydrophobic interactions [38], which may lead to reduced activities of FLamA and FLamB as well. Pefabloc, a serine protease inhibitor, negatively influenced the enzyme activity probably by binding irreversibly to serine residues nearby the catalytic cleft. In contrary, the reducing properties of DTT

and β-mercaptoethanol might positively influence a cysteine residue adjacent to the nucleophile of the active site and therefore increase the laminarinase activity.

Although FLamA and FLamB share many biochemical properties, the two enzymes exhibit differences in substrate specificities, particularly towards β-1,3-glucans. In comparison to the specific activities of FLamA, FLamB exhibits approximately 40% and more than 70% higher activity towards laminarin and amorphous curdlan, respectively (Table 1). These differences were reflected in the kinetic parameters as well. In the analysis of Labourel et al. [39], significant differences in substrate specificity were caused by an additional loop in protein structure which leads to higher affinities towards mixed-linked glucans. Nevertheless, according to structural predictions, those considerable differences were not observed between FLamA and FLamB (Figure S1). Only minor structural modifications particularly in a loop adjacent to the catalytic cleft possibly result in an upwardly more opened cleft of FLamB. Based on the observations of Labourel et al. [39] and Jeng et al. [40] concerning the enzyme-substrate complexes, the enlarged opening might improve the access for β-1,3-glucans that possess helical conformation, whereas the affinity of FLamB towards mixed-linked glucans with linear conformation is not affected. These results and further investigations will help to improve activities of β-1,3-glucananases towards β-1,3-glucans.

4. Materials and Methods

4.1. Cloning of the Endo-β-1,3-Glucanase Encoding Genes flamA and flamB

Metagenomic DNA was extracted from environmental samples taken from different locations at the hot spring Caldeirão at Furnas Valley (Azores, Portugal) followed by the production and sequencing of a 454 shotgun library as described previously [41]. For the two genes *flamA* and *flamB*, encoding for endo-β-1,3-glucanases from *Fervidobacterium*, no signal peptides were predicted by SignalP [42]. For cloning into the StarGate system (IBA Lifesciences) the two genes were amplified by PCR using the metagenomic DNA as a template and the following primers (primer extending sequences are indicated in boldface):

- *flamA*-for: **AGCGGCTCTTCAATG**AAAGTTAAATATTTCTCAAATATT
- *flamA*-rev: **AGCGGCTCTTCTCCC**CTCATTTTCAAGCTTGTATAC
- *flamB*-for: **AGCGGCTCTTCAATG**AGAGAAAAGTTGCTGT
- *flamB*-rev: **AGCGGCTCTTCTCCC**CTCTTCATCTAATGTATACAC

The PCR products were cloned into the destination vector pASG-IBA33 according to the producer instructions resulting in recombinant fusion genes with C-terminal sequences encoding for hexahistidine tags. After transformation of *Escherichia coli* TOP10 and selection on LB medium containing 100 μg/mL ampicillin and 50 μL/mL X-gal, plasmids of recombinant clones were isolated. The inserts were sequenced for verification. Subsequently, the vectors were used to transform *E. coli* C43(DE3) for protein production.

4.2. Sequence Comparison and Phylogenetic Analysis

From the GenBank database, amino acid sequences of characterized β-1,3-glucanases and β-1,3(4)-glucanases of the family GH 16 were selected. Multiple sequence alignment was performed using ClustalX. Homologous sequence regions were selected and applied in a second multiple sequence alignment. Using MEGA6, a phylogenetic tree was calculated by the neighbor joining method. Bootstrap analysis with resampling of the dataset was performed (*n* = 100) to test the reliability of the tree. Structure prediction of both proteins were done by SWISS-MODEL using the crystal structure of the laminarinase from *Thermotoga maritima* (PDB ID: 3azx) as a template. Structures of FLamA and FLamB were visualized and compared in the UCFS Chimera program by applying the Needleman–Wunsch algorithm and the scoring matrix Blosum62.

4.3. Heterologous Expression of the flamA and flamB Genes and Purification of the Endo-β-1,3-Glucanases

Escherichia coli C43(DE3) harboring the plasmids pASG-IBA33::*flamA* or pASG-IBA33::*flamB* were grown in LB media (100 µg/mL ampicillin) at 37 °C and 160 rpm. Gene expression was induced at OD_{600} 0.6 by adding anhydrotetracycline to a final concentration of 200 ng/µL. After four hours of induction cells were harvested by centrifugation at $9000 \times g$ at 4 °C for 20 min. The resulting cell pellet was stored at −20 °C.

For purification, 0.2 g cells were resuspended per 1 mL lysis buffer (50 mM NaH2PO4, 300 mM NaCl, 10 mM imidazole, pH 8) and disrupted by three passages through a French pressure cell with constant pressure of 1,000 psi (French Pressure Cell Press, SLM-Aminco). Cell debris was removed by centrifugation ($20,000 \times g$, 4 °C, 30 min) and supernatant was loaded onto a 1 ml Ni-NTA Superflow column (Qiagen). Proteins were eluted by an increasing imidazole gradient according to the manufacturer's instructions. Eluted fractions were pooled, washed three times with buffer G (50 mM Na-phosphate buffer, pH 7.2, 150 mM NaCl) by ultrafiltration in an Amicon filter unit (Amicon Ultra-15, 1000 MWCO, Merck Millipore). For final purification via size exclusion chromatography, protein solutions were loaded onto a Superose 12 column (GE Healthcare) previously equilibrated with buffer G. Protein fractions containing the purified β-1,3-glucanases were pooled and stored at 4 °C.

Protein samples were analyzed on a 12.5% SDS-PAGE (12.5%) [43]. Additionally, β-1,3-glucanase activity was determined by zymogram technique. For this, proteins were applied to a denaturing SDS-PAGE and were subsequently renatured by incubation for 1 h in 1% (*v/v*) Triton X-100 and three successive washings for 5 min in 50 mM sodium phosphate buffer (pH 6.5). Then, the gel slices were incubated for 20 min at 80 °C on an agarose gel containing 0.1% (*w/v*) curdlan. For visualization agarose slides were stained for 1 h with 1% (*w/v*) Congo Red and destained in 1 M NaCl. To increase the contrast, the slides were finally overlaid with 0.1 M acetic acid.

Protein concentrations were determined according to Bradford [44], with bovine serum albumin as the standard.

4.4. β-Glucanase Activity Assay

The standard assay was carried out at 90 °C for 7 min in 500 µL reaction mixture using 0.25% (*w/v*) laminarin from *Laminaria digitata* (Merck) as substrate in 20 mM sodium phosphate buffer (pH 6.5) and 50 µL enzyme sample. In advance, it was ensured that product formation per min was constant in the time interval and with a linear correlation to preclude instability effects of the enzymes. Additionally, blank experiments without enzymes were performed by default for all measurement series. The hydrolytic activities of the purified enzymes FLamA and FLamB were detected by measuring the reducing sugars with 3,5-dinitrosalicylic acid (DNS) according to Miller [45] with glucose as the standard. In brief, after enzyme reaction 500 µl reaction mixture were mixed with 500 µL DNS reagent (1% (*w/v*) DNS, 30% (*w/v*) potassium sodium tartrate, 0.4 M NaOH) and were incubated for 5 min at 100 °C. Samples were subsequently cooled on ice to room temperature and absorption was measured at 546 nm. All measurements were done in triplicates. One unit of enzyme activity was defined as the amount of enzyme required to release 1 µmol of reducing sugars per minute.

The influence of temperature was examined by performing the standard assay at temperatures from 20 to 100 °C. To investigate the temperature stability of FLamA and FLamB, the enzymes were preincubated with a concentration of 0.1 mg/mL in 20 mM sodium phosphate buffer (pH 6.5) at 70, 80, and 90 °C. Samples were taken in time intervals up to 26 h and residual activities were measured by using the standard assay.

To investigate the influence of the pH on enzyme activity, a standard assay was performed using Britton–Robinson buffer (50 mM) in a range of pH 2–11 in the reaction mixture [46]. The pH stabilities of both enzymes were tested by preincubation of the enzymes with a concentration of 0.01 mg/mL in 50 mM Britton–Robinson buffer pH 3–11 for 24 h at 4 °C. Residual activity was determined with the standard assay by dilution the incubation mixtures in Britton–Robinson buffer at pH 6. Enzyme activity previous to incubation was defined as 100%.

Additionally, the influences of metal ions on enzyme activity were analyzed by using a standard assay, but with 20 mM maleate buffer (pH 6.5) and the addition of 5 mM $AgNO_3$, $AlCl_3$, $CaCl_2$, $CoCl_2$, $CrCl_3$, $CuCl_2$, $FeCl_2$, KCl, $MgCl_2$, NaCl, $NiCl_2$, RbCl, $SrCl_2$ or $ZnCl_2$. Furthermore, the influences of 3-((3-cholamidopropyl)dimethylammonio)-1-propanesulfonate (CHAPS), SDS, Triton X-100, Tween 20, Tween 80, guanidine hydrochloride, urea, dithiothreitol (DTT), β-mercaptoethanol, EDTA, iodoacetic acid, Pefabloc, cetyltrimethylammonium bromide (CTAB) and sodium azide were examined by the standard assay procedure. All additives were tested in a concentration of 5 mM under standard conditions.

To measure the specific activities of FLamA and FLamB, substrates were used in a final concentration of 0.25% (*w/v*). The CMC and lichenin were obtained from Merck and β-glucan (barley) and curdlan (*Alcaligenes faecalis*) from Megazyme. In case of curdlan an undissolved and a dissolved (amorphous) form was tested. To achieve an amorphous type of curdlan, 0.2 g were first solubilized in 6 mL alkaline solution (0.6 M NaOH) and subsequently neutralized with HCl to a concentration of 0.5% (*w/v*) and pH 6.5 in 20 mM sodium phosphate buffer.

Kinetic parameters were determined by performing the standard assay with twelve different substrate concentrations varying from 0 to 25 mM. The Michaelis constant K_m and the maximum reaction rate at maximum substrate concentration v_{max} were calculated by non-linear regression applying the Michaelis–Menten equation.

4.5. Determination of the Hydrolysis Products

For determination of the hydrolysis products of laminarin and barley β-glucan, 0.25% (*w/v*) substrate were incubated with FLamA or FLamB in standard reaction mixtures at 70 °C for 18 h. After the inactivation of the enzymes at 100 °C for 10 min, samples were centrifuged (20.000 × *g*, 10 min) and the supernatant was filtered using a 0.22 μm membrane filter unit. Hydrolysis products were analyzed by high-performance liquid chromatography (HPLC) under the following conditions: Hi-Plex Na column (Agilent Technologies), 80 °C, 0.2 mL/min flow rate, water as mobile phase, RI detector (Agilent Technologies). Laminaritetraose, laminaritriose, laminaribiose (all from Megazymes) and glucose (Merck) were used as standards.

5. Conclusions

Laminarinases are enzymes which could be applied in diverse fields, from biomass conversion, over yeast extract production, agents against fungal plant pathogens to the production of antiviral and antitumor therapeutics from β-1,3-glucans. The biochemical characterization of the two laminarinases FLamA and FLamB derived from a *Fervidobacterium* species revealed high specific activities and resistance to elevated temperatures and various additives which make both enzymes suitable candidates for application under harsh conditions. Moreover, the comparative analysis of both enzymes showed differences in their thermal stability and catalytic efficiency towards β-1,3-glucans, like laminarin and curdlan. In conclusion, these results will contribute to our knowledge of sequence-function correlations and will potentially help to improve activity and stability of laminarinases and other related glucanases.

References

1. Stone, B.A. Chemistry, biochemistry, and biology of (1-3)-β-glucans and related polysaccharides. In *Chemistry of β-Glucans*, 1st ed.; Bacic, A., Fincher, G.B., Stone, B.A., Eds.; Elsevier Academic Press: Amsterdam, The Netherlands, 2009; pp. 5–46.
2. Ebringerová, A.; Hromádková, Z.; Heinze, T. Hemicellulose. In *Polysaccharides, I. Structure, Characterisation and Use*; Heinze, T.T., Ed.; Springer-Verlag GmbH: Berlin/Heidelberg, Germany, 2005; pp. 1–67.
3. Painter, T.J. Algal polysaccharides. In *The Polysaccharides*; Aspinall, G.O., Ed.; Academic Press (Molecular biology): New York, NY, USA, 1983; pp. 195–285. Volume 2.
4. Schiener, P.; Black, K.D.; Stanley, M.S.; Green, D.H. The seasonal variation in the chemical composition of the kelp species Laminaria digitata, Laminaria hyperborea, Saccharina latissima Alaria esculenta. *J. Appl. Phycol.* **2015**, *27*, 363–373. [CrossRef]
5. Wei, N.; Quarterman, J.; Jin, Y.-S. Marine macroalgae: an untapped resource for producing fuels and chemicals. *Trends Biotechnol.* **2013**, *31*, 70–77. [CrossRef] [PubMed]
6. Jung, K.A.; Lim, S.-R.; Kim, Y.; Park, J.M. Potentials of macroalgae as feedstocks for biorefinery. *Bioresour. Technol.* **2013**, *135*, 182–190. [CrossRef] [PubMed]
7. Hehemann, J.-H.; Boraston, A.B.; Czjzek, M. A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Curr. Opin. Struct. Biol.* **2014**, *28*, 77–86. [CrossRef] [PubMed]
8. Shi, P.; Yao, G.; Yang, P.; Li, N.; Luo, H.; Bai, Y.; Wang, Y.; Yao, B. Cloning, characterization, and antifungal activity of an endo-1,3-β-D-glucanase from *Streptomyces* sp. S27. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1483–1490. [CrossRef] [PubMed]
9. Vetvicka, V. Glucan-immunostimulant, adjuvant, potential drug. *World J. Clin. Oncol.* **2011**, *2*, 115–119. [CrossRef]
10. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–D238. [CrossRef]
11. Sarmiento, F.; Peralta, R.; Blamey, J.M. Cold and Hot Extremozymes: Industrial Relevance and Current Trends. *Front. Bioeng. Biotechnol.* **2015**, *3*, 348. [CrossRef]
12. Krahe, M.; Markl, H.; Antranikian, G. Fermentation of extremophilic microorganisms. *FEMS Microbiol. Rev.* **1996**, *18*, 271–285. [CrossRef]
13. Schroeder, C.; Elleuche, S.; Blank, S.; Antranikian, G. Characterization of a heat-active archaeal beta-glucosidase from a hydrothermal spring metagenome. *Enzyme Microb. Technol.* **2014**, *57*, 48–54. [CrossRef]
14. Zhaxybayeva, O.; Swithers, K.S.; Lapierre, P.; Fournier, G.P.; Bickhart, D.M.; DeBoy, R.T.; Nelson, K.E.; Nesbø, C.L.; Doolittle, W.F.; Gogarten, J.P.; et al. On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5865–5870. [CrossRef] [PubMed]
15. Patel, B.K.C.; Morgan, H.W.; Daniel, R.M. Fervidobacterium nodosum gen. nov. and spec. nov., a new chemoorganotrophic, caldoactive, anaerobic bacterium. *Arch. Microbiol.* **1985**, *141*, 63–69. [CrossRef]
16. Wang, Y.; Wang, X.; Tang, R.; Yu, S.; Zheng, B.; Feng, Y. A novel thermostable cellulase from Fervidobacterium nodosum. *J. Mol. Catal. B: Enzym.* **2010**, *66*, 294–301. [CrossRef]
17. Labourel, A.; Jam, M.; Legentil, L.; Sylla, B.; Hehemann, J.H.; Ferrières, V.; Czjzek, M.; Michel, G. Structural and biochemical characterization of the laminarinase ZgLamCGH16 from Zobellia galactanivorans suggests preferred recognition of branched laminarin. *Acta Crystallogr. D* **2015**, *71 Pt 2*, 173–184. [CrossRef]
18. Palumbo, J.D.; Sullivan, R.F.; Kobayashi, D.Y. Molecular Characterization and Expression in Escherichia coli of Three β-1,3-Glucanase Genes from Lysobacter enzymogenes Strain N4-7. *J. Bacteriol.* **2003**, *185*, 4362–4370. [CrossRef] [PubMed]
19. Hong, T.-Y.; Huang, J.-W.; Meng, M.; Cheng, C.-W. Isolation and biochemical characterization of an endo-1,3-β-glucanase from Streptomyces sioyaensis containing a C-terminal family 6 carbohydrate-binding module that binds to 1,3-β-glucan. *Microbiology* **2002**, *148 Pt 4*, 1151–1159. [CrossRef]
20. Qian, W.; Zhang, J. Gene dosage and gene duplicability. *Genetics* **2008**, *179*, 2319–2324. [CrossRef]
21. Wagner, A. Energy Constraints on the Evolution of Gene Expression. *Mol. Boil. Evol.* **2005**, *22*, 1365–1374. [CrossRef]

22. Francino, M.P. An adaptive radiation model for the origin of new gene functions. *Nat. Genet.* **2005**, *37*, 573–578. [CrossRef]

23. Bratlie, M.S.; Johansen, J.; Sherman, B.T.; Huang, D.W.; Lempicki, R.A.; Drabløs, F. Gene duplications in prokaryotes can be associated with environmental adaptation. *BMC Genom.* **2010**, *11*, 588. [CrossRef]

24. Cota, J.; Alvarez, T.M.; Citadini, A.P.; Santos, C.R.; Neto, M.D.O.; Oliveira, R.R.; Pastore, G.M.; Ruller, R.; Prade, R.A.; Murakami, M.T.; et al. Mode of operation and low-resolution structure of a multi-domain and hyperthermophilic endo-β-1,3-glucanase from Thermotoga petrophila. *Biochem. Biophys. Res. Commun.* **2011**, *406*, 590–594. [CrossRef] [PubMed]

25. Zverlov, V.V.; Volkov, I.Y.; Velikodvorskaya, T.V.; Schwarz, W.H. Thermotoga neapolitana bglB gene, upstream of lamA, encodes a highly thermostable beta-glucosidase that is a laminaribiase. *Microbiology* **1997**, *143 Pt 11*, 3537–3542. [CrossRef]

26. Gueguen, Y.; Voorhorst, W.G.B.; van der Oost, J.; de Vos, W.M. Molecular and biochemical characterization of an endo-β -1,3-glucanase of the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **1997**, *272*, 31258–31264. [CrossRef] [PubMed]

27. Schwarz, W.H.; Schimming, S.; Staudenbauer, W.L. Isolation of a Clostridium thermocellum gene encoding a thermostable β-1,3-glucanase (laminarinase). *Biotechnol. Lett.* **1988**, *10*, 225–230. [CrossRef]

28. Kasai, N.; Harada, T. Ultrastructure of Curdlan. In *Fiber Diffraction Methods*; French, A.D., Ed.; American Chemical Society (ACS symposium series): Washington, DC, USA, 1980; pp. 363–383. Volume 141.

29. Meng, D.D.; Wang, B.; Ma, X.Q.; Ji, S.Q.; Lu, M.; Li, F.L. Characterization of a thermostable endo-1,3(4)-beta-glucanase from Caldicellulosiruptor sp. strain F32 and its application for yeast lysis. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 4923–4934. [CrossRef] [PubMed]

30. Kobayashi, T.; Uchimura, K.; Kubota, T.; Nunoura, T.; Deguchi, S. Biochemical and genetic characterization of beta-1,3 glucanase from a deep subseafloor Laceyella putida. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 203–214. [CrossRef]

31. Masuda, S.; Endo, K.; Koizumi, N.; Hayami, T.; Fukazawa, T.; Yatsunami, R.; Fukui, T.; Nakamura, S. Molecular identification of a novel beta-1,3-glucanase from alkaliphilic Nocardiopsis sp. strain F96. *Extremophiles* **2006**, *10*, 251–255. [CrossRef] [PubMed]

32. Spilliaert, R.; Hreggvidsson, G.O.; Kristjansson, J.K.; Eggertsson, G.; Palsdottir, A. Cloning and Sequencing of a Rhodothermus marinus Gene, bglA, Coding for a Thermostable beta-Glucanase and its Expression in Escherichia coli. *JBIC J. Boil. Inorg. Chem.* **1994**, *224*, 923–930. [CrossRef]

33. Krah, M.; Misselwitz, R.; Politz, O.; Thomsen, K.K.; Welfle, H.; Borriss, R. The laminarinase from thermophilic eubacterium *Rhodothermus marinus*. Conformation, stability, and identification of active site carboxylic residues by site-directed mutagenesis. *Eur. J. Biochem.* **1998**, *257*, 101–111. [CrossRef]

34. Fuchs, K.-P.; Zverlov, V.V.; Velikodvorskaya, G.A.; Lottspeich, F.; Schwarz, W.H. Lic16A of *Clostridium thermocellum*, a non-cellulosomal, highly complex endo-beta-1,3-glucanase bound to the outer cell surface. *Microbiology* **2003**, *149 Pt 4*, 1021–1031. [CrossRef]

35. Woo, C.-B.; Kang, H.-N.; Lee, S.-B. Molecular cloning and anti-fungal effect of endo-β-1,3-glucanase from Thermotoga maritima. *Food Sci. Biotechnol.* **2014**, *23*, 1243–1246. [CrossRef]

36. Ilari, A.; Fiorillo, A.; Angelaccio, S.; Florio, R.; Chiaraluce, R.; Van Der Oost, J.; Consalvi, V. Crystal structure of a family 16 endoglucanase from the hyperthermophile Pyrococcus furiosus- structural basis of substrate recognition. *FEBS J.* **2009**, *276*, 1048–1058. [CrossRef] [PubMed]

37. Gaiser, O.J.; Piotukh, K.; Ponnuswamy, M.N.; Planas, A.; Borriss, R.; Heinemann, U. Structural basis for the substrate specificity of a *Bacillus* 1,3-1,4-beta-glucanase. *J. Mol. Biol.* **2006**, *357*, 1211–1225. [CrossRef] [PubMed]

38. Pereira, J.D.C.; Giese, E.C.; Moretti, M.M.D.S.; Gomes, A.C.D.S.; Perrone, M.B.O.M.; Da Silva, R.; Martins, D.A.B. Effect of Metal Ions, Chemical Agents and Organic Compounds on Lignocellulolytic Enzymes Activities. In *Enzyme Inhibitors and Activators*; IntechOpen: London, UK, 2017; pp. 139–164.

39. Labourel, A.; Jam, M.; Jeudy, A.; Hehemann, J.-H.; Czjzek, M.; Michel, G. The β-glucanase ZgLamA from *Zobellia galactanivorans* evolved a bent active site adapted for efficient degradation of algal laminarin. *J. Biol. Chem.* **2014**, *289*, 2027–2042. [CrossRef] [PubMed]

40. Jeng, W.-Y.; Wang, N.-C.; Lin, C.-T.; Shyur, L.-F.; Wang, A.H.-J. Crystal structures of the laminarinase catalytic domain from *Thermotoga maritima* MSB8 in complex with inhibitors: essential residues for β-1,3- and β-1,4-glucan selection. *J. Biol. Chem.* **2011**, *286*, 45030–45040. [CrossRef] [PubMed]

41. Sahm, K.; John, P.; Nacke, H.; Wemheuer, B.; Grote, R.; Daniel, R.; Antranikian, G. High abundance of heterotrophic prokaryotes in hydrothermal springs of the Azores as revealed by a network of 16S rRNA gene-based methods. *Extremophiles* **2013**, *17*, 649–662. [CrossRef] [PubMed]

42. Petersen, T.N.; Brunak, S.; Von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **2011**, *8*, 785–786. [CrossRef] [PubMed]

43. Laemmli, U.K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227*, 680–685. [CrossRef] [PubMed]

44. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]

45. Miller, G.L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.* **1959**, *31*, 426–428. [CrossRef]

46. Britton, H.T.S.; Robinson, R.A. CXCVIII.—Universal buffer solutions and the dissociation constant of veronal. *J. Chem. Soc.* **1931**, 1456–1462. [CrossRef]

Article

Expression, Characterisation and Homology Modelling of a Novel Hormone-Sensitive Lipase (HSL)-Like Esterase from *Glaciozyma antarctica*

Hiryahafira Mohamad Tahir [1,2], Raja Noor Zaliha Raja Abd Rahman [1,3], Adam Thean Chor Leow [1,4] and Mohd Shukuri Mohamad Ali [1,2,*]

[1] Enzyme and Microbial Technology Research Center, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia; firatahir3@gmail.com (H.M.T.); rnzaliha@upm.edu.my (R.N.Z.R.A.R.); adamleow@upm.edu.my (A.T.C.L.)
[2] Department of Biochemistry, Faculty of Biotechnology and Biomolecular Science, Universiti Putra Malaysia, Serdang Selangor 43400, Malaysia
[3] Department of Microbiology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, Serdang Selangor 43400, Malaysia
[4] Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, Serdang Selangor 43400, Malaysia
* Correspondence: mshukuri@upm.edu.my; Tel.: +603-9769-6721

Abstract: Microorganisms, especially those that survive in extremely cold places such as Antarctica, have gained research attention since they produce a unique feature of the protein, such as being able to withstand at extreme temperature, salinity, and pressure, that make them desired for biotechnological application. Here, we report the first hormone-sensitive lipase (HSL)-like esterase from a *Glaciozyma* species, a psychrophilic yeast designated as GlaEst12-like esterase. In this study, the putative lipolytic enzyme was cloned, expressed in *E. coli*, purified, and characterised for its biochemical properties. Protein sequences analysis showed that GlaEst12 shared about 30% sequence identity with chain A of the bacterial hormone-sensitive lipase of E40. It belongs to the H group since it has the conserved motifs of Histidine-Glycine-Glycine-Glycine (HGGG)and Glycine-Aspartate-Serine-Alanine-Glycine (GDSAG) at the amino acid sequences. The recombinant GlaEst12 was successfully purified via one-step Ni-Sepharose affinity chromatography. Interestingly, GlaEst12 showed unusual properties with other enzymes from psychrophilic origin since it showed an optimal temperature ranged between 50–60 °C and was stable at alkaline pH conditions. Unlike other HSL-like esterase, this esterase showed higher activity towards medium-chain ester substrates rather than shorter chain ester. The 3D structure of GlaEst12, predicted by homology modelling using Robetta software, showed a secondary structure composed of mainly α/β hydrolase fold, with the catalytic residues being found at Ser^{232}, Glu^{341}, and His^{371}.

Keywords: psychrophilic yeast; hormone-sensitive lipase; *Glaciozyma antarctica*; Antarctica and homology modelling

1. Introduction

The lipolytic enzyme consists of esterases (EC 3.1.1.1) and lipases (EC 3.1.1.3) that catalyse both the cleavage and formation of ester bonds [1]. Although they have similar secondary structures, i.e., α/β hydrolase fold, esterase prefers to hydrolyse fatty-acids esters with acyl chain with less than 10 carbon atoms, whereas lipase is able to hydrolyse long-chain fatty acids with more than 10 carbon atoms [2]. Based on the sequence similarity, these protein have been classified into four groups, namely, C (cholinesterases and fungal lipase), L (lipoprotein and bacterial lipase), H (mammalian

hormone-sensitive lipase and hormone-sensitive lipase (HSL)-like family), and X (α/β hydrolase and does not belong to any of the other groups) [3].

The H group consists of two members that are HSL and HSL-like, in which both of them having conserved motifs, such as GDSAG or Glycine-Threonine-Serine-Alanine-Glycine (GTSAG) and HGGG motifs. Hormone-sensitive lipase is an enzyme that is mostly found in mammalian tissue and stimulated by several hormones, such as catecholamines, ACTH, and glucagon, to hydrolyse the triglyceride into free fatty acids and glycerol, which makes it play a pivotal role in providing the major source of energy for most tissues [4,5]. Another member of the H group is HSL-like, which is mostly originated from microbial sources that have similar protein sequences with HSL, especially at the C-terminal catalytic domain [6]. Although the mechanism of catalysis and the function of N-terminal domain HSL-like in microorganisms is still scarcely explored, the discovery of this new enzyme provides biotechnological application, such as in biosensors to detect foodborne bacteria and organophosphate pesticides [7,8]. Besides that, HSL-like enzymes also have potential to be used in the pharmaceutical, biodiesel, and detergents industry [6,9,10].

Several HSL-like enzymes have been reported from microbial sources, such as RmEstB from the thermophilic fungus *Rhizomucor miehei* [11], PMGL2 from a permafrost bacterium *Psychrobacter cryohalolentis* [12], and E25 HSL esterase from a surface sediment sample E505 collected from the South China Sea [13]. Even though there are many reported HSL-like esterase from psychrophilic microorganisms on heterologous expression and biochemical characterisation, there are few reports on HSL-like esterase specifically from psychrophilic yeast. Discovery of the new HSL-like esterase from psychrophilic yeast not only provides an opportunity in biotechnology application but also gives crucial information on novel sequence, characterisation, and the structure–function relationship.

Glaciozyma antarctica strain PI12 is a member of the phylum Basidiomycota that was previously known as the *Leucosporidium antarcticum* [14]. This psychrophilic yeast was isolated from sea ice near the Casey Research Station in Antarctica and had optimum temperature growth at 12 °C but can grow at up to 18 °C [15]. A few reported proteins have been successfully expressed from *G. antarctica* such as proteases, antifreeze protein, α-amylase, and chitinase [16–19]. In this work, we report the heterologous expression, purification, biochemical characterisation, and structural prediction of the first HSL-like esterase from the *Glaciozyma Antarctica* species, and we also believe this enzyme is the first HSL-like esterase from psychrophilic basidiomycete yeast.

2. Results and Discussion

2.1. Sequence Analysis of GlaEst12

The amino acid sequence of *Glaciozyma antarctica* hormone-sensitive lipase (GlaEst12) esterase was searched for similarity against the protein data bank at the National Centre of Biotechnology (NCBI) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using BLASTP. The search results showed that GlaEst12 had low sequence similarity (about 30% identity) with chain A of the crystal structure of esterase 40 from the bacterial HSL family and apparently no homology to the HSL-like esterase from psychrophilic yeast or bacteria. No similarity of the GlaEst12 sequence with psychrophilic microorganism may have two possible explanations. Firstly, there is less discovery of this type of enzyme from a cold environment, so the possibility of this GlaEst12 sequence being similar to other mesophiles and thermophiles is quite high. Hormone-sensitive lipase from *Psycrobacter* sp. that has been isolated from Antarctic seawater also showed similar reports in that the enzyme is closely related to the HSL-like esterase from the mesophilic enzyme [20]. Another explanation is because GlaEst12 showed mesophilic or thermophilic characteristics rather than psychrophilic features. The limited sequence of the GlaEst12 to the known HSL-like esterase sequences indicated less conserved residues, which provide novelty properties. The nucleotide sequence of GlaEst12 revealed an open reading frame (ORF) of 1200 nucleotide, which encoded 399 amino acids with a predicted molecular weight of 44.5 kDa. This esterase lacks signal peptide and has a theoretical isoelectric point (pI) value of 7.72.

The multiple sequence alignment was performed using ENDscript with the other seven proteins that have higher percentage of sequence similarity with GlaEst12 esterase, in which they are chain A of esterase from the bacterial HSL family (PDB ID: 4XVC A); chain A of mutant S202w/203f of the Esterase E40 (PDB ID: 5GMS A); chain C of mutant M3+s202w/i203f of Esterase E40 (PDB ID: 5GMR C); chain A of esterase/lipase from uncultured bacterium (PDB ID: 3V9A A); chain A of Hormone-sensitive lipase-like Este5 (PDB ID: 3FAK A), chain A of hyper-thermophilic carboxylesterase from archaeon *Archaeoglobus fulgidus* (PDB ID: 1JJI A), and chain A of MGS-MT1, an alpha/beta hydrolase enzyme from a Lake Matapan deep-sea metagenome library. Surprisingly, the GlaEst12 sequence is closely related to the mesophilic and thermophilic esterase. None of them are from either a psychrophilic or Antarctica environment. This finding might give new insight into the highly similar protein sequences that are not usually from the same environment.

Multiple sequence alignment showed that GlaEst12 belongs to the H group of the lipolytic group, which consists of a type of protein that has sequence similarity with the HSL subfamily. Most of the members of the H group have two highly conserved motifs, which are also present in the GlaEst12 sequence, such as His-Gly-Gly-Gly at upstream of the catalytic triad and the residue of serine at GDSAG motif [3]. Figure 1 shows GlaEst12 adhered to the characteristic of the H group, which is indicated by the red residue for HGGG and GDSAG motifs. The alignment with other proteins showed the possibility of the catalytic residue of GlaEst12 at position Ser^{232}, Glu^{341}, and His^{371}. The hormone-sensitive lipase-like family (HSL-like) can be widely found in microorganisms, animals, and plants. Most of the microbial HSL-like family consists of two subfamilies, GDSAG and GTSAG [21]. Since the serine residue was located at the pentapeptide motif, which is in the middle between aspartate acid and alanine, we proposed that GlaEst12 is a new member of the GDSAG subfamily of the HSL family.

Furthermore, the phylogenetic tree was constructed based on the amino acid sequence that aligned with closely related proteins and with the other members of HSL-like esterase from prokaryotic and eukaryotic microorganisms (PDB ID: 4QO5; 4Q30; 4WY8; 4WY5; Accession number: WP_012330536.1, ADH59412, QBH67630.1, KX580963.1). The results showed that GlaEst12 is grouped under the GDSAG motif subfamily (Figure 2) together with other proteins containing a GDSAG conserved sequence. Interestingly, GlaEst12 was assigned at a different sub-branch with other GDSAG subfamily members, indicating the differences of this sequence with the other esterases. The contrast may be due to the presence of extra α-helix at the N-terminal region, which was absent in all other esterases. Apart from that, this esterase is mostly related to eukaryotic proteins, such as RmEsTA (PDB 4WY5) and RmEsTB (PDB 4WY8), since they come from fungi species.
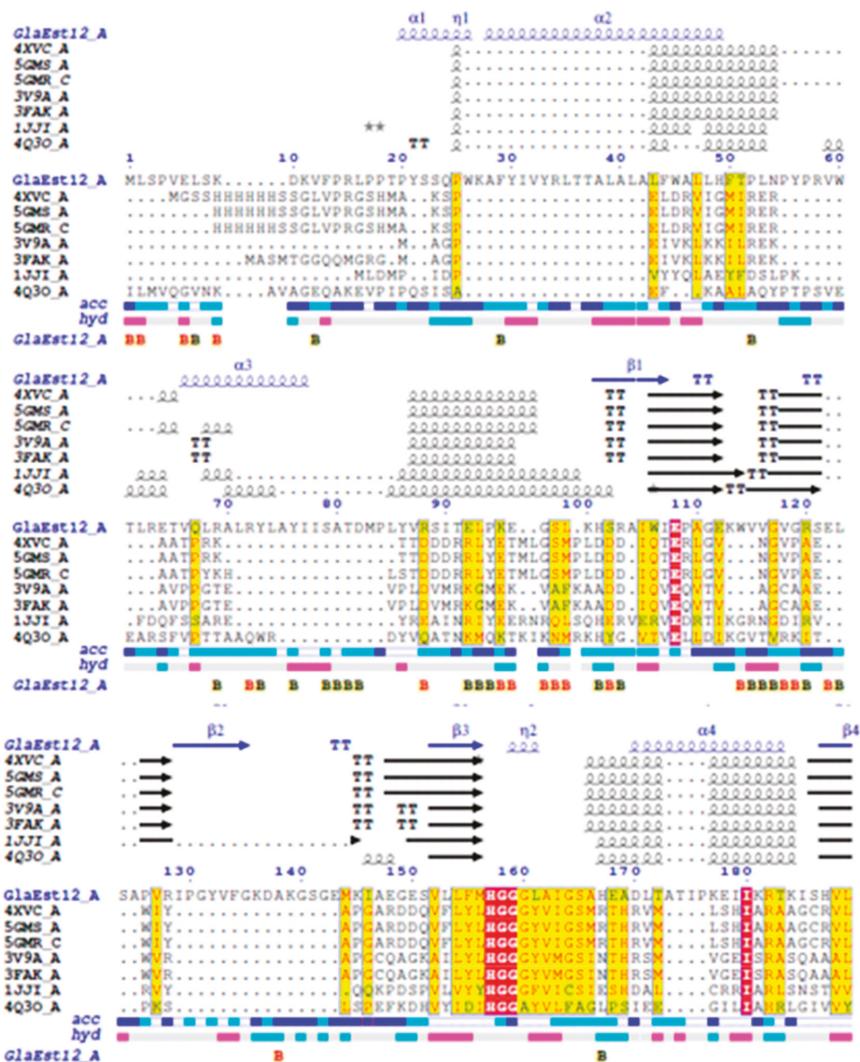
Figure 1. *Cont.*
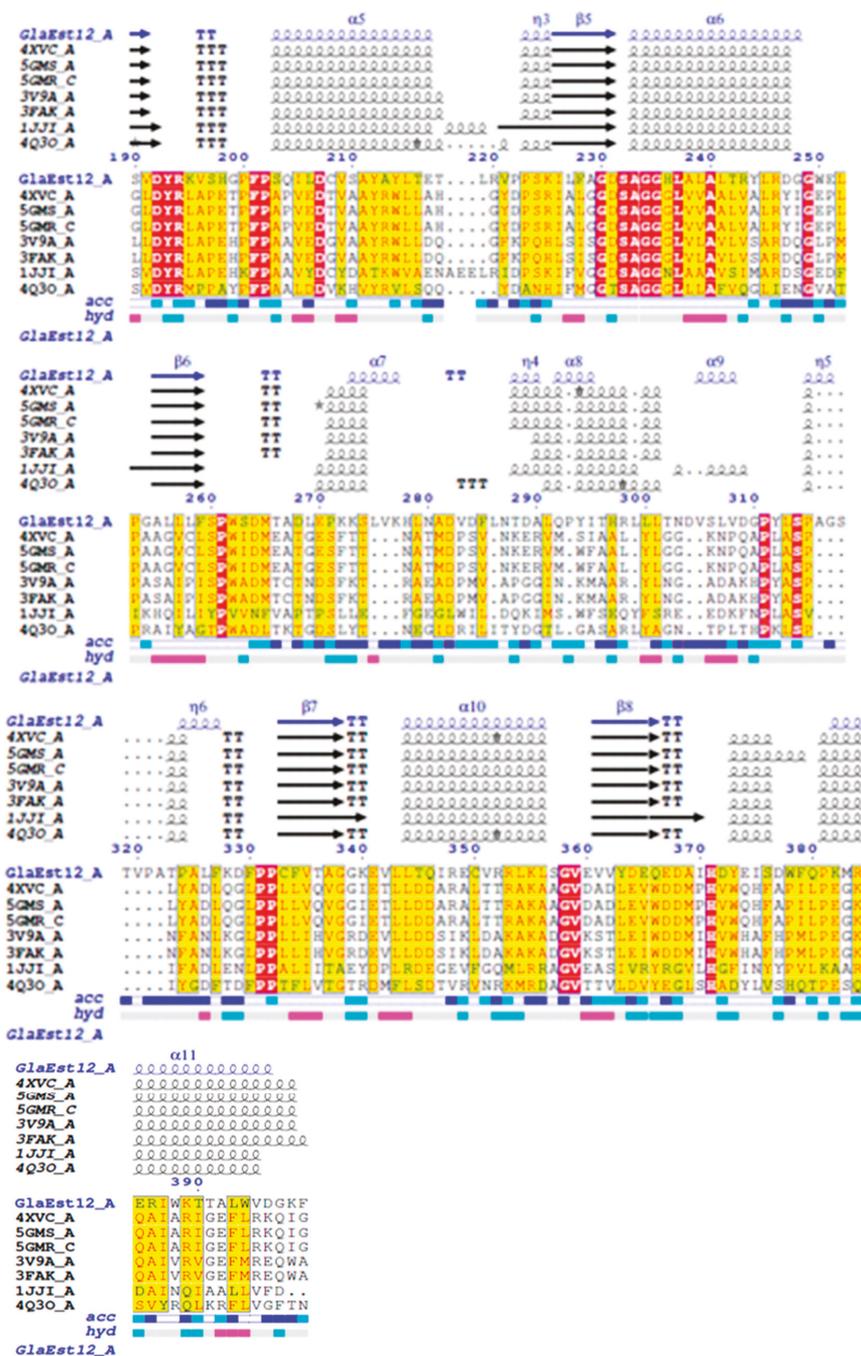
Figure 1. Multiple alignments of amino acid and secondary structure protein sequences from *Glaciozyma Antarctica* of hormone-sensitive lipase (GlaEst12) esterase with other related proteins. Squiggles indicate helices, arrows indicate β-strands, TT letters indicate a turn, ᵑ letters indicate random coil and the catalytic triad are indicated by an arrow symbol. The identical and highly conserved residues are indicated by red and yellow colour, respectively.

Figure 2. Phylogenetic tree of representative esterase sequences from microbial hormone-sensitive lipase (HSL) family generated using MEGA 7.0. The amino acid sequences were retrieved from the National Centre of Biotechnology (NCBI) and Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) database. The neighbour-joining method was built with a Jones-Taylor-Thornton matrix-based model to estimate the phylogenetic tree. The black box indicates GlaEst12.

2.2. Expression and Purification of Recombinant GlaEst12

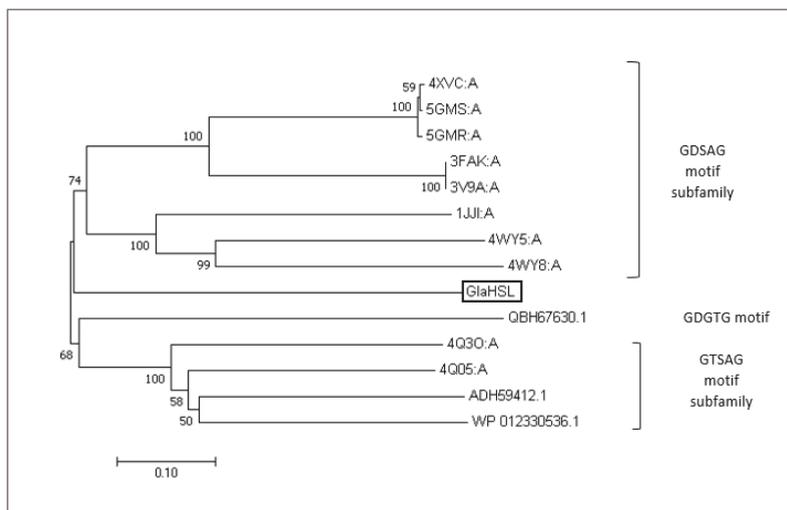GlaEst12-like esterase was cloned and expressed in an *E. coli* BL21(De3)/pET32b(+) expression system, which resulted in the accumulation of expressed GlaEst12 in the form of inclusion bodies. It is well known that high-level expression recombinant protein in *E. coli* is usually formed of partially folded or misfolded protein. HSL-like esterases from *Psychrobacter* sp. TA144 [20] and *Mycobacterium tuberculosis* LIPY [22] were also expressed as inclusion bodies. In the case of recombinant GlaEst12, the active enzyme was successfully renatured (Figure 2). A protein band corresponding to GlaEst12-like esterase with an expected size of 63 kDa was obtained as visualised by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). High expression of GlaEst12 in inclusion bodies leads to solubilisation and refolding to recover the bioactive protein. The proper protein folding of the aggregated protein in inclusion bodies was achieved by solubilising it with a high concentration of urea and then refolded by dilution. The refolded GlaEst12 esterase showed the highest esterase activity that revealed the successful solubilisation of GlaEst12 from inclusion bodies. The expression of recombinant GlaEst12 was optimised with the temperature, induction time, and isopropyl β-D-1-thiogalactopyranoside concentration at 16 °C, 20 h, and 10 μM, respectively.

The crude refolded GlaEst12 was purified in one-step purification using nickel sepharose affinity chromatography. N-terminal of polyhistidine (His-tag) in pET32b(+) vector was fused and expressed together with esterase, which enables the protein-containing polyhistidine to bind the specific immobilised Nickel (II) ions [23]. The crude refolded GlaEst12 was loaded into a nickel sepharose column, and the bound of protein was eluted using an ascending wise gradient of imidazole concentration. The bound protein was eluted at 300 mM Imidazole concentration and was checked for the presence of target protein by performing a lipase assay and SDS-PAGE. Figure 3 shows a single band on SDS-PAGE with a molecular weight of 63 kDa, which is consistent with the size of GlaEst12 (i.e., 45 kDa) fused with an 18 kDa pET32b(+) vector, indicating the successful purification of GlaEst12 esterase. The esterase was purified for homogeneity with a 40% yield and a purification fold of 1.72. The size of protein GlaEst12 was compared with the protein marker. Unlike with other HSL esterases,

the molecular mass of GlaEst12 is 45 kDa, which is slightly higher than the reported range of molecular mass between 30–40 kDa [24–26].
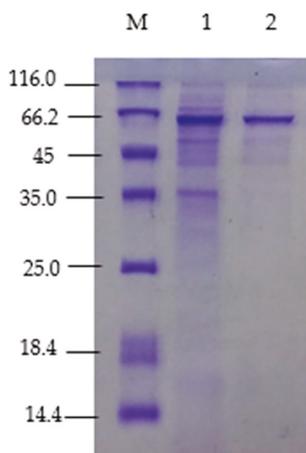

Figure 3. The SDS-PAGE analysis of purified GlaEst12 esterase. Lane M: unstained protein marker (Thermo Scientific, Waltham, MA, USA); Lane 1: refolded GlaEst12; Lane 2: purified GlaEst12-like esterase via Ni-Sepharose affinity chromatography.

2.3. Characterisation of Purified GlaEst12

2.3.1. Effect of Temperature on GlaEst12 Esterase Activity and Stability

The effect of temperature on GlaEst12 esterase activity and stability was studied by measuring the activity from 10–70 °C with an interval of 10 °C. Interestingly, the purified GlaEst12 has a broad temperature from 10–70 °C with an optimum temperature at 60 °C (Figure 4a) with 980 U/mL, and the activity of GlaEst12 dropped drastically at 70 °C. This indicated that GlaEst12 esterase exhibits thermophilic characteristics rather than psychrophilic, in which most of the reported enzymes from psychrophilic organisms have activity at low temperature with an optimal temperature range between 0–30 °C [27]. However, this is not the first reported enzyme from a psychrophilic microorganism that has broad temperature because there are already reported microbes isolated from the cold environment that appeared to produce thermotolerant lipases, such as AMS3 lipase from *Pseudomonas* sp. [28] and lipase ZJB09193 from *Candida antarctica* [29]. Another explanation of why GlaEst12 is able to withstand at a higher temperature is because of the presence of three cysteines in the amino acid composition. The cysteine consists of a thiolate group in the side chain that will form a disulphide bond that increases the rigidity of the protein, which plays a role in thermostability [27]. Figure 4b shows the thermostability of GlaEst12, which was tested by incubating the enzyme at 10–70 °C for 30 min. The results showed GlaEst12 was most stable at 50 °C and when incubated at a lower temperature range from 10–40 °C, the reduction of enzyme activity less than 40%. Thermal stability is one of the important criteria for making an enzyme to be used for industrial purposes [30]. Esterases or lipases that have optimal activity at low or high temperatures make them a versatile biocatalyst [31].
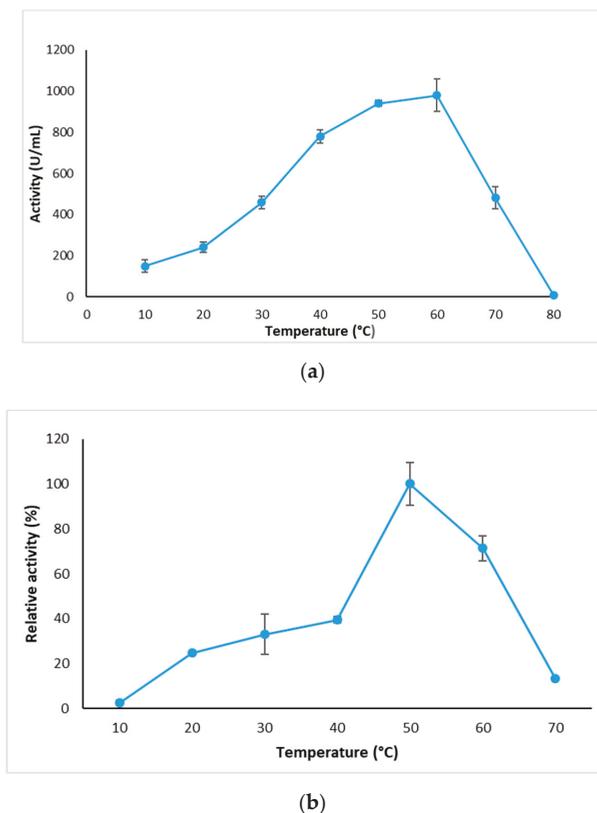
(**a**)

(**b**)

Figure 4. Effect of temperature on enzyme (**a**) activity and (**b**) stability of purified GlaEst12. The optimal temperature was determined by measuring the enzymatic activity at different temperatures ranged from 10–70 °C by using C10 as a substrate. The maximum optimal activity was observed at 60 °C. The temperature stability of purified esterase was determined by measuring the residual activities after the enzyme had been incubated for 30 min at different temperatures (10–80 °C) and the assay was performed at optimum temperature. Error bar represents standard deviation (*n* = 3). The absence of the bar indicates the error smaller than symbols.

2.3.2. Effects of pH on GlaEst12 Activity and Stability

The effect of pH on purified GlaEst12 esterase was tested on the different buffers with different pH ranging from 4–11. This esterase showed maximum activity at pH 8 using the Tris-HCl buffer. The GlaEst12 tended to be stable at a pH ranging from 7 to 9. Figure 5a shows the increasing trends of enzyme activity from pH 6 of sodium acetate to pH 7 and 8 of sodium phosphate, which peaked at pH 8 of Tris-HCl and then decreased gradually from pH 9 of Tris-HCl to pH 9–10 of Glycine-OH. Extreme acidic and alkaline buffers (i.e., pH less than 6 and pH more than 10, respectively) exhibited unfavourable conditions for this esterase with enzyme activity less than 100 U/mL. The pH stability of GlaEst12 was studied by treating the enzyme with various buffers for 30 min. Then, the residual activity after incubation was measured, and the highest activity was denoted as 100% relative activity, as shown in Figure 5b. The pH stability shows a similar pattern to the effects of pH on the enzyme activity since the GlaEst12 esterase showed the most stable in Tris-HCl pH 8 and more than 50% of residual activity stable at pH 7–9. Moreover, the result showed similar findings as reported in the other hormone-sensitive lipases, such as from *Psychrobacter* sp. TA144 and *R. miehei*, which have higher

activity and tend to be stable at pH 8 [11,20]. The buffer with pH range between 4–6 showed less than 10% enzyme activity, suggesting that the extreme acidic condition may affect the secondary structures, which ultimately leads to the reduction of the esterase activity.
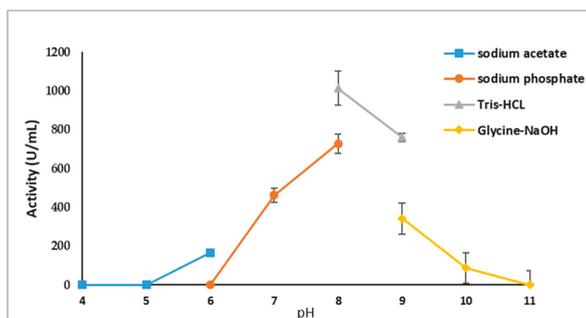
(a)
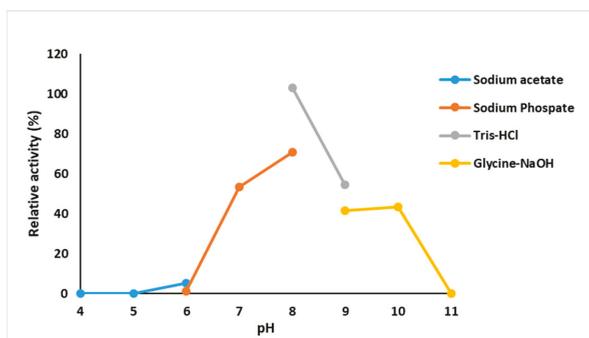
(b)

Figure 5. Effect of purified GlaEst12 esterase on enzyme (**a**) activity and (**b**) stability. The optimal pH was determined by measuring enzyme activity using pNP (C10) as a substrate in different buffer systems ranging from pH 4–11. The pH stability was determined by incubating an enzyme at different buffers for 30 min and measuring the residual activity at optimum pH. The buffer systems were used: sodium acetate (pH 4–6) (blue, filled square); sodium phosphate (pH 6–8) (orange, filled circle); Tris-HCl (pH 8–9) (grey, filled triangle), and glycine-OH (pH 9–11) (yellow, filled diamond). Error bar represents standard deviation (*n* = 3). The absence of the bar indicates the error smaller than symbols.

2.3.3. Substrate Specificity of GlaEst12 Esterase

The study on the substrate specificity of GlaEst12 was examined using various *p*-nitrophenyl (pNP) esters with an acyl chain length from C2–C16 using standard assay. The esterase showed high substrate specificity toward middle chain esters, pNP decanoate rather than a shorter or a longer chain of pNP ester. Figure 6 shows that more than 80% of activity was achieved when GlaEst12 used C8 as a substrate, while about 50% activity dropped when this esterase was assayed with a longer chain that is more than 10 carbons. The shorter chain of carbon, such as C2 and C4, had the lowest activity, which was less than 10%, indicating that the GLA has a low specificity toward the shorter carbon chain. The results showed differently from other esterases, for example, HSL esterase that has a protein sequence similarity with GlaEst12, E40 had the highest activity toward pNP butyrate (C4) [32], RmEstA from *R. miehei* prefers pNP hexanoate (C5) [24] and Est22 from deep-sea metagenomic library has the highest activity on pNP butyrate [33].
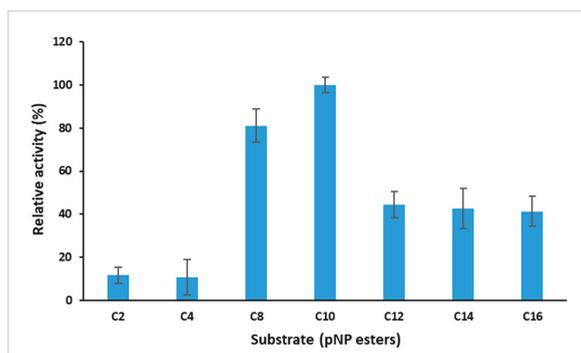
Figure 6. Effect of different pNP esters on purified GlaEst12 esterase. The activity of esterase was measured using different pNP esters at 60 °C using 50 mM Tris-HCl pH 8. The highest activity with p-nitrophenol decanoate (C10) substrate is shown as 100%. Error bar represents standard deviation ($n = 3$).

2.3.4. Effect of Metal Ions on Esterase Activity

The importance of metal ions in enzyme catalysis is well established since there are many reported metal-dependent enzymes that enhanced enzyme activity. Each metal ion has different roles since they may play an important role in a redox reaction, stabilisation of negative charges, and activation of substrates by virtue of their Lewis acid properties [34]. Effect of metal ions on GlaEst12-like esterase was conducted by treating the enzyme with various metals ions at the concentration of 1 mM and 5 mM. Figure 7 shows that metal ions (Na^+, K^+, Ca^{2+}, and Mn^{2+}) enhanced the activity, which was higher than that in the control (enzyme without metal ions). However, 1 mM and 5 mM of Mg^{2+}, Ni^{2+}, and Cu^{2+} decreased and abolished the GlaEst12 esterase activity. For Rb^+, lower concentration showed an increase in the enzyme activity, but the high concentration of Rb^+ had a negative effect on the enzyme activity. Most of the experiments that involved the HSL-like esterases showed that Ni^{2+} and Cu^{2+} tended to decrease the enzyme activity, such as EstAG1 from *Staphylococcus saprophyticus* and RmEsT from *R. miehei* [11,35].
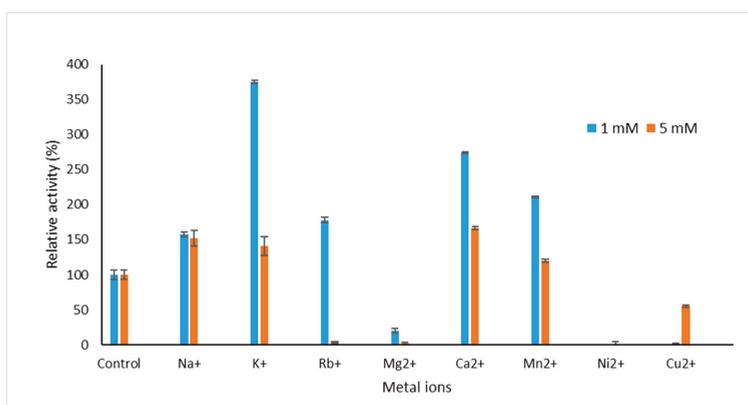

Figure 7. Effect of metal ions on the activity of purified GlaEst12-like esterase. The relative activity of the unincubated enzyme without metal ions (control) was taken 100%. Error bar represents standard deviation ($n = 3$).

2.3.5. Effect of Organic Solvents on GlaEst12

The stability of GlaEst12 esterase on organic solvents was studied by incubating the enzyme with polar and non-polar organic solvents based on log P-value. Figure 8 reveals the activity of GlaEst12 that is increasing with dimethyl sulfoxide (DMSO) (104%), 1-propanol (123%), and Toluene (113%) compared to the control. However, other solvents such as methanol, acetonitrile, benzene, octanol, xylene, and *n*-hexane caused instability in the protein. Among organic solvents, DMSO tends to give better stability to the GlaEst12 and other HSL esterases, consistent with many previous studies that have reported that this solvent is able to enhance lipolytic activities, such as RmEstB esterase from *R. miehei*, rEst1 from *Rheinheimera* sp., and EstAG1 from *S. saprophyticus* [11,35,36]. Based on these results, the enzyme showed less tolerance to the organic solvents since they were unable to resist the denaturation by the organic solvent, and the presence of these solvents may prevent accessibility of substrate to the active site [37].
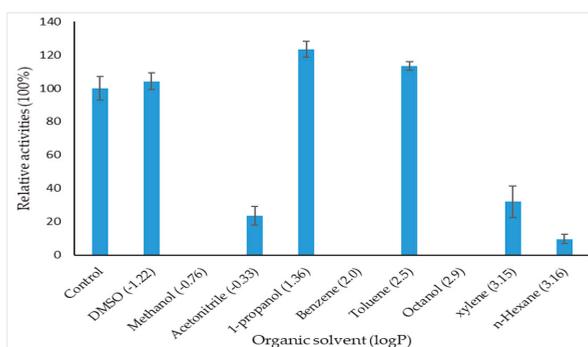

Figure 8. Effect of various organic solvents on the activity of purified GlaEst12-like esterase. The relative activity of the unincubated enzyme without organic solvent (control) was taken 100%. The log P is the logarithm of the partition coefficient, P, of the solvent between n-octanol and water, and is used as a quantitative measure of the polarity. Error bar represents standard deviation ($n = 3$).

2.4. Homology Modelling and Validation of GlaEst12

The homology modelling of GlaEst12 was done using the Robetta server (http://robetta.bakerlab.org/). The software uses two approaches to predict the structure, namely, comparative modelling or *de novo* structure prediction method. The de novo method was used when the query sequence was not matched with the template sequence and is known as the de novo Rosetta fragment insertion method [38]. Based on the multiple sequence alignments result, a crystal structure of Esterase 40 from the bacterial HSL family was chosen as the template to generate a 3D structure of GlaEst12 because it gave a higher score of sequence identity (about 30%), and the structure E40 was already solved using X-ray diffraction method [32]. Figure 9a shows the predicted model of this esterase that exists as a dimer comprised of two monomers of the subunits. Each monomer is dominated by 33.08% of α-helix followed by 9.52% and 57.39% of β-sheets and others, respectively. A higher number of α-helix present in the structure might be helping the survival of this enzyme in the Antarctica environment because an increase in the number of α-helix in protein structure tends to make the enzyme more flexible, which is responsible for enzyme activity at low temperatures [39]. The active site of GlaEst12-like esterase was predicted to be at position Ser 232, Glu 341, and His 371 (Figure 9b), which plays an important role in allowing the accessibility of the substrate. Serine residue located at the active site acts as a nucleophile, which is responsible for attacking the carbonyl group of the substrate, and this reaction later forms a tetrahedral intermediate together with the substrates, i.e., histidine and glutamate. In contrast with other esterases and lipases, mammalian HSL and the HSL-like esterase group exhibited conserved motif HGGG sequences. This sequence usually forms a loop in the secondary structure that is located

in close proximity to the active site and contributes to the formation and stabilisation of the oxyanion hole [40].

(**a**)

(**b**)
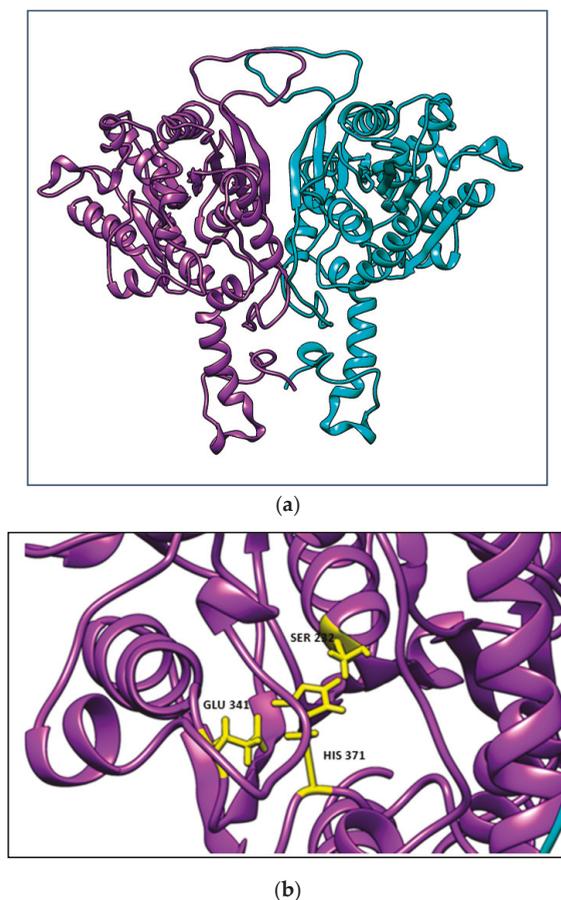
Figure 9. (**a**) The predicted GlaEst12 exists as dimer composed of two chain A (purple) and B (cyan) of α-helix structure, β-sheet, and coiled structures. (**b**) The catalytic triad of GlaEst12 was positioned at Ser^{232}, Glu^{341}, and His^{371} and depicted in yellow.

The assessment of protein models with 3D profiles was performed using online websites with the predicted structure of HSL esterase as a subject (Table 1). VERIFY 3D was used to determine the accuracy of the atomic model (3D), where the result was generated by comparing the subject with the structures that had already been solved by crystallography or the nuclear magnetic resonance (NMR) method. From the results, it showed that GlaEst12 has 87.8% residues of amino acid that scored equal and above 0.2 in Verify 3D. Although the value of score was lower than 90%, this structure is accepted because the residues have low scores at the N-terminal region, and the GLA esterase sequence is mostly conserved only at the central region as revealed by the multiple sequence alignment (Figure 1). This result showed consistency with the previous study, which stated that HSL lipase from the psychrophilic *Psychrobacter* sp. has sequence similarity with other homologous HSL proteins at the central region to the catalytic region. However, this psychrophilic enzyme has an additional sequence at the N-terminal region, which is expected to be the additional domain unique to the cold-adapted protein [20]. The addition of four α-helix domains at the N-terminal in GlaEst12 comparing to the other HSLs might

support the facts of the additional domain in HSL lipase from *Psychrobacter* since both of them are from the psychrophilic Antarctic. Besides that, the Errat tool was used to determine the accuracy and exactness of the atom distribution in the protein region and GlaEst12 has a high score that is more than 90%. The predicted structure was validated using a Ramachandran plot and revealed that 84.8% of it, which is about 328 residues, was located at the favoured region, while the remaining 14.8%, 0.3%, and 0.1% located at allowed, general, and outlier, respectively. The residues located at the disallowed region contributed about 0.1% of the total residues together with one of the catalytic triad, which is serine at position 232. The presence of the catalytic triad serine at negative region suggested that the enzyme is an active conformation. In contrast, the predicted structure of AMS8 lipase revealed that the catalytic serine is located at the allowed region and the protein is a closed conformation since it has the lid structure that covers the active site [41].

Table 1. The summary score for the predicted structure of GlaEst12 esterase using online web tools.

	Validation Tools	Score (%)
(A)	Verify 3D	87.8
(B)	Errat	91.3
(C)	Ramachandran Plot	
	Most favoured region	84.8
	Additional allowed region	14.8
	Generously allowed region	0.3
	Disallowed region	0.1

3. Materials and Methods

3.1. Sequence Analysis of GlaEst12

Previously, a psychrophilic yeast named *G. antarctica* was successfully isolated from sea ice near Casey Research Station, Antarctica. The whole-genome sequencing of this organism was done using 454 pyrosequencing and Illumina technology, with the protein information of *G. antarctica* being deposited in the *Glaciozyma antarctica* Genome Database (GanDB) (www.mgi-nibm.my/glaciozyma_antarctica) [42]. The gene encoding for putative esterase was chosen and known as *Glaciozyma antarctica* hormone-sensitive lipase (GlaEst12) esterase. The protein sequence of GlaEst12 was analysed using the GenBank database BLASTp (http://www.ncbi.nih.gov) from the NCBI to search the protein similarity with the other proteins. The amino acid composition, molecular weight, and predicted pI value of GlaEst12 were determined using Expasy Tools (https://web.expasy.org/protparam/). The presence of the signal peptide was predicted using the online tool SignalP-5.0 server (http://www.cbs.dtu.dk/services/SignalP/). The sequences similarity and secondary structure information from aligned sequences were performed using ENDscript 2.0 (http://endscript.ibcp.fr). The phylogenetic tree was constructed using MEGA 7.0, whereby the GlasEst12 protein was aligned with eight additional proteins (accession numbers: WP_012330536.1, ADH59412, QBH67630.1, KX580963.1, 4WY8, 4WY5, 4QO5, and 4Q30). The alignment was generated using Clustal W, and the evolutionary history was inferred by using the Neighbour Joining method with a Jones-Taylor-Thornton (JTT) method.

3.2. Gene Synthesis, Bacteria Strains, and Plasmids

The sequence of GlaEst12-like esterase that encoded for 1200 nucleotides was sent for gene synthesis. Codon optimisation was performed based on preferred codons by *E. coli* to enhance GLA HSL lipase expression in the *E. coli* host system. This codon-optimised gene was synthesised together with restriction endonuclease EcoR1 and Xho1 placed at the beginning and at the end of the gene sequence (Integrated DNA Technologies, Coralville, IA, USA). This gene was also cloned into a cloning vector (pUCIDT) and supplied in the dried plasmids. Since the pUCIDT/GLA HSL plasmid was in the form of powder, plasmid resuspension was carried out according to the manufacturer's protocol (Integrated DNA Technologies, Coralville, IA, USA). For cloning and expression of the protein, pET32b

(Merck, Kenilworth, NJ, USA) was used together with *E. coli* BL21(De3) as vector and expression host, respectively.

3.3. Cloning of GlaEst12 in E. coli

The gene that encoded GLA HSL lipase gene was amplified by PCR using recombinant plasmid pUCIDT/GLA HSL as a template. A set of forward and reverse primers with EcoRI and XhoI restriction sites were designed based on an optimised GLA HSL esterase gene sequence. The forward and reverse sequences are 5′CGTGAATTCGATGTTGAGTCCTG-3′ and 5′GAGCTCTTAAAACTTCCCGTCTA-3′, respectively, in which the underlined nucleotide sequences represent the sequences of EcoRI and XhoI. The PCR product was purified using a Gel Extraction kit (GeneAll, Seoul, Korea) and then digested with restriction enzymes EcoRI and XhoI. The digested PCR product was cloned into a pET32b vector (Merck, Kenilworth, NJ, USA) and transformed into *E. coli* BL21(De3) in tributyrin-containing ampicillin agar plates. The agar plates were incubated at 37 °C for 16 h and followed by incubation at 4 °C for 24 h. The positively transformed colonies were indicated by the formation of halo zones of colonies in tributyrin agar supplemented with ampicillin.

3.4. Expression, Solubilisation, and Refolding of GlaEst12 Inclusion Bodies

The recombinant GlaEst12 was expressed using pET32(b) + vector and transformed into *E. coli* BL21(De3). The expression was induced using 10 μM IPTG at 16 °C for 20 h. Solubilisation of GlaEst12 was conducted as the enzyme was mostly expressed as inclusion bodies. The *E. coli* cell was harvested by centrifugation at 10,000× *g* for 15 min. Then, the supernatant was discarded, and the pellet was resuspended with 20 mL of 50 mM Tris-HCl (pH 8) and subjected to sonication for 6 min under the output of 2 and duty cycle of 20 (Sonifer® SLP150 Branson, Danbury, CT, USA). The clear lysate was centrifuged at 10,000× *g* for 15 min, and the pellet-containing insoluble protein was further resuspended with Tris-HCl buffer (pH 8) containing 8 M of urea. The resuspend mixture was then incubated at 4 °C for 4 h with constant agitation. After incubation, the mixture was centrifuged with the same condition stated above, and the supernatant was used for further reaction. Renaturation of the supernatant containing the GlaEst12-like esterase was achieved by a 10× dilution of the denaturant in 50 mM Tris-HCl buffer (pH 8). The solubilised protein was diluted in one-step with a peristaltic pump of the flow rate of 0.5 mL/min and stirred thoroughly at 4 °C. The refolded protein was then subjected to enzyme assay.

3.5. Purification of Recombinant of GlaEst12-Like Esterase

The His-tagged of recombinant GlaEst12 was purified by single-step Ni-sepharose affinity chromatography. The filtered crude protein was loaded onto a Nickel-Sepharose HP column (XK16/20) (GE Healthcare, Boston, MA, USA). The binding buffer [20 mM Sodium phosphate, 10 mM imidazole, 500 mM NaCl (pH 7.4)] was used to equilibrate the column at a flow rate of 1 mL/min. Then, the crude protein was loaded onto the column, and the bound protein was eluted with an ascending step gradient of elution buffer [20 mM Sodium phosphate, 500 mM imidazole, 500 mM NaCl (pH 7.4)]. The eluted proteins were collected in 2 mL per fraction. The fractions containing the protein of interest were confirmed through pNP assay and SDS-PAGE. The fractions containing protein of interest were pooled, dialysed with buffer [50 mM Tris-HCl, 50 mM NaCl (pH 8)], and stored at 4 °C for further characterisation. The molecular weight of GlaEst12 was determined by using SDS-PAGE with 6% stacking gel and 12% separating gel, as described by Laemmli., 1970, with some modification [43]. The gel was stained using Coomassie Brilliant Blue R 250 (BioRad, Hercules, CA, USA) and destained with a destaining solution. The molecular mass of the protein was estimated using a broad range of protein standard markers (unstained protein marker 18.4–116 kDa, Thermo Scientific, Waltham, MA, USA).

3.6. Enzyme Assay

A spectrophotometric method was used to determine the GlaEst12 activity using pNP substrate. The pNP released from the substrate was measured according to the method described by Sumby et al., 2009, with some modifications [44]. The mixture reaction consisted of 950 μL of 50 mM Tris-HCl (pH 8), 25 μL of 10 mM *p*-nitrophenyl decanoate (C10:0), and 25 μL of 0.1 mg/mL enzyme. The mixture was assayed with shaking at 150 rpm, 60 °C for 10 min. Then, the liberation of pNP was measured using Biochrom WPA UV/Visible spectrophotometer (Cambridge, UK) at 410 nm. The absorbance of the sample was deduced with the control that the mixture stated above without the enzyme. One unit of esterase was defined as 1.0 μmol of pNP released per min under the conditions stated above.

3.7. Characterisation of Purified GlaEst12

3.7.1. Effect of Temperature on Activity and Stability

The determination of the effective temperature of purified GlaEst12-like esterase on its activity was conducted by measuring the esterase activity (as mentioned in Section 3.6) assayed at different temperatures of 10–80 °C (10 °C interval) for 10 min. For thermostability, 25 μL of the enzyme was first incubated with 50 mM Tris-HCl pH8 at different temperatures of 10–70 °C (10 °C interval) for 30 min without substrate. Then, the residual of enzyme activity was assayed together with 10 mM *p*-nitrophenyl decanoate (C10) as substrate at the optimum temperature of 60 °C for 10 min.

3.7.2. Effect of pH and pH Stability

Different buffers were used to study and determine the optimum buffer for GlaEst12-like esterase under pH range from 4–11. The buffers used were 50 mM sodium acetate (pH 4.0–6.0), 50 mM sodium phosphate (pH 6.0–8.0), 50 mM Tris-HCl (pH 8.0–9.0), and 50 mM glycine-NaOH (pH 9.0–11.0). The pH stability was investigated by incubating the enzyme with different buffers as stated above at 60 °C for 30 min and followed by enzyme assay (same as in point 3.6).

3.7.3. Effect of Substrate Specificity

The substrate specificity was determined by p-nitrophenyl esters with various chain lengths, including *p*-nitrophenyl acetate (C2), *p*-nitrophenyl butyrate (C4), *p*-nitrophenyl octanoate (C8), *p*-nitrophenyl decanoate (C10), *p*-nitrophenyl laurate (C12), *p*-nitrophenyl myristate (C14), and *p*-nitrophenyl palmitate (C16). The reaction mixtures containing 25 μL of the purified enzyme, 950 μL of 50 mM Tris-HCl pH 8, and 10 mM of different substrates were assayed at 60 °C for 10 min.

3.7.4. Effect of Metals Ions

GlaEst12-like esterase was treated with 1 mM and 5 mM metal ions (i.e., Li^+, Na^+, K^+, Rb^{2+}, Mg^{2+}, Ca^{2+}, Mn^{2+}, Ni^{2+}, Cu^{2+}). The treated enzyme was then subjected to enzyme assay. For 1 mM of metal ions, 940 μL of Tris-HCl buffer pH 8, 25 μL of the enzyme was treated with 10 μL of metal ions for 30 min at 60 °C. Then, 25 μL of 10 mM *p*-nitrophenyl decanoate (C10) was added to the mixture and assay, as mentioned in point 3.6. For 5 mM, all the composition are same as 1 mM except for the composition of buffer and metal ions, which used 900 μL and 50 μL, respectively. The stability was determined as the relative activity to the control (i.e., without a metal ion).

3.7.5. Effect of Organic Solvents

The esterase was incubated for 30 min at 60 °C with various organic solvents at a concentration of 25% (*v/v*). The solvents were selected based on their log P values (in parentheses): DMSO (−1.22), methanol (−0.76), acetonitrile (−0.33), 1-propanol (1.36) benzene (2.0), toluene (2.5), octanol (2.9), xylene (3.15), and *n*-hexane (3.16). The mixtures pre-incubate for 30 min, which contained 700 μL of 50 mM of Tris-HCl (pH 8), 25 μL of the enzyme, and 250 μL of organic solvents and later were assayed together

with 10 mM of *p*-nitrophenyl decanoate (C10) at 60 °C. The stability was determined as the relative activity to the control (i.e., without organic solvent).

3.8. Homology Modelling and Structure Validation

The homology modelling was used to predict 3D structure using templates deposited in the Protein Data Bank (PDB) that have high similarity to GlaEst12. The 3D structure of GlaEst12 was generated by using the Robetta server (http://robetta.bakerlab.org) that provides automated tools for protein structure prediction, while the figures were prepared using the Chimera visual system (www.cgl.ucsf.edu/chimera). The validation of protein structure was done using online software such as Ramachandran Plot (http://www-cryst.bioc.cam.ac.uk/), Errat [45], and VERIFY 3D [46].

4. Conclusions

A novel HSL-like esterase family known as GlaEst12 is being introduced from *G. antarctica*, a psychrophilic yeast. Multiple sequence alignment with another hormone-sensitive lipase proteins revealed GlaEst12 as a new member of the GDSAG motif subfamily of the HSL family. GlaEst12-like esterase was successfully expressed in *E. coli* and purified with single-step nickel-sepharose affinity chromatography. Biochemical characterisation of this esterase showed interestingly higher activity and stability at a higher temperature, which gives a unique feature to HSL-like esterase that was isolated from psychrophilic yeast. Besides that, this esterase was activated when treated with metal ions (Na^+, K^+, Ca^{2+}, Mn^{2+}) and stabilised when incubated with *1*-propanol and toluene. Homology modelling of this GlaEst12-like esterase showed the predicted structure of this enzyme that is composed of a typical α/β hydrolase fold with the catalytic residues found at Ser [232], Glu [341], and His [371]. The characterisation of GlaEst12 that can withstand a broad temperature and remain stable in an alkaline environment make it a potential catalyst in industrial application.

References

1. Jensen, M.B.V.; Horsfall, L.E.; Wardrope, C.; Togneri, P.D.; Marles-Wright, J.; Rosser, S.J. Characterisation of a New Family of Carboxyl Esterases with an OsmC Domain. *PLoS ONE* **2016**, *11*, e0166128. [CrossRef]

2. Fojan, P. What Distinguishes an Esterase from a Lipase: A Novel Structural Approach. *Biochimie* **2000**, *82*, 1033–1041. [CrossRef]

3. De Simone, G.; Mandrich, L.; Menchise, V.; Giordano, V.; Febbraio, F.; Rossi, M.; Pedone, C.; Manco, G. A Substrate-Induced Switch in the Reaction Mechanism of a Thermophilic Esterase: KINETIC EVIDENCES AND STRUCTURAL BASIS. *J. Biol. Chem.* **2004**, *279*, 6815–6823. [CrossRef]

4. Lampidonis, A.D.; Rogdakis, E.; Voutsinas, G.E.; Stravopodis, D.J. The Resurgence of Hormone-Sensitive Lipase (HSL) in Mammalian Lipolysis. *Gene* **2011**, *477*, 1–11. [CrossRef]

5. Langin, D.; Laurell, H.; Holst, L.S.; Belfrage, P.; Holm, C. Gene Organization and Primary Structure of Human Hormone-Sensitive Lipase: Possible Significance of a Sequence Homology with a Lipase of Moraxella TA144, an Antarctic Bacterium. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4897–4901. [CrossRef]

6. Kim, T.D. Bacterial Hormone-Sensitive Lipases (BHSLs): Emerging Enzymes for Biotechnological Applications. *J. Microbiol. Biotechnol.* **2017**, *27*, 1907–1915. [CrossRef]

7.   Pöhlmann, C.; Wang, Y.; Humenik, M.; Heidenreich, B.; Gareis, M.; Sprinzl, M. Rapid, Specific and Sensitive Electrochemical Detection of Foodborne Bacteria. *Biosens. Bioelectron.* **2009**, *24*, 2766–2771. [CrossRef]

8.   Febbraio, F.; Merone, L.; Cetrangolo, G.P.; Rossi, M.; Nucci, R.; Manco, G. Thermostable Esterase 2 from *Alicyclobacillus acidocaldarius* as Biosensor for the Detection of Organophosphate Pesticides. *Anal. Chem.* **2011**, *83*, 1530–1536. [CrossRef]

9.   Jaeger, K.E.; Eggert, T. Lipases for Biotechnology. *Curr. Opin. Biotechnol.* **2002**, *13*, 390–397. [CrossRef]

10. Bassegoda, A.; Fillat, A.; Pastor, F.I.J.; Diaz, P. Special *Rhodococcus* sp. CR-53 Esterase Est4 Contains a GGG(A)X-Oxyanion Hole Conferring Activity for the Kinetic Resolution of Tertiary Alcohols. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 8559–8568. [CrossRef]

11. Yan, Q.; Yang, S.; Duan, X.; Xu, H.; Liu, Y.; Jiang, Z. Characterization of a Novel Hormone-Sensitive Lipase Family Esterase from *Rhizomucor miehei* with Tertiary Alcohol Hydrolysis Activity. *J. Mol. Catal. B Enzym.* **2014**, *109*, 76–84. [CrossRef]

12. Petrovskaya, L.E.; Novototskaya-Vlasova, K.A.; Gapizov, S.S.; Spirina, E.V.; Durdenko, E.V.; Rivkina, E.M. New Member of the Hormone-Sensitive Lipase Family from the Permafrost Microbial Community. *Bioengineered* **2017**, *8*, 420–423. [CrossRef]

13. Li, P.Y.; Ji, P.; Li, C.Y.; Zhang, Y.; Wang, G.L.; Zhang, X.Y.; Xie, B.B.; Qin, Q.L.; Chen, X.L.; Zhou, B.C.; et al. Structural Basis for Dimerization and Catalysis of a Novel Esterase from the GTSAG Motif Subfamily of the Bacterial Hormone-Sensitive Lipase Family. *J. Biol. Chem.* **2014**, *289*, 19031–19041. [CrossRef]

14. Turchetti, B.; Thomas Hall, S.R.; Connell, L.B.; Branda, E.; Buzzini, P.; Theelen, B.; Müller, W.H.; Boekhout, T. Psychrophilic Yeasts from Antarctica and European Glaciers: Description of Glaciozyma Gen. Nov., *Glaciozyma martinii* sp. Nov. and *Glaciozyma watsonii* sp. Nov. *Extremophiles* **2011**, *15*, 573–586. [CrossRef]

15. Boo, S.Y.; Wong, C.M.V.L.; Rodrigues, K.F.; Najimudin, N.; Murad, A.M.A.; Mahadi, N.M. Thermal Stress Responses in Antarctic Yeast, *Glaciozyma antarctica* PI12, Characterized by Real-Time Quantitative PCR. *Polar Biol.* **2013**, *36*, 381–389. [CrossRef]

16. Alias, N.; Mazian, A.; Salleh, A.B.; Basri, M.; Rahman, R.N.Z.R. Molecular Cloning and Optimization for High Level Expression of Cold-Adapted Serine Protease from Antarctic Yeast *Glaciozyma antarctica* PI12. *Enzym. Res.* **2014**, *2014*, 1–20. [CrossRef]

17. Hashim, N.H.F.; Sulaiman, S.; Abu Bakar, F.D.; Illias, R.M.; Kawahara, H.; Najimudin, N.; Mahadi, N.M.; Murad, A.M.A. Molecular Cloning, Expression and Characterisation of Afp4, an Antifreeze Protein from *Glaciozyma antarctica*. *Polar Biol.* **2014**, *37*, 1495–1505. [CrossRef]

18. Ramli, A.N.M.; Azhar, M.A.; Shamsir, M.S.; Rabu, A.; Murad, A.M.A.; Mahadi, N.M.; Illias, R.M. Sequence and Structural Investigation of a Novel Psychrophilic α-Amylase from *Glaciozyma antarctica* PI12 for Cold-Adaptation Analysis. *J. Mol. Model.* **2013**, *19*, 3369–3383. [CrossRef]

19. Ramli, A.; Mahadi, N.; Rabu, A.; Murad, A.; Bakar, F.; Illias, R. Molecular Cloning, Expression and Biochemical Characterisation of a Cold-Adapted Novel Recombinant Chitinase from *Glaciozyma antarctica* PI12. *Microb. Cell Fact.* **2011**, *10*, 94. [CrossRef]

20. De Santi, C.; Tutino, M.L.; Mandrich, L.; Giuliani, M.; Parrilli, E.; Del Vecchio, P.; de Pascale, D. The Hormone-Sensitive Lipase from *Psychrobacter* sp. TA144: New Insight in the Structural/Functional Characterization. *Biochimie* **2010**, *92*, 949–957. [CrossRef]

21. Li, Z.; Su, L.; Wang, L.; Liu, Z.; Gu, Z.; Chen, J.; Wu, J. Novel Insight into the Secretory Expression of Recombinant Enzymes in *Escherichia coli*. *Process Biochem.* **2014**, *49*, 599–603. [CrossRef]

22. Deb, C. A Novel Lipase Belonging to the Hormone-Sensitive Lipase Family Induced under Starvation to Utilize Stored Triacylglycerol in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **2006**, *281*, 3866–3875. [CrossRef] [PubMed]

23. Bornhorst, J.A.; Falke, J.J. Purification of Proteins Using Polyhistidine Affinity Tags. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 2000; Volume 326, pp. 245–254.

24. Liu, Y.; Xu, H.; Yan, Q.; Yang, S.; Duan, X.; Jiang, Z. Biochemical Characterization of a First Fungal Esterase from *Rhizomucor miehei* Showing High Efficiency of Ester Synthesis. *PLoS ONE* **2013**, *8*, e77856. [CrossRef]

25. Jeon, J.H.; Lee, H.S.; Kim, J.T.; Kim, S.J.; Choi, S.H.; Kang, S.G.; Lee, J.H. Identification of a New Subfamily of Salt-Tolerant Esterases from a Metagenomic Library of Tidal Flat Sediment. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 623–631. [CrossRef]

26. Jin, P.; Pei, X.; Du, P.; Yin, X.; Xiong, X.; Wu, H.; Zhou, X.; Wang, Q. Overexpression and Characterization of a New Organic Solvent-Tolerant Esterase Derived from Soil Metagenomic DNA. *Bioresour. Technol.* **2012**, *116*, 234–240. [CrossRef]

27. Salwoom, L.; Rahman, R.A.; Zaliha, R.N.; Salleh, A.B.; Convey, P.; Ali, M.; Shukuri, M. New Recombinant Cold-Adapted and Organic Solvent Tolerant Lipase from Psychrophilic *Pseudomonas* sp. LSK25, Isolated from Signy Island Antarctica. *Int. J. Mol. Sci.* **2019**, *20*, 1264. [CrossRef]

28. Latip, W.; Rahman, R.N.; Leow, A.T.; Shariff, F.M.; Ali, M.S. Expression and Characterization of Thermotolerant Lipase with Broad PH Profiles Isolated from an Antarctic *Pseudomonas* sp. Strain AMS3. *PeerJ* **2016**, *4*, e2420. [CrossRef]

29. Liu, Z.Q.; Zheng, X.B.; Zhang, S.P.; Zheng, Y.G. Cloning, Expression and Characterization of a Lipase Gene from the *Candida antarctica* ZJB09193 and Its Application in Biosynthesis of Vitamin A Esters. *Microbiol. Res.* **2012**, *167*, 452–460. [CrossRef]

30. Joseph, B.; Ramteke, P.W.; Thomas, G. Cold Active Microbial Lipases: Some Hot Issues and Recent Developments. *Biotechnol. Adv.* **2008**, *26*, 457–470. [CrossRef]

31. Lopez-Lopez, O.; Cerdan, M.; Siso, M. New Extremophilic Lipases and Esterases from Metagenomics. *Curr. Protein Pept. Sci.* **2014**, *15*, 445–455. [CrossRef]

32. Li, P.Y.; Chen, X.L.; Ji, P.; Li, C.Y.; Wang, P.; Zhang, Y.; Xie, B.B.; Qin, Q.L.; Su, H.N.; Zhou, B.C.; et al. Interdomain Hydrophobic Interactions Modulate the Thermostability of Microbial Esterases from the Hormone-Sensitive Lipase Family. *J. Biol. Chem.* **2015**, *290*, 11188–11198. [CrossRef] [PubMed]

33. Huang, J.; Huo, Y.Y.; Ji, R.; Kuang, S.; Ji, C.; Xu, X.W.; Li, J. Structural Insights of a Hormone Sensitive Lipase Homologue Est22. *Sci. Rep.* **2016**, *6*, 28550. [CrossRef] [PubMed]

34. Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G.L.; Thornton, J.M. Metal Ions in Biological Catalysis: From Enzyme Databases to General Principles. *J. Biol. Inorg. Chem.* **2008**, *13*, 1205–1218. [CrossRef] [PubMed]

35. Gricajeva, A.; Bikutė, I.; Kalėdienė, L. Atypical Organic-Solvent Tolerant Bacterial Hormone Sensitive Lipase-like Homologue EstAG1 from Staphylococcus Saprophyticus AG1: Synthesis and Characterization. *Int. J. Biol. Macromol.* **2019**, *130*, 253–265. [CrossRef] [PubMed]

36. Virk, A.P. A New Esterase, Belonging to Hormone-Sensitive Lipase Family, Cloned from *Rheinheimera* sp. Isolated from Industrial Effluent. *J. Microbiol. Biotechnol.* **2011**, *21*, 667–674. [CrossRef] [PubMed]

37. Bansal, V.; Delgado, Y.; Fasoli, E.; Ferrer, A.; Griebenow, K.; Secundo, F.; Barletta, G.L. Effect of Prolonged Exposure to Organic Solvents on the Active Site Environment of Subtilisin Carlsberg. *J. Mol. Catal. B Enzym.* **2010**, *64*, 38–44. [CrossRef] [PubMed]

38. Kim, D.E.; Chivian, D.; Baker, D. Protein Structure Prediction and Analysis Using the Robetta Server. *Nucleic Acids Res.* **2004**, *32*, W526–W531. [CrossRef]

39. D'Amico, S.; Collins, T.; Marx, J.C.; Feller, G.; Gerday, C. Psychrophilic Microorganisms: Challenges for Life. *EMBO Rep.* **2006**, *7*, 385–389. [CrossRef]

40. Rehdorf, J.; Behrens, G.A.; Nguyen, G.S.; Kourist, R.; Bornscheuer, U.T. Pseudomonas Putida Esterase Contains a GGG(A)X-Motif Confering Activity for the Kinetic Resolution of Tertiary Alcohols. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 1119–1126. [CrossRef]

41. Ali, M.S.M.; Ganasen, M.; Rahman, R.N.Z.R.A.; Chor, A.L.T.; Salleh, A.B.; Basri, M. Cold-Adapted RTX Lipase from Antarctic *Pseudomonas* sp. Strain AMS8: Isolation, Molecular Modeling and Heterologous Expression. *Protein J.* **2013**, *32*, 317–325. [CrossRef]

42. Firdaus-Raih, M.; Hashim, N.H.F.; Bharudin, I.; Abu Bakar, M.F.; Huang, K.K.; Alias, H.; Lee, B.K.B.; Mat Isa, M.N.; Mat-Sharani, S.; Sulaiman, S.; et al. The *Glaciozyma antarctica* Genome Reveals an Array of Systems That Provide Sustained Responses towards Temperature Variations in a Persistently Cold Habitat. *PLoS ONE* **2018**, *13*, e0189947. [CrossRef] [PubMed]

43. Laemmli, U.K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227*, 680–685. [CrossRef] [PubMed]

44. Sumby, K.M.; Matthews, A.H.; Grbin, P.R.; Jiranek, V. Cloning and Characterization of an Intracellular Esterase from the Wine-Associated Lactic Acid Bacterium *Oenococcus oeni*. *Appl. Environ. Microbiol.* **2009**, *75*, 6729–6735. [CrossRef] [PubMed]

45. Colovos, C.; Yeates, T.O. Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions. *Protein Sci.* **1993**, *2*, 1511–1519. [CrossRef] [PubMed]

46. Luthy, R.; Bowie, J.U.; Eisenberg, D. Assesment of Protein Models with 3 Dimensional Profile. *Nature* **1992**, *356*, 83–85. [CrossRef]

Article

Characterization of the Novel Ene Reductase Ppo-Er1 from *Paenibacillus Polymyxa*

David Aregger, Christin Peters and Rebecca M. Buller *

Competence Center for Biocatalysis, Institute of Chemistry and Biotechnology, Department of Life Sciences and Facility Management, Zurich University of Applied Sciences, Einsiedlerstrasse 31, 8820 Waedenswil, Switzerland; David.Aregger@zhaw.ch (D.A.); Christin.peters@zhaw.ch (C.P.)
* Correspondence: rebecca.buller@zhaw.ch; Tel.: +41-58-934-5438

Abstract: Ene reductases enable the asymmetric hydrogenation of activated alkenes allowing the manufacture of valuable chiral products. The enzymes complement existing metal- and organocatalytic approaches for the stereoselective reduction of activated C=C double bonds, and efforts to expand the biocatalytic toolbox with additional ene reductases are of high academic and industrial interest. Here, we present the characterization of a novel ene reductase from *Paenibacillus polymyxa*, named Ppo-Er1, belonging to the recently identified subgroup III of the old yellow enzyme family. The determination of substrate scope, solvent stability, temperature, and pH range of Ppo-Er1 is one of the first examples of a detailed biophysical characterization of a subgroup III enzyme. Notably, Ppo-Er1 possesses a wide temperature optimum (T_{opt}: 20–45 °C) and retains high conversion rates of at least 70% even at 10 °C reaction temperature making it an interesting biocatalyst for the conversion of temperature-labile substrates. When assaying a set of different organic solvents to determine Ppo-Er1's solvent tolerance, the ene reductase exhibited good performance in up to 40% cyclohexane as well as 20 vol% DMSO and ethanol. In summary, Ppo-Er1 exhibited activity for thirteen out of the nineteen investigated compounds, for ten of which Michaelis–Menten kinetics could be determined. The enzyme exhibited the highest specificity constant for maleimide with a k_{cat}/K_M value of 287 mM^{-1} s^{-1}. In addition, Ppo-Er1 proved to be highly enantioselective for selected substrates with measured enantiomeric excess values of 92% or higher for 2-methyl-2-cyclohexenone, citral, and carvone.

Keywords: biocatalysis; ene reductase; enzyme sourcing; old yellow enzyme; solvent stability

1. Introduction

Many bioactive molecules contain at least one chiral center rendering the development of effective asymmetric synthesis methods essential for the chemical industry. Besides the well-established metal- and organocatalytic approaches [1], biocatalytic strategies offer an interesting alternative to install chirality into small molecules. To date, industrial biocatalysis has mastered a range of enzyme families including ketoreductases [2], transaminases [3], and imine reductases [4]. Looking forward, the increasing power of genomic mining and enzyme engineering will allow industrial access to even more enzyme families leading to an expansion of the available biocatalytic toolbox [5].

The families of enzymes collectively known as ene reductases (ERs) catalyze the stereoselective trans- and, more rarely, cis-hydrogenation of activated alkenes [6–9]. Thus, ene reductases offer a valuable access route to asymmetric compounds, which is complementary to the chemical cis-hydrogenation catalyzed by chiral rhodium or ruthenium phosphine catalysts [10,11]. Today, ene reductases are classified into five enzyme groups, which differ in structure, reaction mechanism, substrate spectrum, and stereoselectivity (Figure 1) [12]. While enoate reductases, medium- and short-chain dehydrogenases/reductases (MDR and SDR), as well as the recently discovered quinone reductase-like ene reductases [13], are currently being investigated in terms of their industrial

potential [14], enzymes stemming from the old yellow enzyme (OYE) family are established members of the biocatalytic toolbox and are the best characterized and most extensively employed ene reductases today [6].
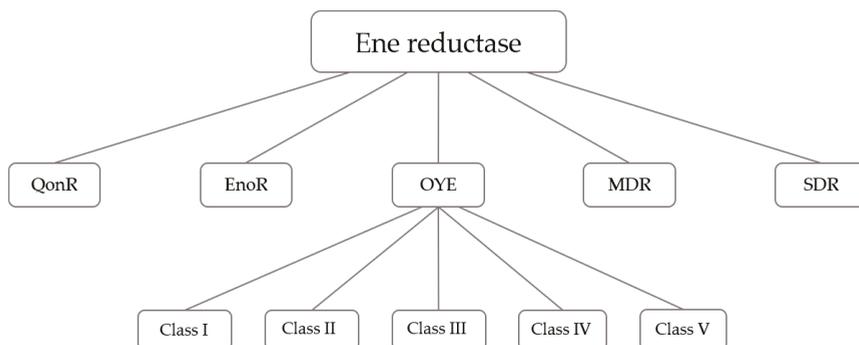

Figure 1. Overview of the classification within the ene reductase family [15]. QnoR (NADPH-dependent quinone reductase like ene-reductases), EnoR (enoate reductase), OYE (old yellow enzyme), MDR (medium-chain dehydrogenase/reductase), and SDR (short-chain dehydrogenase/reductase); Class I (classical OYE); Class II (thermophilic-like OYE) and Class V (fungal OYE).

Isolated in 1932 by Warburg and Christian from bottom-fermented brewer's yeast (*Saccharomyces pastorianus*), the first such ene reductase was named "yellow enzyme" [16]. After the discovery of several additional members belonging to the same enzyme family the "yellow enzyme" was renamed to "old yellow enzyme" (OYE1) [17]. OYEs preferentially accept α,β-unsaturated ketones, aldehydes, nitroalkenes, and some carboxylic acids as substrates [7]. In the last decade, the catalytic mechanism of OYEs has been exhaustively investigated and its general principle is well understood: The enzymes follow a bi–bi ping–pong mechanism, which can be divided into a reductive and an oxidative half reaction [18]. In the reductive half-reaction, flavin mononucleotide (FMN) is reduced through hydride transfer from NAD(P)H, whereas in the oxidative half reaction a hydride is transferred from the reduced flavin to the C_β of the activated alkene. The missing proton for the C_α is transferred via a tyrosine residue from the opposite site [18,19], ultimately leading to an anti-addition hydrogenation.

The catalytic machinery of OYE enzymes is supported by a typical (α,β) 8-barrel (TIM-barrel) fold with additional secondary structural elements present (e.g., four β-strands and five α-helices in OYE1 [20]; six β-strands and two α-helices in 12-oxophytodienoate reductase OPR [18]). The folded domain is known to occur in different oligomeric states, such as monomers (PETN reductase) [21], dimers (OYE1) [20], tetrameters (dimers of dimers such as YqjM [22] or TOYE [23]), octamers, and dodecamers [23]. The oligomerization state is described to be often governed by the position and amino acid composition of surface loops [7]. In addition, the constitution of the loops can have an influence on thermostability [23].

Notably, amino acid sequence alignments of OYE homologs show high conservation in specific regions of the proteins, such as residues involved in catalysis, FMN, and substrate binding [7,15,23]. To account for these differences in sequence and the resulting structural features, the old yellow enzyme family can be further divided into five subclasses [15]. While enzyme members of the subclass I, also termed "classical" old yellow enzymes, and class II, introduced by Scrutton's group in 2010 and dubbed "thermophilic-like" [23], have been well explored [7,14], the recently described class III–V are less well investigated [15,24].

Synthetic applications of ene reductases are manifold and range from the preparation of profens [25–27] and chiral γ-amino acids [28–30] to the synthesis of chiral phosphonates [31] and nitroalkanes [32], precursors in the synthesis of pharmaceutically active ingredients. To further

promote an off-the-shelve synthetic use of ene reductases, which can reduce the time and cost of the implementation of a biocatalytic step into a process significantly, we set out to expand the available biocatalytic toolbox [15]. In this context, not only the discovery and engineering of novel ene reductases is of great utility [33], but also a careful characterization of the new biocatalysts is needed as it may lead to the construction of a more targeted enzyme library associated with reduced screening time and costs.

Herein, we showcase the detailed characterization of Ppo-Er1 from *Paenibacillus polymyxa*, an OYE subclass III enzyme, and highlight the enzyme's substrate scope, kinetic parameters, solvent tolerance, as well as pH and temperature profile. The data presented may facilitate future screening and engineering studies and, in selected cases, thus, lead to the faster adoption of an ene reductase in chemical process development.

2. Results and Discussion

The enzyme Ppo-Er1 from *P. polymyxa* was discovered during the screening of 19 bacterial wild-type strains from the Culture Collection of Switzerland, as previously described [15]. Ppo-Er1 (41.3 kDa) is characterized by a substantial sequence similarity with the old yellow enzyme YqiG from *Bacillus subtilis* (50%) [34], Bac-OYE2 from *Bacillus* sp. (50%) [35], Lla-Er from *Lactococcus lactis* (39%) [15], and LacER from *Lactobacillus paracasei* (47%) [36], all of which belong to the subclass III of the OYE family. In detail, Ppo-Er1 contains a specific combination of motifs known from the classical and thermophilic-like groups that has been found to be characteristic for class III enzymes [15]: Gln104 and Arg228 predicted to interact with the pyrimidine ring of FMN [22], His 171, and Asn 175 proposed to interact with N1 and N3 of FMN [22,37]; Thr30 suggested to interact with isoalloxazine ring O4 of FMN [38]; and Met29, Leu324, and Arg321, which presumably interact with the dimethyl benzene moiety of FMN. As expected, subclass III old yellow enzyme Ppo-Er1 is thus phylogenetically positioned between classical and thermophilic-like OYEs.

2.1. Expression and Characterization of Ppo-Er1

The ready-to-use plasmid consisting of pET-28b(+) vector and the Ppo-Er1 sequence was assembled by Twist Bioscience and a C-terminal His$_6$ tag for protein purification by affinity chromatography was included. The soluble recombinant expression of Ppo-Er1 in *Escherichia coli* BL21 (DE3) was achieved in terrific broth (TB) medium at 25 °C. Ppo-Er1 was purified by affinity chromatography using Ni-NTA resin (Figure S1) and the cofactor FMN was reconstituted before further analysis. FMN reconstitution (100 μM) proved necessary to obtain a fully active enzyme as without this step the enzyme preparation only exhibited 8% (0.05 U/mg for cyclohexanone) of the expected activity (0.61 U/mg for cyclohexanone). This effect was also described for the OYEs LacER [36] and Lla-Er [15]. In the case of LacER, for example, the addition of FMN after purification by DEAE ion exchange chromatography increased the activity by a factor of 92 from 0.0018 to 0.168 U/mg for the substrate *trans*-2-hexen-1-al. This observation suggests that—similar to other known OYEs—the binding affinity of Ppo-Er1 to FMN under purification conditions is low, a fact that has to be kept in mind for any following activity analysis. The storage stability of the purified Ppo-Er1 proved to be very good, boding well for the enzyme's incorporation in potential enzyme screens: At −20 °C and in the presence of 20% glycerol, the enzyme did not lose any activity even when stored for an extended period of time (one week), whereas an activity drop of approximately 20% was observed after incubation for 10 days at 4 °C (no additives). In contrast to a number of reported OYEs [15,39], we found that NADPH and NADH are equally preferred physiological cofactors of Ppo-ER1 (Figure S14) allowing for maximum flexibility in the choice of recycling system during process development. Both, the coupled-enzyme approach [40] or the use of alternative hydride sources [41,42] will thus be conceivable options to avoid having to add stoichiometric amounts of the coenzymes.

The oligomeric state of Ppo-Er1 was determined via gel filtration by correlation with a commercial gel filtration standard containing proteins of specific size. Based on this comparison, Ppo-Er1 mostly

occurs as a monomer (Figure S2) as do for example PETN from *Enterobacter cloacae* [21] and RmER from *Ralstonia metallidurans* [43], both thermophilic-like ene reductases.

Further relevant parameters for application such as optimum pH, optimum temperature, and long-term temperature stability were determined using the substrate cyclohexenone. The pH profile of Ppo-Er1 was measured in Davies buffer covering pH 5 to pH 10 [44], in which the enzyme reached about 50% of the activity observed in 50 mM phosphate buffer (Figure S3). The pH profile was found to be bell-shaped, exhibiting a narrow optimum at pH 6.5–7.5 (Figure 2). Beyond this range, enzyme activity decreases rapidly, especially when the enzyme was pre-incubated for a longer time period (24 h) in the measurement buffers (Figure 2). In the case of other characterized class III OYEs such as LacER [36] and YqiG [15,34], a similar pH profile was determined albeit with a wider pH working range as indicated by the reported optimum activities in the range of pH_{opt} 8–9 and pH_{opt} 6–9, respectively. Notably, OYE enzymes belonging to other subclasses exhibit similar pH profiles as reported for Ppo-Er1, e.g., the "classical" XenB [45] and NemA [45] with a pH_{opt} of 6–7.5, the "thermophilic-like" YqjM [46] and Chr-OYE3 [47] with a pH_{opt} of 6–8, and the class IV enzyme Ppo-Er3 [15] with a pH_{opt} of 7–8.5.
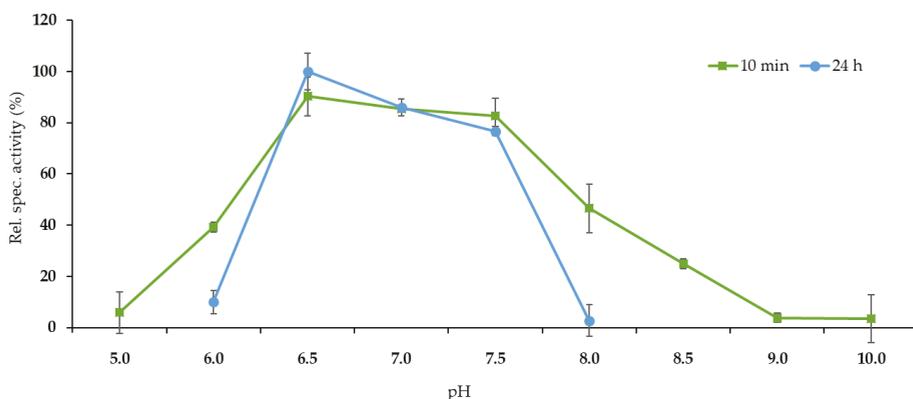
Figure 2. pH profile of Ppo-Er1 measured between pH 5 and pH 10 in Davies buffer [44]. The enzyme was preincubated at 25 °C in the respective measurement buffer solution for 10 min and 24 h, respectively, to determine the stability and activity of Ppo-Er1 in dependence of pH. Relative specific activity corresponds 100% to an activity of 0.41 U/mg for cyclohexenone. The error bars show the standard deviation of triplicates.

In terms of thermal robustness, Ppo-Er1 possesses interesting long-term stability. After 24 h incubation at 20 °C, enzyme activity toward cyclohexenone remained virtually unchanged, whereas residual activity of approximately 70% was detected after an equally long incubation time at 30 °C. Furthermore, short-term exposure of Ppo-Er1 to 45 °C led to only a marginal loss in activity (<10%) allowing the enzyme to be used for applications that require higher temperatures (Figure 3). These results are in line with data obtained for other class III and IV enzymes such as YqiG and Ppo-Er3, which have reported T_{opt} values of 25–40 °C [15,34]. Strikingly, Ppo-Er1 retained a relative specific activity of >70% at temperatures as low as 10 °C making the enzyme an interesting candidate to be used for the transformation of thermolabile substrates such as aldehydes (Figure 3). Overall, our Ppo-Er1 data confirm that the temperature profile of class III enzymes resembles those of their mesophilic counterparts of class I, for example NemA [45] with a reported T_{opt} of 30–50 °C and OYE2p [48] with a T_{opt} of 25–40 °C. Finally, we employed the *Thermo*FAD technique to determine the melting temperature of Ppo-Er1 and found that the ene reductase unfolds at $T_m = 46.5 \pm 1$ °C (Figure S15).

Figure 3. The temperature profile and the temperature stability of Ppo-Er1. For the temperature profile Ppo-Er1 was incubated fo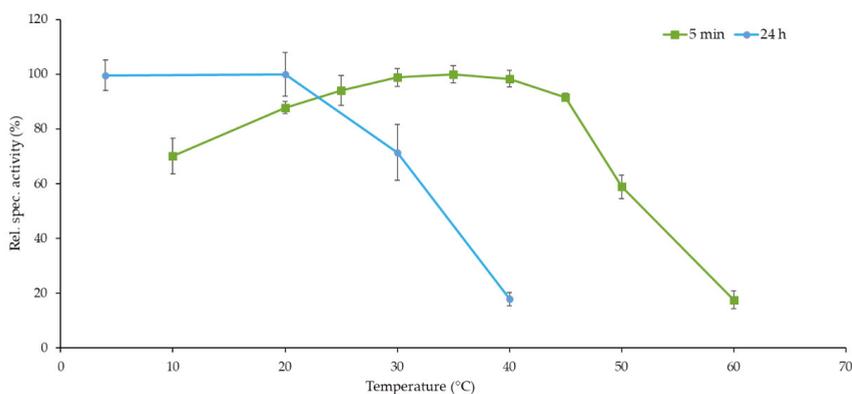r 5 min at different temperatures (10–60 °C) and directly measured for the conversion of substrate cyclohexenone (1 mM). For the temperature stability measurement, Ppo-Er1 was incubated at four different temperatures (4–40 °C) and measured after 24 h at 25 °C. The error bars show the standard deviation of triplicates. Relative specific activity corresponds 100% to an activity of 0.52 U/mg for cyclohexenone.

The use of cosolvents is often a "must" in biocatalytic processes due to the presence of high concentrations of various organic substrates. Consequently, in many instances the solvent stability of enzymes needs to be optimized by enzyme engineering to generate catalysts that are compatible with the process conditions [49]. To verify the stability of Ppo-Er1 in the presence of a set of typical solvents, we thus determined the enzymatic activity over a concentration range of 10–40% of DMSO, DMF, cyclohexane, ethanol, and ethyl acetate. The enzyme performed best in cyclohexane (assayed substrate: 1 mM hexenal), which did not cause a significant loss in activity even when supplemented to a final volume of up to 40% in the assay. Alternatively, DMSO could be considered as a viable cosolvent for Ppo-Er1 as the enzyme was virtually unaffected up to a concentration of 20% *v/v*. Even at a concentration of 30% *v/v* DMSO, Ppo-Er1 retained a relative activity of approximately 80% (assay substrate: 1 mM cyclohexenone). The solvent ethanol was shown to also be a suitable choice for this enzyme, as it was tolerated well up to a concentration of 10% *v/v*. DMF or ethyl acetate, however, should not be used in combination with Ppo-Er1 as their presence was found to be detrimental for enzymatic activity. Already at a concentration of 10% *v/v* activity drops of 30% and 85% were observed, respectively (Figure 4).

In comparison to most known old yellow enzymes, Ppo-Er1 exhibits similar solvent resistance: The thermophilic-like OYE YqjM [46] has been reported to remain active in an analogous concentration range of DMSO, DMF, and ethyl acetate as Ppo-Er1. However, an ethanol concentration of 10% *v/v* led to a strong reduction of the half-life of YqjM, which we did not observe in the case of Ppo-Er1. TOYE [23], another thermophilic-like OYE, was reported to exhibit a 50% loss of activity at an ethanol concentration of 45% corresponding to a higher stability toward this solvent compared to Ppo-Er1, whereas the classical PETNR [50] already lost 50% activity in the presence of an ethanol concentration of 20% *v/v*. In this context, it should be noted that organic-solvent-tolerant ene reductases have also been reported: FOYE1, originating from an acidophilic iron oxidizer, was shown to perform well in many solvent systems with up to 20% *v/v* solvents (ethanol, methanol, acetone, isopropanol, DMSO, THF) clearly outperforming all abovementioned ene reductases in terms of solvent stability [51].
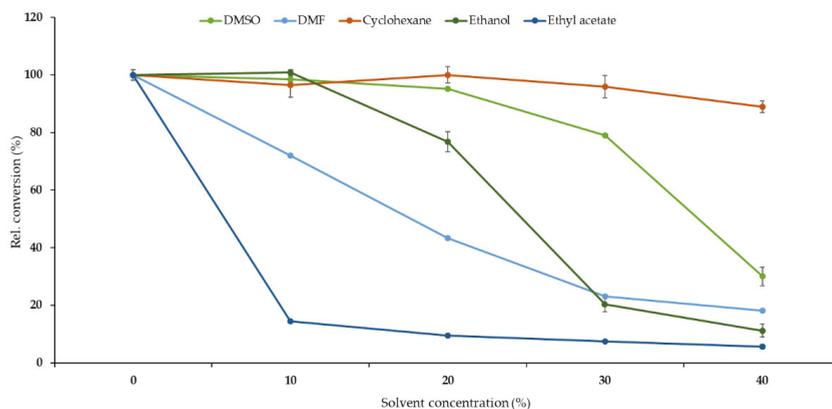
Figure 4. Overview of the solvent stability of Ppo-Er1 in DMSO (dimethyl sulfoxide), DMF (dimethyl formamide), cyclohexane, ethanol, and ethyl acetate in a concentration range of 10%–40% *v/v*. The standard enzyme assay was performed while the concentration of solvents was varied (substrate for cyclohexane: 1 mM hexenal, all other solvents: 1 mM cyclohexenone). Data are shown as values relative to an enzyme assay without cosolvent in which 100% relative conversion corresponds to the production of 0.84 mM cyclohexanone or 0.49 mM hexanal, respectively. The error bars show the standard deviation of triplicates, except for the 30% *v/v* cyclohexane point for which only two measurements were available.

2.2. Substrate Scope, Determination of Michaelis–Menten Parameters, and Stereoselectivity

To determine the substrate profile of Ppo-Er1, the enzyme was tested for the conversion of nineteen structurally diverse aliphatic and cyclic alkenes bearing ketone, aldehyde, nitro, carboxylic acid, or ester moieties as electron-withdrawing groups. For thirteen substrates, product formation by Ppo-Er1 could be detected. Cyclohexenone, hexenal, 2-methyl-2-pentenal, 4-phenyl-3-buten-2-one, cinnamic aldehyde, maleimide, and carvone (at 5 mM concentration) were converted especially well, and >99% conversion was obtained within 4 h (Table 1). Substrates not accepted by Ppo-Er1 included α,β-unsaturated carboxylic acids such as butenic acid, cinnamic acid, and citraconic acid as well as the ketones 3-methyl-2-cyclohexenone and 3-methyl-2-cyclopentenone, which are characterized by an additional methyl group in the β-position. The α,β-unsaturated ester ethyl crotonate was also not converted.

Based on the obtained data, it can be concluded that the overall substrate profile of Ppo-Er1 resembles that of other subclass III enzymes such as YqiG [15,34] and Lla-Er [15]. For example, 5 mM of cinnamic aldehyde and cyclohexenone are also well converted by Lla-Er [15] (65% ± 4.2% and 23% ± 3.1%) and YqiG [15] (58% ± 2.4% and 55% ± 6.1%) after 1 h at 30 °C. Notably, however, marked differences in substrate acceptance by class III enzymes occur for some of the investigated substrates highlighting the importance of an in-depth substrate profiling: Whereas carvone and maleimide are very well converted by Ppo-Er1 (both: >99%), Lla-Er, for example, accepts this compound only poorly (carvone: 2.6% ± 0.1%, maleimide: not converted) [15]. Diethylbenzylidenemalonate conversion by YqiG [15,34] (11% ± 1.3%), on the other hand, significantly exceeded the detected product formations achieved by Lla-Er (<1%) [15] and Ppo-Er1 (1.2%). Moreover, 3-methyl-2-cyclopentenone, which is not converted by Ppo-Er1, Lla-Er [15], and YqiG [15,34], has been shown to be accepted by LacER [36]. Generally, we noted that Ppo-Er1 has a restricted substrate acceptance for cyclic β-methylated substrates such as 3-methyl-2-cyclohexenone and 3-methyl-2-cyclopentenone, which possibly results from a difficulty in accepting substituents at the C_β position of cyclic compounds in the active site in analogy to other class II, III, and IV enzymes [15,39]. In addition, carboxylic acids

and esters seem to be non-optimal alkene activating groups for this enzyme as conversion of the corresponding substrates was low or not detectable.

Table 1. Conversion, steady state kinetics,[a] and enantiomeric excess (ee) of various substrates converted with purified enzymes as determined after 4 h at 20 °C (n.d.: not detected; n.s.: not soluble). The given uncertainties show the standard deviation of triplicates.

Substrate		Conversion	ee	k_{cat}/K_m	K_m	k_{cat}
Name	Structure	(%)	(%)	(mM^{-1} s^{-1})	(mM)	(s^{-1})
Maleimide		≥99 ± 3.7		287.8 ± 0.12	0.10 ± 0.01	28.78 ± 0.62
trans-β-Methyl-β-nitrostyrene		81 ± 1.0		41.4 ± 0.23	0.12 ± 0.03	4.97 ± 0.36
2-Methyl-2-pentenal		≥99 ± 7.4	(S) 63	15.3 ± 0.09	0.41 ± 0.04	6.27 ± 0.11
Cinnamaldehyde		≥99 ± 1.5		14.6 ± 0.14	0.36 ± 0.05	5.27 ± 0.18
Hexenal		≥99 ± 3.4		3.3 ± 0.10	2.22 ± 0.21	7.42 ± 0.01
Carvone		≥99 ± 2.1	(R) 98	0.5 ± 0.16	4.35 ± 0.69	2.20 ± 0.08
Cyclohexenone		≥99 ± 0.5		0.4 ± 0.08	13.42 ± 1.0	5.25 ± 0.1
Citral		29 ± 1.4	(S) 94	0.2 ± 0.93	1.12 ± 1.0	0.17 ± 0.04
2-Methyl-2-cyclohexenone		76.2 ± 0.4	(R) 92	0.1 ± 0.23	14.93 ± 3.3	1.30 ± 0.08
Cyclopentenone		59 ± 1.7		0.03 ± 0.17	57.24 ± 9.4	1.75 ± 0.16
4-Phenyl-3-buten-2-one		≥99 ± 1.0			n.s.	
Butylacrylate		22 ± 6.5				
Diethyl benzyldienemalonate		1.2 ± 0.0				
3-Methyl-2-cyclohexenone		n.d.				
3-Methyl-2-cyclopentenone		n.d.				
Etylcrotonate		n.d.				
Butenoic acid		n.d.				
Cinnamic acid		n.d.				
Citraconic acid		n.d.				

[a] Reactions (1 mL) were performed in potassium phosphate buffer (50 mM, pH 7.0) containing NADPH (175 μM) and substrate (20 μM–80 mM), depending on substrate, Ppo-Er1 (0.61 μM), and DMSO to solubilize the substrates. The reactions were followed continuously by monitoring NADPH oxidation at 340 nm for 90 sec at 25 °C.

To complement the substrate acceptance profile, Michaelis–Menten parameters of Ppo-Er1 for ten diverse substrates were determined (Table 1, Figures S4–S13). Within the tested substrate range, Ppo-Er1 showed the highest catalytic efficiency for maleimide (k_{cat}/K_m = 287 mM^{-1} s^{-1}) followed by trans-β-methyl-β-nitrostyrene (k_{cat}/K_m = 41 mM^{-1} s^{-1}). In combination with the conversion data, the

measured kinetic parameters (Table 1) indicate a general preference for alkenes carrying a phenyl substituent at the C_β position of the substrates. Overall, Ppo-Er1's specific activity for other typical ene reductase substrates such as carvone (k_{cat}/K_m = 0.5 mM^{-1} s^{-1}) and cyclohexanone (k_{cat}/K_m = 0.4 mM^{-1} s^{-1}) was found to be in a similar range as those described for other well-known OYEs such as the classical PETNR (carvone: k_{cat}/K_m = 2 mM^{-1} s^{-1}; cyclohexanone: k_{cat}/K_m = 5 mM^{-1} s^{-1}) [50] and the thermophilic-like YqjM (cyclohexanone: k_{cat}/K_m = 6.4 mM^{-1} s^{-1}) [46] (Table 2). Maleimide, however, is better converted by ene reductases from photosynthetic extremophiles such as CtOYE (k_{cat}/K_m = 1940 mM^{-1} s^{-1}) or GsOYE (k_{cat}/K_m = 399 mM^{-1} s^{-1}) [52] the thermophilic-like OYERo2 (k_{cat}/K_m = 10,800 mM^{-1} s^{-1}) [53] or the class III OYE YqiG (k_{cat}/K_m = 800 mM^{-1} s^{-1}) (Table 2).

Table 2. Comparison of the catalytic efficiencies (mM^{-1} s^{-1}) of a range of known old yellow enzymes (OYEs) (YqiG [34], PETNR [50], YqjM [46], TOYE [23], DrER [43], RmER [43], and OYERo2 [53]) from class I–III.

Substrate	Class I	Class II					Class III	
	PETNR	YqjM	TOYE	DrER	RmER	OYERo2	Ppo-Er1	YqiG
Cyclohexenone	5	6.4	0.5	2.1	0.7		0.4	22
2-Methyl-cyclohexenone	4	1.0					0.1	
Cyclopentenone	<0.5	1.9	0.6				0.03	
Hexenal		0.60					3.3	
Citral	9	0.02	0.05				0.2	6.7
2-Methyl-2-pentenal	61		0.14				15.3	18
Cinnamaldehyde	8						14.6	
Carvone	2		1.5				0.5	7.5
Maleimide						10,800	287.8	800
trans-β-Methyl-β-nitrostyrene							41.4	

In addition to determining the steady-state kinetic parameters, we also investigated the stereopreference of Ppo-Er1. Based on our results with four selected substrates, Ppo-Er1 displays a similar stereopreference to other reported OYE class III enzymes (Table 3), preferentially forming the *S*-product when converting 2-methy-2-pentenal and citral and forming the *R*-product when transforming carvone and 2-methyl-2-cyclohexenone. Notably, the detected ee values of Ppo-Er1 are generally superior to values determined for YqiG and Lla-Er [15] with the only exception being the enantiomeric excess reported for the conversion of carvone by Lla-Er (>99.9% ee). It should be noted, however, that Lla-Er displayed a low conversion of 2.6% of 5 mM substrate after 1 h at 30 °C compared to the >99% conversion of 5 mM substrate by Ppo-Er1 after 4 h at 20 °C.

Table 3. The enantiomeric excess of some selected OYEs (YqiG [15], Lla-Er [15], Ppo-Er3 [15], OPR1 [54], OPR3 [54], PETNR [50], YqjM [54], TOYE [23]) from classes I–IV. The values presented for YqjM were measured as a reference for Ppo-Er1 and compared with the literature [54].

Substrate	Class I			Class II		Class III			Class IV
	OPR1	OPR3	PETNR	YqjM	TOYE	Ppo-Er1	YqiG	Lla-ER	Ppo-Er3
2-Methyl-2-pentenal	(R) 47	(S) 78		(R) 20	(S) 55	(S) 63	(S) 33	(S) 5	(S) 67
Carvone			(R) 95	(R) 82	(R) 95	(R) 98	(R) 89	(R) >99.9	(R) 91
2-Methyl-2-cyclohexenone	(R) 77	(R) 62		(R) 81		(R) 92	(R) 83	(R) 11	(R) 86
Citral	(S) >95	(S) >95		(S) 95	(S) 91	(S) 94			

3. Materials and Methods

3.1. Materials

All chemicals were purchased from Merck (Darmstadt, Germany), VWR (Hannover, Germany), or Carl Roth (Karlsruhe, Germany). The purchased chemicals were of the highest available purity or of analytical grade and were used without further purification unless otherwise specified. NADPH tetrasodium salt was ordered from Oriental Yeast Co. Ltd. (Tokyo, Japan). The plasmid (pET 28b(+)

incl. Ppo-Er1) was ordered from Twist Bioscience (San Francisco, CA, USA). The HisTrap FF and the HiTrap Desalting columns were ordered from GE Healthcare (Uppsala, Sweden).

3.2. Plasmid

Twist Bioscience (San Franscisco, CA, USA) cloned the synthetic gene of the codon optimized Ppo-Er1 (Accession Nr: WP_013369181) with *Nde*I and *Xho*I in the commercial pET28b(+) vector.

3.3. Bacterial Strains and Culture Conditions

E. coli BL21 (DE3) [*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*] was purchased from New England Biolabs (Beverly, MA, USA). *E. coli* strains were cultured routinely in Lysogeny broth (LB) or TB media and were supplemented with kanamycin (50 μg mL^{-1}). Bacterial cultures were incubated in baffled Erlenmeyer flasks in a New Brunswick Innova 42 orbital shaker at 200 rpm and 37 °C. Bacteria on agar plates were incubated in a HERATherm Thermo Scientific incubator under air. All materials and biotransformation media were sterilized by autoclaving at 121 °C for 20 min. Aqueous stock solutions were sterilized by filtration through 0.22 μm syringe filters. Agar plates were prepared with LB medium supplemented by 1.5% (w/v) agar.

3.4. Expression

The expression of Ppo-Er1 in *E. coli* BL21 (DE3) was performed by inoculation of TB media (400 mL) supplemented with kanamycin (50 μg mL^{-1}) with an overnight culture (4 mL; 1:100). The culture was incubated at 37 °C and 180 rpm until optical density OD$_{600}$ = 0.5–0.8 was reached. Afterward expression was induced by the addition of 100 μM IPTG, and incubation was continued at 25 °C for 18 h. Cells were harvested by centrifugation at 4500× *g* for 10 min at 4 °C and either used directly or the pellet was stored by freezing at −20 °C.

3.5. Enzyme Purification

The cell disruption was performed by resuspending the pellet from a 400 mL culture in 20 mL buffer (100 mM sodium phosphate buffer pH 7.5, 300 mM NaCl, supplemented by 30 mM imidazole) and a single passage through a French press (2000 psi). The crude extract was separated from the cell debris by centrifugation at 8000× *g* for 45 min. Purification was achieved by affinity chromatography exploiting the C-terminal His-Tag using an automated Äkta purifier system. The crude extract was filtered (0.45 μm) and applied to a pre-equilibrated 5 mL HisTrap FF column. The unbound protein was washed with five column volumes of buffer supplemented with 45 mM imidazole. The elution of Ppo-Er1 was accomplished by a three-column volume of buffer supplemented with 300 mM of imidazole. The resulting fractions were collected and analyzed by SDS-PAGE. The fractions with a high content of Ppo-Er1 were pooled and desalted using 50 mM sodium phosphate buffer (pH 7.5) to remove the imidazole. This step was performed employing the Äkta purifier system using three coupled 5 mL HiTrap desalting columns. After the system was equilibrated, the Ppo-Er1-containing sample was applied and fractioned. The protein fractions were analyzed via the integrated online absorption measurement at 280 nm. The protein content of the pooled purified sample was determined by measuring the adsorption with a NanoDrop One (Thermo Fisher Scientific) system and using the molecular weight (41.3 kDa) and extinction coefficient ($\epsilon_{\lambda = 280\,nm} = 38'390$ M^{-1} cm^{-1}) of Ppo-Er1 for the calculation. The extinction coefficient was obtained by using the online calculation tool Prot pi [55].

3.6. Activity Assay

The activity measurements were recorded spectrophotometrically by observing NADPH consumption at 340 nm for 60–90 s in a 1 mL (1 cm) plastic cuvette in the Lambda 465 (PDA UV/VIS) system from Perkin Elmer. The biocatalytic experiments to obtain the pH and the temperature profile were conducted in sodium phosphate buffer (50 mM, pH 7.5) using 175 μM NADPH, 1 mM

cyclohexenone, and 0.61 µM purified Ppo-Er1. For the determination of the Michaelis–Menten parameters, the substrate concentration was varied in the range of 20 µM–80 mM depending on the substrate while the enzyme concentration was kept constant at 0.61 µM. For the pH profile, Davies buffer [44] was used. All measurements were done in triplicates. Background NADPH consumption was determined in assays in which either the enzyme or the substrate had been eliminated. The substrates were solubilized as 1 M stock in DMSO.

3.7. Biocatalysis Reaction

The in in vitro biocatalysis reaction were performed by using desalted Ppo-Er1 (with a concentration of 12.1 µM), 5 mM substrate (1 M stock in DMSO) supplemented with 100 µM NADPH, 10 mM glucose, and 5 µL GDH (20% w/v cell suspension). The reaction volume was adjusted to 1 mL in a glass vial by using sodium phosphate buffer (200 mM, pH 7.0) and incubated for 4 h at 20 °C and 1000 rpm. To determine the solvent stability of Ppo-Er1, the biocatalysis reaction conditions were adapted to include 2.4 µM Ppo-Er1 and 0%–40% v/v solvent (ethanol, ethyl acetate, DMSO, DMF, cyclohexane) in a total reaction volume of 1 mL for 50 min at 20 °C and 1000 rpm. All biocatalysis reactions were done in triplicate, biocatalysis results were verified by control reactions omitting the enzyme.

3.8. GC-Analysis

One milliliter biocatalysis reactions were extracted once with 500 µL methyl *tert*-butyl ether (incl. 1 g/L 1-octanol as internal standard). The phase separation was achieved by centrifugation of the biphasic sample, and the organic phase was separated and subjected to GC analysis (Table S1).

3.9. Gel Filtration

For the determination of the oligomeric state of Ppo-Er1, the Äkta purifier system employing a HiLoad 16/600 Superdex 75 pg column (GE Healthcare (Uppsala, Sweden)) and sodium phosphate buffer (50 mM, pH 7.5) was used. In a first step, the system was calibrated by using the gel filtration standard from Bio Rad (1.35–670 kDa Prod. no.: #1511901). Then flavin-saturated Ppo-Er1 was applied to system under identical conditions.

3.10. Melting Temperature

The unfolding temperature was determined by a *Thermo*FAD assay [56] using Rotor-Gene Q RT-PCR machine. Protein samples (0.5–0.3 mg/mL) in 20 µL sodium phosphate buffer pH 7 were measured using a temperature gradient from 25 to 90 °C, performing fluorescence measurements every 0.5 °C increase after a 10 s delay for signal stabilization. The measurements were performed in triplicates using 470 nm excitation wavelength and 510 nm emission wavelength.

4. Conclusions

Ppo-Er1 is a well-expressed, easy to purify, old yellow enzyme belonging to the recently introduced subclass III designation. In terms of cofactor preference, the enzyme accepts NADPH and NADH equally well, whereas pH and optimum temperature resemble those of previously described OYEs. Notably, the enzyme exhibits only slightly reduced performance (>70% conversion of 1 mM cyclohexenone) at lowered temperatures (10 °C) making it a possible candidate for the transformation of labile substrates such as some aldehydes. In addition, the enzyme was shown to have noteworthy stability in the presence of the solvents cyclohexane (up to at least 40% v/v), DMSO, and ethanol (up to 20% v/v).

The substrate profile analysis with a set of 19 representative alkenes allowed the establishment of Ppo-Er1's substrate scope highlighting its acceptance of a variety of linear and cyclic compounds with often excellent transformation efficiencies and exquisite stereoselectivity (e.g., 98% ee for carvone). Complementing this analysis with the determination of steady-state kinetics for ten of the substrates allowed us to conclude that Ppo-Er1 classifies well with other subgroup III old yellow enzymes.

In summary, our in-depth characterization of Ppo-Er1 allows the enlargement of the available panel of ene reductases with a versatile biocatalyst having interesting synthetic properties. Its introduction in the biocatalytic toolbox may further facilitate academic and industrial efforts when screening for biocatalysts capable of asymmetric double bond reduction. Looking forward, Ppo-Er1's performance could be further optimized via enzyme and process engineering.

Author Contributions: Conceptualization: C.P. and R.M.B.; experimental work: D.A.; writing: D.A., C.P., and R.M.B. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. List, B.; Yang, J.W. Chemistry. The organic approach to asymmetric catalysis. *Science* **2006**, *313*, 1584–1586. [CrossRef] [PubMed]

2. Huffman, M.A.; Fryszkowska, A.; Alvizo, O.; Borra-Garske, M.; Campos, K.R.; Canada, K.A.; Devine, P.N.; Duan, D.; Forstater, J.H.; Grosser, S.T.; et al. Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **2019**, *366*, 1255–1259. [CrossRef] [PubMed]

3. Savile, C.K.; Janey, J.M.; Mundorff, E.C.; Moore, J.C.; Tam, S.; Jarvis, W.R.; Colbeck, J.C.; Krebber, A.; Fleitz, F.J.; Brands, J.; et al. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **2010**, *329*, 305–309. [CrossRef] [PubMed]

4. Schober, M.; MacDermaid, C.; Ollis, A.A.; Chang, S.; Khan, D.; Hosford, J.; Latham, J.; Ihnken, L.A.F.; Brown, M.J.B.; Fuerst, D.; et al. Chiral synthesis of lsd1 inhibitor gsk2879552 enabled by directed evolution of an imine reductase. *Nat. Catal.* **2019**, *2*, 909–915. [CrossRef]

5. Adams, J.P.; Brown, M.J.B.; Diaz-Rodriguez, A.; Lloyd, R.C.; Roiban, G.D. Biocatalysis: A pharma perspective. *Adv. Synth. Catal.* **2019**, *361*, 2421–2432. [CrossRef]

6. Toogood, H.S.; Scrutton, N.S. Discovery, characterisation, engineering and applications of ene reductases for industrial biocatalysis. *ACS Catal.* **2019**, *8*, 3532–3549. [CrossRef]

7. Toogood, H.S.; Gardiner, J.M.; Scrutton, N.S. Biocatalytic reductions and chemical versatility of the old yellow enzyme family of flavoprotein oxidoreductases. *Chemcatchem* **2010**, *2*, 892–914. [CrossRef]

8. Winkler, C.K.; Faber, K.; Hall, M. Biocatalytic reduction of activated cc-bonds and beyond: Emerging trends. *Curr. Opin. Chem. Biol.* **2018**, *43*, 97–105. [CrossRef]

9. Shimoda, K.; Ito, D.I.; Izumi, S.; Hirata, T. Novel reductase participation in the syn-addition of hydrogen to the c=c bond of enones in the cultured cells of *Nicotiana tabacum*. *J. Chem. Soc. Perkin Trans.* **1996**, *1*, 355–358. [CrossRef]

10. Knowles, W.S. Asymmetric hydrogenations (nobel lecture). *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 1999–2007.

11. Noyori, R. Asymmetric catalysis: Science and opportunities (nobel lecture). *Angew. Chem. Int. Ed.* **2002**, *41*, 2008–2022. [CrossRef]

12. Knaus, T.; Toogood, H.S.; Scrutton, N.S. Ene-reductases and their applications. In *Green Biocatalysis*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2016; pp. 473–488.

13. Steinkellner, G.; Gruber, C.C.; Pavkov-Keller, T.; Binter, A.; Steiner, K.; Winkler, C.; Lyskowski, A.; Schwamberger, O.; Oberer, M.; Schwab, H.; et al. Identification of promiscuous ene-reductase activity by mining structural databases using active site constellations. *Nat. Commun.* **2014**, *5*, 4150. [CrossRef] [PubMed]

14. Hecht, K.; Buller, R. Ene-reductases in pharmaceutical chemistry. In *Pharmaceutical Biocatalysis: Chemoenzymatic Synthesis of Active Pharmaceutical Ingredients*; Grunwald, P., Ed.; Jenny Stanford Publishing: Singapore, 2019.

15. Peters, C.; Frasson, D.; Sievers, M.; Buller, R. Novel old yellow enzyme subclasses. *Chembiochem* **2019**, *20*, 1569–1577. [CrossRef] [PubMed]

16. Warburg, O.; Christian, W. Ein zweites sauerstoffübertragendes ferment und sein absorptionsspektrum. *Naturwissenschaften* **1932**, *20*, 688. [CrossRef]

17. Haas, E. Isolierung eines neuen gelben ferments. *Biochem. Z* **1938**, *298*, 369–390.

18. Breithaupt, C.; Strassner, J.; Breitinger, U.; Huber, R.; Macheroux, P.; Schaller, A.; Clausen, T. X-ray structure of 12-oxophytodienoate reductase 1 provides structural insight into substrate binding and specificity within the family of oye. *Structure* **2001**, *9*, 419–429. [CrossRef]

19. Kohli, R.M.; Massey, V. The oxidative half-reaction of old yellow enzyme: The role of tyrosine 196. *J. Biol. Chem.* **1998**, *273*, 32763–32770. [CrossRef]

20. Fox, K.M.; Karplus, P.A. Old yellow enzyme at 2 å resolution: Overall structure, ligand binding, and comparison with related flavoproteins. *Structure* **1994**, *2*, 1089–1105. [CrossRef]

21. Barna, T.M.; Khan, H.; Bruce, N.C.; Barsukov, I.; Scrutton, N.S.; Moody, P.C.E. Crystal structure of pentaerythritol tetranitrate reductase: "Flipped" binding geometries for steroid substrates in different redox states of the enzyme. *J. Mol. Biol.* **2001**, *310*, 433–447. [CrossRef]

22. Kitzing, K.; Fitzpatrick, T.B.; Wilken, C.; Sawa, J.; Bourenkov, G.P.; Macheroux, P.; Clausen, T. The 1.3 å crystal structure of the flavoprotein yqjm reveals a novel class of old yellow enzymes. *J. Biol. Chem.* **2005**, *280*, 27904–27913. [CrossRef]

23. Adalbjörnsson, B.V.; Toogood, H.S.; Fryszkowska, A.; Pudney, C.R.; Jowitt, T.A.; Leys, D.; Scrutton, N.S. Biocatalysis with thermostable enzymes: Structure and properties of a thermophilic 'ene'-reductase related to old yellow enzyme. *ChemBioChem* **2010**, *11*, 197–207. [CrossRef] [PubMed]

24. Nizam, S.; Verma, S.; Borah, N.N.; Gazara, R.K.; Verma, P.K. Comprehensive genome-wide analysis reveals different classes of enigmatic old yellow enzyme in fungi. *Sci. Rep.* **2014**, *4*, 4013. [CrossRef] [PubMed]

25. Pietruszka, J.; Schölzel, M. Ene reductase-catalysed synthesis of (r)-profen derivatives. *Adv. Synth. Catal.* **2012**, *354*, 751–756. [CrossRef]

26. Li, Z.N.; Wang, Z.X.; Meng, G.; Lu, H.; Huang, Z.D.; Chen, F.E. Identification of an ene reductase from yeast *Kluyveromyces marxianus* and application in the asymmetric synthesis of (R)-profen esters. *Asian. J. Org. Chem.* **2018**, *7*, 763–769. [CrossRef]

27. Waller, J.; Toogood, H.S.; Karuppiah, V.; Rattray, N.J.W.; Mansell, D.J.; Leys, D.; Gardiner, J.M.; Fryszkowska, A.; Ahmed, S.T.; Bandichhor, R.; et al. Structural insights into the ene-reductase synthesis of profens. *Org. Biomol. Chem.* **2017**, *15*, 4440–4448. [CrossRef] [PubMed]

28. Fryszkowska, A.; Fisher, K.; Gardiner, J.M.; Stephens, G.M. A short, chemoenzymatic route to chiral β-aryl-γ-amino acids using reductases from anaerobic bacteria. *Org. Biomol. Chem.* **2010**, *8*, 533–535. [CrossRef] [PubMed]

29. Winkler, C.K.; Clay, D.; Davies, S.; O'Neill, P.; McDaid, P.; Debarge, S.; Steflik, J.; Karmilowicz, M.; Wong, J.W.; Faber, K. Chemoenzymatic asymmetric synthesis of pregabalin precursors via asymmetric bioreduction of β-cyanoacrylate esters using ene-reductases. *J. Org. Chem.* **2013**, *78*, 1525–1533. [CrossRef]

30. Winkler, C.K.; Clay, D.; Turrini, N.G.; Lechner, H.; Kroutil, W.; Davies, S.; Debarge, S.; O'Neill, P.; Steflik, J.; Karmilowicz, M.; et al. Nitrile as activating group in the asymmetric bioreduction of beta-cyanoacrylic acids catalyzed by ene-reductases. *Adv. Synth. Catal.* **2014**, *356*, 1878–1882. [CrossRef]

31. Janicki, I.; Kiełbasiński, P.; Turrini, N.G.; Faber, K.; Hall, M. Asymmetric bioreduction of β-activated vinylphosphonate derivatives using ene-reductases. *Adv. Synth. Catal.* **2017**, *359*, 4190–4196. [CrossRef]

32. Bertolotti, M.; Brenna, E.; Crotti, M.; Gatti, F.G.; Monti, D.; Parmeggiani, F.; Santangelo, S. Substrate scope evaluation of the enantioselective reduction of β-alkyl-β-arylnitroalkenes by old yellow enzymes 1–3 for organic synthesis applications. *ChemCatChem* **2016**, *8*, 577–583. [CrossRef]

33. Dobrijevic, D.; Benhamou, L.; Aliev, A.E.; Méndez-Sánchez, D.; Dawson, N.; Baud, D.; Tappertzhofen, N.; Moody, T.S.; Orengo, C.A.; Hailes, H.C.; et al. Metagenomic ene-reductases for the bioreduction of sterically challenging enones. *RSC Adv.* **2019**, *9*, 36608–36614. [CrossRef]

34. Sheng, X.; Yan, M.; Xu, L.; Wei, M. Identification and characterization of a novel old yellow enzyme from *Bacillus subtilis* str.168. *J. Mol. Catal. B Enzym.* **2016**, *130*, 18–24. [CrossRef]

35. Zhang, H.; Gao, X.; Ren, J.; Feng, J.; Zhang, T.; Wu, Q.; Zhu, D. Enzymatic hydrogenation of diverse activated alkenes. Identification of two bacillus old yellow enzymes with broad substrate profiles. *J. Mol. Catal. B Enzym.* **2014**, *105*, 118–125. [CrossRef]

36. Gao, X.; Ren, J.; Wu, Q.; Zhu, D. Biochemical characterization and substrate profiling of a new nadh-dependent enoate reductase from *Lactobacillus casei*. *Enzyme Microb. Technol.* **2012**, *51*, 26–34. [CrossRef] [PubMed]

37. Brown, B.J.; Deng, Z.; Karplus, P.A.; Massey, V. On the active site of old yellow enzyme: Role of histidine 191 and asparagine 194. *J. Biol. Chem.* **1998**, *273*, 32753–32762. [CrossRef]

38. Spiegelhauer, O.; Dickert, F.; Mende, S.; Niks, D.; Hille, R.; Ullmann, M.; Dobbek, H. Kinetic characterization of xenobiotic reductase a from *Pseudomonas putida* 86. *Biochemistry* **2009**, *48*, 11412–11420. [CrossRef]

39. Scholtissek, A.; Tischler, D.; Westphal, A.; van Berkel, W.; Paul, C. Old yellow enzyme-catalysed asymmetric hydrogenation: Linking family roots with improved catalysis. *Catalysts* **2017**, *7*, 130. [CrossRef]

40. Hummel, W.; Gröger, H. Strategies for regeneration of nicotinamide coenzymes emphasizing self-sufficient closed-loop recycling systems. *J. Biotechnol.* **2014**, *191*, 22–31. [CrossRef]

41. Knaus, T.; Paul, C.E.; Levy, C.W.; de Vries, S.; Mutti, F.G.; Hollmann, F.; Scrutton, N.S. Better than nature: Nicotinamide biomimetics that outperform natural coenzymes. *J. Am. Chem. Soc.* **2016**, *138*, 1033–1039. [CrossRef]

42. Lee, S.H.; Choi, D.S.; Pesic, M.; Lee, Y.W.; Paul, C.E.; Hollmann, F.; Park, C.B. Cofactor-free, direct photoactivation of enoate reductases for the asymmetric reduction of c=c bonds. *Angew. Chem. Int. Ed.* **2017**, *56*, 8681–8685. [CrossRef]

43. Litthauer, S.; Gargiulo, S.; van Heerden, E.; Hollmann, F.; Opperman, D.J. Heterologous expression and characterization of the ene-reductases from *Deinococcus radiodurans* and *Ralstonia metallidurans*. *J. Mol. Catal. B Enzym.* **2014**, *99*, 89–95. [CrossRef]

44. Davies, M.T. A universal buffer solution for use in ultra-violet spectrophotometry. *Analyst* **1959**, *84*, 248–251. [CrossRef]

45. Peters, C.; Kölzsch, R.; Kadow, M.; Skalden, L.; Rudroff, F.; Mihovilovic, M.D.; Bornscheuer, U.T. Identification, characterization, and application of three enoate reductases from *Pseudomonas putida* in in vitro enzyme cascade reactions. *ChemCatChem* **2014**, *6*, 1021–1027. [CrossRef]

46. Pesic, M.; Fernández-Fueyo, E.; Hollmann, F. Characterization of the old yellow enzyme homolog from *Bacillus subtilis* (yqjm). *ChemistrySelect* **2017**, *2*, 3866–3871. [CrossRef]

47. Xu, M.-Y.; Pei, X.-Q.; Wu, Z.-L. Identification and characterization of a novel "thermophilic-like" old yellow enzyme from the genome of *Chryseobacterium* sp. Ca49. *J. Mol. Catal. B Enzym.* **2014**, *108*, 64–71. [CrossRef]

48. Zheng, L.; Lin, J.; Zhang, B.; Kuang, Y.; Wei, D. Identification of a yeast old yellow enzyme for highly enantioselective reduction of citral isomers to (R)-citronellal. *Bioresour. Bioprocess.* **2018**, *5*, 1–12. [CrossRef]

49. Rudroff, F.; Mihovilovic, M.D.; Gröger, H.; Snajdrova, R.; Iding, H.; Bornscheuer, U.T. Opportunities and challenges for combining chemo- and biocatalysis. *Nat. Catal.* **2018**, *1*, 12–22. [CrossRef]

50. Fryszkowska, A.; Toogood, H.; Sakuma, M.; Gardiner, J.M.; Stephens, G.M.; Scrutton, N.S. Asymmetric reduction of activated alkenes by pentaerythritol tetranitrate reductase: Specificity and control of stereochemical outcome by reaction optimisation. *Adv. Synth. Catal.* **2009**, *351*, 2976–2990. [CrossRef]

51. Tischler, D.; Gadke, E.; Eggerichs, D.; Gomez Baraibar, A.; Mugge, C.; Scholtissek, A.; Paul, C.E. Asymmetric reduction of (r)-carvone through a thermostable and organic-solvent-tolerant ene-reductase. *Chembiochem* **2019**. [CrossRef]

52. Robescu, M.S.; Niero, M.; Hall, M.; Cendron, L.; Bergantino, E. Two new ene-reductases from photosynthetic extremophiles enlarge the panel of old yellow enzymes: Ctoye and gsoye. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 2051–2066. [CrossRef]

53. Riedel, A.; Mehnert, M.; Paul, C.E.; Westphal, A.H.; van Berkel, W.J.; Tischler, D. Functional characterization and stability improvement of a 'thermophilic-like' ene-reductase from *Rhodococcus opacus* 1cp. *Front. Microbiol.* **2015**, *6*, 1073. [CrossRef] [PubMed]

54. Hall, M.; Stueckler, C.; Ehammer, H.; Pointner, E.; Oberdorfer, G.; Gruber, K.; Hauer, B.; Stuermer, R.; Kroutil, W.; Macheroux, P.; et al. Asymmetric bioreduction of c=c bonds using enoate reductases opr1, opr3 and yqjm: Enzyme-based stereocontrol. *Adv. Synth. Catal.* **2008**, *350*, 411–418. [CrossRef]

55. Josuran, R. Prot pi. Available online: https://www.protpi.ch/ (accessed on 22 January 2020).
56. Forneris, F.; Orru, R.; Bonivento, D.; Chiarelli, L.R.; Mattevi, A. Thermofad, a thermofluor-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS J.* **2009**, *276*, 2833–2840. [CrossRef] [PubMed]

Cofactor F$_{420}$-Dependent Enzymes: An Under-Explored Resource for Asymmetric Redox Biocatalysis

Mihir V. Shah [1,†]**, James Antoney** [1,2,†]**, Suk Woo Kang** [1,2]**, Andrew C. Warden** [1]**, Carol J. Hartley** [1]**, Hadi Nazem-Bokaee** [1]**, Colin J. Jackson** [1,2] **and Colin Scott** [1,*]

[1] CSIRO Synthetic Biology Future Science Platform, Canberra 2601, Australia; Mihir.Shah@csiro.au (M.V.S.); James.Antoney@anu.edu.au (J.A.); Suk.Kang@anu.edu.au (S.W.K.); andrew.warden@csiro.au (A.C.W.); Carol.Hartley@csiro.au (C.J.H.); Hadi.Nazem-Bokaee@csiro.au (H.N.-B.); colin.jackson@anu.edu.au (C.J.J.)

[2] Research School of Chemistry, Australian National University, Canberra 2601, Australia

* Correspondence: colin.scott@csiro.au; Tel.: +61-2-6246-4090

† M.V.S. and J.A. contributed equally to this review.

Abstract: The asymmetric reduction of enoates, imines and ketones are among the most important reactions in biocatalysis. These reactions are routinely conducted using enzymes that use nicotinamide cofactors as reductants. The deazaflavin cofactor F$_{420}$ also has electrochemical properties that make it suitable as an alternative to nicotinamide cofactors for use in asymmetric reduction reactions. However, cofactor F$_{420}$-dependent enzymes remain under-explored as a resource for biocatalysis. This review considers the cofactor F$_{420}$-dependent enzyme families with the greatest potential for the discovery of new biocatalysts: the flavin/deazaflavin-dependent oxidoreductases (FDORs) and the luciferase-like hydride transferases (LLHTs). The characterized F$_{420}$-dependent reductions that have the potential for adaptation for biocatalysis are discussed, and the enzymes best suited for use in the reduction of oxidized cofactor F$_{420}$ to allow cofactor recycling in situ are considered. Further discussed are the recent advances in the production of cofactor F$_{420}$ and its functional analog F$_O$-5′-phosphate, which remains an impediment to the adoption of this family of enzymes for industrial biocatalytic processes. Finally, the prospects for the use of this cofactor and dependent enzymes as a resource for industrial biocatalysis are discussed.

Keywords: cofactor F$_{420}$; deazaflavin; oxidoreductase; hydride transfer; hydrogenation; asymmetric synthesis; cofactor biosynthesis

1. Introduction

Enzymes that catalyze the asymmetric reduction of activated double bonds are among the most important in biocatalysis, allowing access to chiral amines from imines (C=N), *sec*-alcohols from ketones C=O), and enantiopure products derived from enoates (C=C). To date, the reduction of imines, ketones and enoates has been achieved largely using enzymes that draw their reducing potential from the nicotinamide cofactors NADH and NADPH; e.g., imine reductases, ketoreductases and Old Yellow Enzymes [1–4]. However, there has been recent interest in an alternative reductive cofactor, cofactor F$_{420}$ (8-hydroxy-5-deazaflavin) [5,6].

Cofactor F$_{420}$ is a deazaflavin that is structurally similar to flavins (Figure 1), with a notable difference at position 5 of the isoalloxazine ring, which is a nitrogen in flavins and a carbon in deazaflavins. Additionally, while C-7 and C-8 are methylated in riboflavin, they are not in cofactor F$_{420}$: C-7 is hydroxylated and C-8 is unsubstituted. These structural differences cause significant differences in the electrochemical properties of cofactor F$_{420}$ and flavins: a −360–340 mV the redox mid-point

potential of cofactor F$_{420}$ is not only lower than that of the flavins (-205 mV to -220 mV), but it is also lower than that of the nicotinamides (-320 mV) [7]. Additionally, as a consequence of the substitution of N-5 for a carbon, cofactor F$_{420}$ cannot form a semiquinone (Figure 1), which means that unlike other flavins, cofactor F$_{420}$ can only perform two-electron reductions.
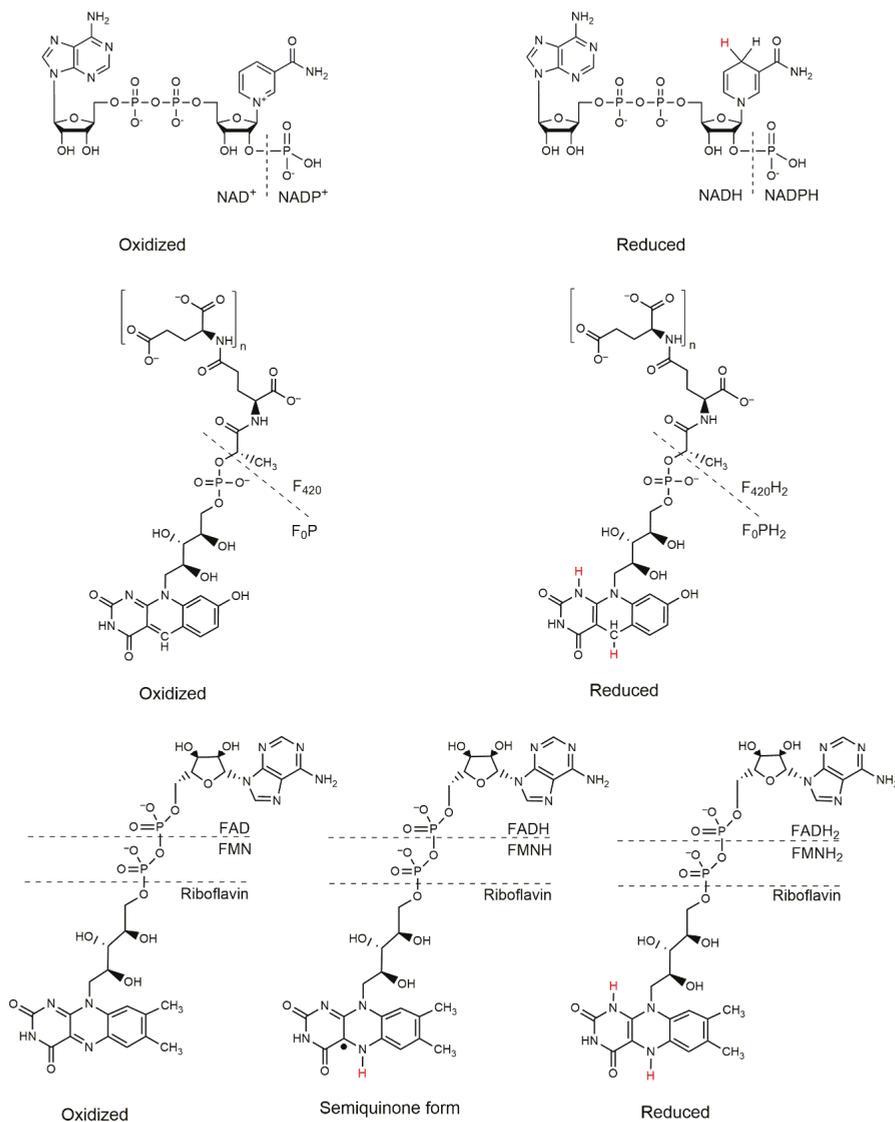
Figure 1. The structures of NAD(P) (top), cofactor F$_{420}$ and its synthetic analog F$_O$P (center) and common flavins (riboflavin, FMN and FAD; bottom). The oxidized and reduced forms are shown, as is the flavin semiquinone. Dashed lines indicate the differences in the structures of F$_O$P and cofactor F$_{420}$, and riboflavin, FMN and FAD.

Cofactor F$_{420}$ was originally described in methanogenic archaea, where it plays a pivotal role in methanogenesis [8,9]. Cofactor F$_{420}$ has since been described in a range of soil bacteria supporting

a range of metabolic activities, including catabolism of recalcitrant molecules (such as picric acid) and the production of secondary metabolites, such as antibiotics [7]. A comprehensive review of the biochemistry and physiological roles of cofactor F_{420} was recently published by Greening and coworkers [7]. This review considers the potential of F_{420}-dependent enzymes in industrial biocatalysis, focusing on the enzyme families relevant to biocatalytic applications and the reactions that they catalysis. Cofactor recycling strategies and cofactor production are also discussed, with a focus on the prospects for achieving low-cost production at scale in the latter case.

2. Families of F_{420}-Dependent Enzymes Relevant to Biocatalysis

With respect to their prospective biocatalytic applications, the two most important families of F_{420}-dependent enzymes are the Flavin/Deazaflavin Oxidoreductase (FDOR) and Luciferase-Like Hydride Transferase (LLHT) families, albeit F_{420}-dependent enzyme from other families have also been shown to have catalytic activities of interest (e.g., TomJ, the imine reducing flavin-dependent monooxygenase or OxyR, the tetracycline oxidoreductase) [10,11]. The FDOR and LLHT families are large and contain highly diverse flavin/deazaflavin-dependent enzymes. In both families, there are enzymes with preferences for flavins, such as flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), as well as those that use cofactor F_{420} [12,13]. Moreover, there are F_{420}-dependent FDORs that have been shown to be able to promiscuously bind FMN and use it in oxidation reactions [14]. In this section, the FDOR and LLHT families and the classes of reaction that they catalyze are discussed.

2.1. The FDOR Superfamily

The FDOR superfamily (PFAM Clan CL0336) can be broadly divided into two groups: the FDOR-As (which includes a sub-group called the FDOR-AAs) and the FDOR-Bs. The FDOR-As are restricted to *Actinobacteria* and *Chloroflexi* and to date no FDOR-As have been described that use cofactors other than F_{420} [7,12]. The FDOR-Bs are found in a broader range of bacterial genera than the FDOR-A enzymes, and in addition to F_{420}-dependent enzymes, this group also includes heme oxygenases, flavin-sequestering proteins, pyridoxine 5' oxidases and a number of proteins of unknown function [12,15–17]. Both groups of FDOR are highly diverse, with many homologs often found within a single bacterial genome (e.g., *Mycobacterium smegmatis* has 28 FDORs) [18]. In addition, the majority of the enzymes of this family are yet to be characterized with respect to either their biochemical or physiological function, and therefore the FDORs represent a currently under-explored source of enzymes for biocatalysis.

The FDOR enzymes share a characteristic split β-barrel fold that forms part of the cofactor-binding pocket. The majority of the protein sequences of enzymes currently identified as belonging to this family are small single-domain proteins. The topologies of the two FDOR subgroups are broadly similar (Figure 2), with the split-barrel core composed of 7–8 strands and with 4–5 helices interspersed. All FDOR-Bs studied so far have been demonstrated to be dimeric, with stands β2, β3, β5 and β6 making up the core of the dimer interface (Figure 2). In structures of full-length FDOR-As solved to date, the N-terminal helix (if present) lies on the opposite face of the beta sheet to that in FDOR-Bs. Thus, the N-terminus occupies part of the dimer interface region and prevents interaction between the sheets of adjacent monomers. In contrast to the FDOR-Bs, the oligomerization state of the FDOR-As is more varied. While a number of FDOR-As have been determined to be monomeric [18], the deazaflavin-dependent nitroreductase (DDN) from *M. tuberculosis* forms soluble aggregates through the amphipathic N-terminal helix [19]. DDN and the FDOR-AA subgroup have been shown to be membrane-associated [20–22], and FDOR-AAs have been associated with fatty acid metabolism [12]. No structures of FDOR-AAs have been solved to date.
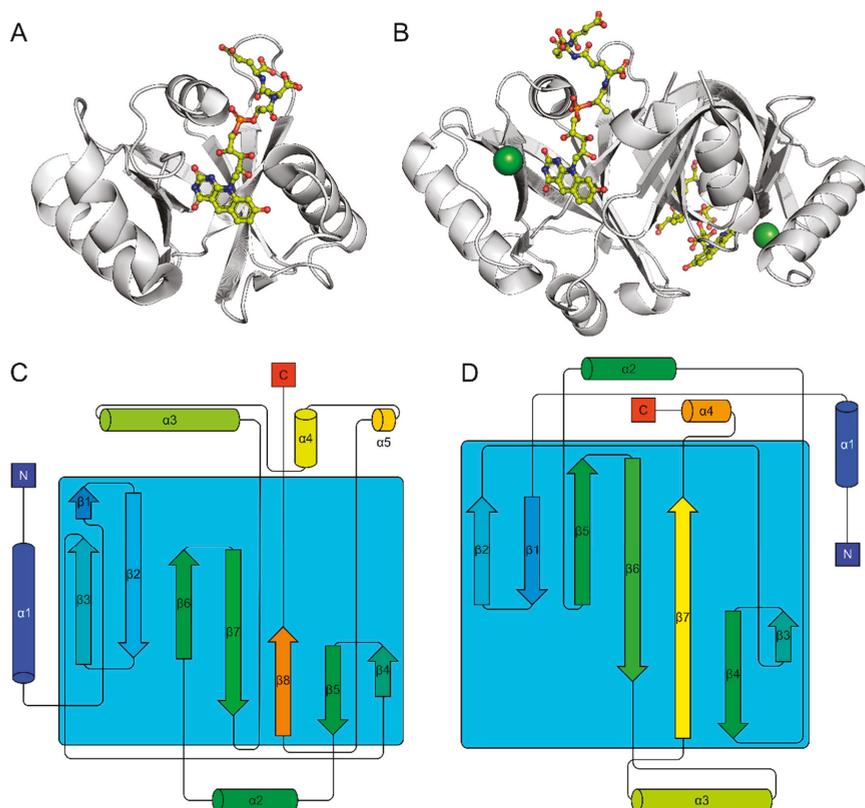
Figure 2. Representative structures of F_{420}-dependent FDOR-A (PDB: 3R5Z, panels **A** and **C**) and FDOR-B (PDB: 5JAB, panels **B** and **D**). Both are predominantly composed of a single β-sheet forming a split barrel. The N-terminal helices are spatially displaced between the two families, falling on opposite faces of the β-sheet.

2.2. The LLHT Family:

The LLHT family form part of the Luciferase-Like Monooxygenase family (PFAM PF00296). They adopt an $(\alpha/\beta)_8$ TIM-barrel fold with three insertion regions, IS1–4 (Figure 3). IS1 contains a short loop and forms part of the substrate cleft. IS2 contains two antiparallel β-strands, and IS3 contains a helical bundle at the C-terminus of the β-barrel and contains the remainder of the substrate-binding pocket (Figure 3). All structures solved to date from the LLHT family contain a non-prolyl *cis* peptide in β3 [23–26]. Recent phylogenetic reconstructions have shown that the F_{420}-dependent LLHTs form two clades: the F_{420}-dependent reductases and the F_{420}-depented dehydrogenases [27]. The F_{420}-reductases contain methylenetetrahydromethanopterin reductases (MERs), which catalyze the reversable, ring-opening cleavage of a carbon-nitrogen bond during the biosynthesis of folate in some archaea [28–30]. The F_{420}-dependent dehydrogenases can be further divided into three subgroups. The first contains F_{420}-dependent secondary alcohol dehydrogenases (ADFs) and the hydroxymycolic acid reductase from *M. tuberculosis* [31]. The second contains the F_{420}-dependent glucose-6-phosphate dehydrogenases (FGDs) from *Mycobacteria* and *Rhodococcus*, while the third appear to be more general sugar-phosphate dehydrogenases [27]. In contrast to the heterodimeric structure of bacterial luciferase, the F_{420}-dependent dehydrogenases form homodimers with the dimer interface burying a relatively large portion of the surface area of the monomers ($\approx$2000 Å^2, roughly 15%

of the total surface area) [24–26]. A number of enzymes involved in the F_{420}-dependent degradation of nitroaromatic explosives, such as picrate and 2,4-dinitroanisole, appear to belong to the LLHT family as well [32,33].
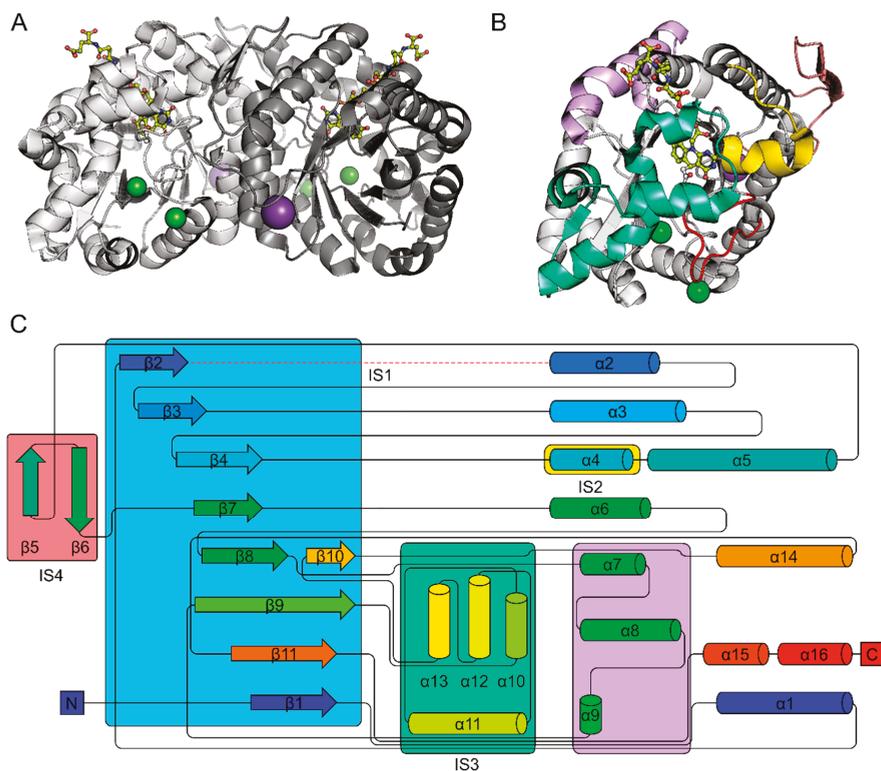
Figure 3. Structure of representative luciferase-like hydride transferase (LLHT) (PDB: 1RHC). (**A**) A 3D representation of the biologically relevant dimer (panel **A**). Monomer of an LLHT with insertion sequences IS1–4 highlighted, along with the helical bundle composed of α7–9 (panel **B**). Topology diagram showing $(\alpha/\beta)_8$ fold with insertion sequences highlighted: IS1, red; IS2, orange; IS3, light green, IS4, pink. The helical bundle of α7–9 is highlighted in purple (panel **C**).

2.3. Cofactor F_{420}-Dependent Reactions with Relevance to Biocatalysis

From the perspective of biocatalysis, cofactor F_{420}-dependent enzymes catalyze a number of key reductions including the reduction of enoates, imines, ketones and nitro-groups (Table 1; Figure 4).
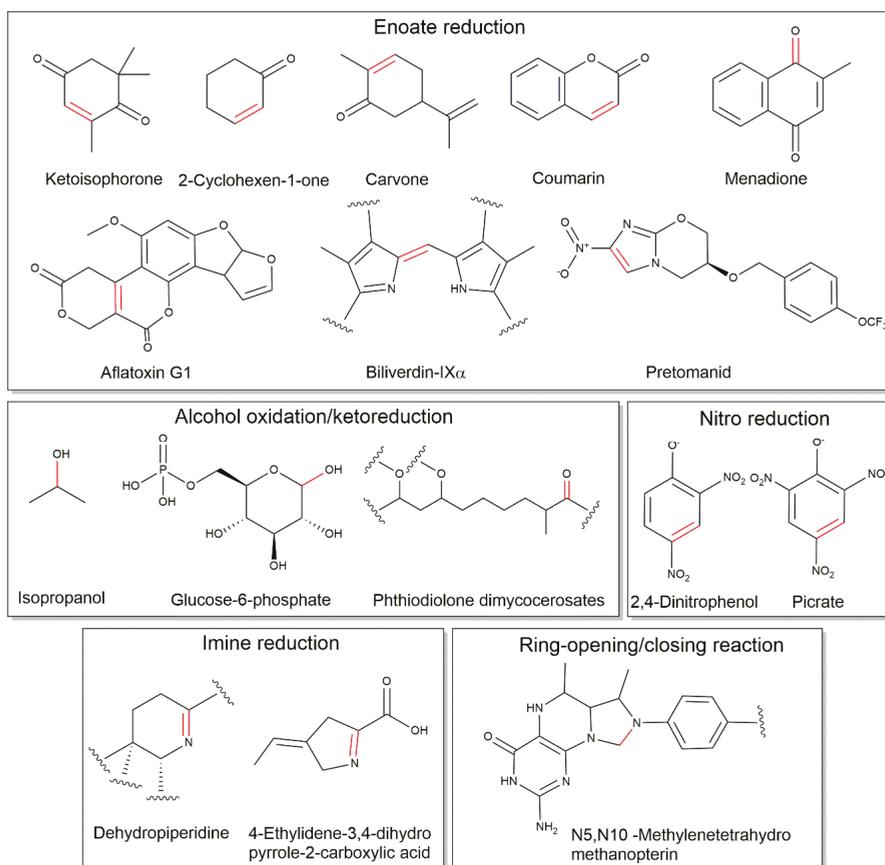
Figure 4. Representative cofactor F_{420}-dependent oxidoreductions with the potential for adaptation to biocatalytic applications. Included are: nitroreduction, enoate reduction, ketoreduction and imine reduction (from top to bottom). For clarity, only the dehydropiperidine ring of the thiopeptide is shown and partial structures for biliverdin-Ixα and phthiodiolone dimycocerosates are shown.

For enoate reductions, a small number of FDORs have been studied. However, the substrate range for most of these enzymes is yet to be fully elucidated. The ability of the mycobacterial FDORs to reduce activated C=C double bonds was first identified when DDN was shown to be responsible for activating the bicyclic nitroimidazole PA-824 in *M. tuberculosis*. These enzymes were then shown to also reduce enoates in aflatoxins, coumarins, furanocoumarins and quinones [6,12,14,16,34–38]. Recent studies have shown that these enzymes are promiscuous and can use cyclohexen-1-one, malachite green and a wide range of other activated ene compounds as substrates [35]. However, there have been a few FDOR studies to date that have examined their kinetic properties and stereospecificity. In one of these studies, FDORs from *Mycobacterium hassiacum* (FDR-Mha) and *Rhodococcus jostii* RHA1 (FDR-Rh1 and FDR-Rh2) were shown to reduce a range of structurally diverse enoates with conversions ranging from 12 to >99% and *e.e.* values of up to >99% [6]. Interestingly, it has been proposed that both the hydride and proton transfer from $F_{420}H_2$ in these reactions was directed to the same face of the activated double bond (Figure 5), which results in the opposite enantioselectivity compared to that of the FMN-dependent Old Yellow Enzyme family of enoate reductases [6]. This suggests that the F_{420}-dependent FDORs may provide a stereocomplementary enoate reductase toolbox. However, other studies suggest that protonation of the substrate is mediated by solvent or an enzyme side-chain (as it

is in Old Yellow Enzyme) [37]. Further structure/function studies are needed to fully understand the mechanistic diversity of this family of enzymes.
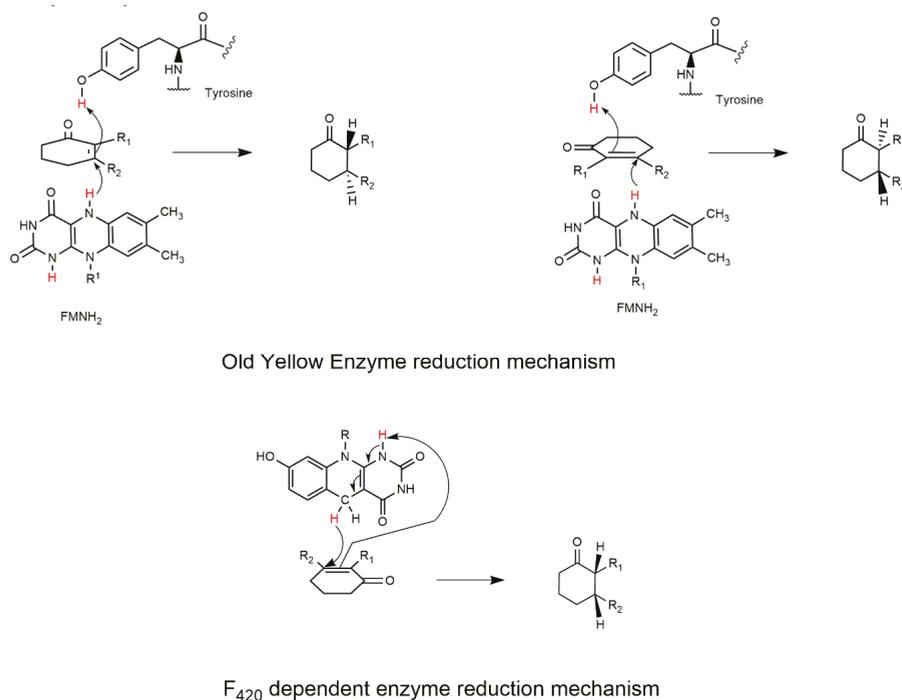

Old Yellow Enzyme reduction mechanism

F$_{420}$ dependent enzyme reduction mechanism

Figure 5. Enoate reduction by a flavin-dependent enzyme (Old Yellow Enzyme) and the proposed mechanism for cofactor F$_{420}$-dependent reduction. Notably the mechanism of reduction yields *trans*-hydrogenation products for Old Yellow Enzyme and *cis*-hydrogenation products for the F$_{420}$-dependent enzymes.

The LLHT family contains several enzymes with alcohol oxidase or ketoreductase activity (Table 1; Figure 4). The F$_{420}$-dependent glucose-6-phosphate dehydrogenases of several species have been investigated [25,26,39]. Although an extensive survey of their substrate ranges has yet to be conducted, it has been demonstrated that glucose is a substrate for the *Rhodococcus jostii* RHA1 enzymes [26]. An F$_{420}$-dependent alcohol dehydrogenase (ADH) from *Methanogenium liminatans* has been shown to catalyze the oxidation of the short chain aliphatic alcohols 2-propanol, 2-butanol and 2-pentanol (85, 49 and 23.1 s^{-1} k_{cat}, 2.2, 1.2 and 7.2 mM K_M respectively) [40], but it was unable to oxidize primary alcohols, polyols or secondary alcohols with more than five carbons. It is unclear whether these alcohol oxidations are reversible, but in the oxidative direction, these reactions provide enzymes that can be used to recycle reduced cofactor F$_{420}$ (see Section 4). Alcohol oxidation can also be used to produce ketones as intermediates in biocatalytic cascades that can then be used in subsequent reactions, such those catalyzed by transaminases or amine dehydrogenases in chiral amine synthesis [1,41–43] or by ketoreductases or alcohol dehydrogenases in chiral *sec*-alcohol synthesis (i.e., deracemization or stereoinversion of *sec*-alcohols). This approach can be achieved in a one pot cascade if different cofactors are used for the oxidation and reduction (Figure 6) [44].

At least one F$_{420}$-dependent ketoreductase has been described. The mycobacterial F$_{420}$-dependent phthiodiolone ketoreductase catalyzes a key reduction in the production of phthiocerol dimycocerosate, a diacylated polyketide found in the mycobacterial cell wall [45]. Although the physiological role of

this enzyme has been elucidated, biochemical studies of the catalytic properties and substrate range are required to assess this enzymes' potential for use as a biocatalyst.
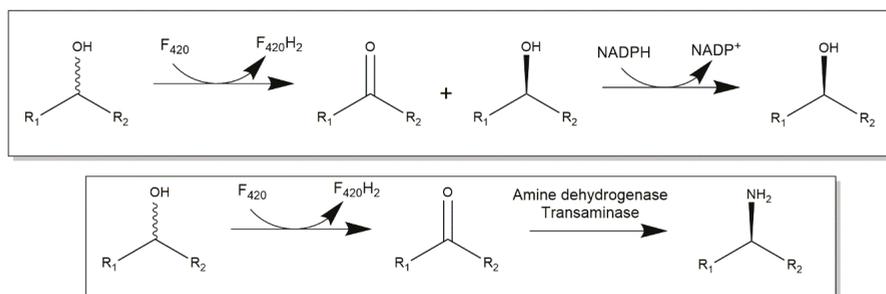

Figure 6. Proposed scheme for one-pot, enzyme cascades for deracemization/steroinversion of *sec*-alcohols (top) and chiral amine synthesis (bottom) using cofactor F_{420}-dependent alcohol oxidation.

F_{420}-dependent enzymes have also been shown to reduce imines (Table 1; Figure 4). An FDOR fromr *Streptomyces tateyamensis* (TpnL) is responsible for the reduction of dehydropiperidine in the piperidine-containing series *a* group of thiopeptide antibiotics produced in this bacterium (Figure 4). TpnL was identified as the F_{420}-dependent dehydropiperidine reductase responsible for the reduction of dehydropiperidine ring in thiostrepton A to produce the piperidine ring in the core macrocycle of thiostrepton A [45]. TpnL activity was affected by substrate inhibition at concentrations higher than 2 μM of thiostrepton A, preventing the measurement of the K_M, but its k_{cat}/K_M was measured at 2.80×10^4 M^{-1} S^{-1} [45]. The substrates for phthiodiolone ketoreductase and TpnL are large secondary metabolites and, as yet, it is unclear if it will accept smaller substrates or substrates with larger/smaller heterocycles (e.g., dehydropyrroles).

Another F_{420}-dependent imine reductase (TomJ) has been described from *Streptomyces achromogenes* that reduces the imine in 4-ethylidene-3,4-dehydropyrrole-2-carboxylic acid during the production of the secondary metabolite tomaymycin, which has been shown to have potentially interesting pharmaceutical properties [11]. Additionally, the reduction of a prochiral dihydropyrrole to a pyrrole is a reaction with a number of biocatalytic applications [5].

Nitroreductases have the potential application in the reduction of a prochiral nitro group to form a chiral amine [46]. The LLHT family F_{420}-dependent nitroreductase Npd from *Rhodococcus* catalyzes the two-electron reduction of two nitro groups in picric acid during catabolism of the explosive TNT (Table 1; Figure 4) [47]. While this stops short of reducing the nitro group to an amine, this catalytic activity may contribute to a reductive cascade that achieves this conversion.

The final class of reaction for consideration in this review is the unusual, reversable ring-opening/ closing reaction catalyzed by the MERs (Figure 4; Table 1). This reaction is required for folate biosynthesis in some archaea [23,28–30]. However, ring-closing reactions of this type could be used for producing N-containing heterocycles, which are intermediates in the synthesis of numerous pharmaceuticals [48,49]. The promiscuity of the MERs has not yet been investigated, and so the potential to re-engineer these enzymes is not fully understood.

Table 1. Characterized F_{420}-dependent enzymes with activities that could be adapted for biocatalytic applications.

Reaction	Family	Reference
Enoate reduction		
Aflatoxins	FDOR	[14,18,34]
Coumarins	FDOR	[14,34,35]
Quinones	FDOR	[36]
Biliverdin reduction	FDOR	[12,16]
Nitroimidazoles	FDOR	[36]
Cyclohexenones	FDOR	[6,34,38]
Citral/Neral/Geranial	FDOR	[6]
Carvone	FDOR	[6]
Ketoisophorone	FDOR	[6]
Alcohol oxidation/ketoreduction		
Glucose-6-phosphate	LLHT	[26,50]
Phthiodiolone dimycocerosate	LLHT	[51]
Isopropanol	LLHT	[40]
Imine reductions		
Dehydropiperidine (in thiopeptins)	FDOR	[45]
4-ethylidene-3,4-dihydropyrrole-2-carboxylic acid	Flavin-dependent monooxygenase	[11]
Nitroreductions		
Picrate	LLHT	[47,50]
2,4-DNP	LLHT	[48,50]
Ring opening/closing		
C-N bond cleavage/formation in methylenetetrahydromethanopterin	LLHT	[23,28–30]

3. Cofactor Recycling for Cofactor F_{420}

Cofactor recycling is essential for the practical application of the F_{420}-dependent enzymatic processes in biocatalysis. There are various strategies for cofactor regeneration for NADH and NADPH, including enzymatic, chemical, electrochemical and photochemical methods [52]. In this section, the potential enzymes for the regeneration of cofactor F_{420} are discussed. As most of the industrially relevant F_{420}-dependent reactions are asymmetric reductions, F_{420}-dependent oxidases are required for cofactor regeneration. Figure 7 shows the characterized enzymes that catalyze F_{420}-dependent oxidations that could be applied in cofactor F_{420} reduction.

Emulating methods developed for nicotinamide cofactors, both formate dehydrogenase (FDH) and glucose 6-phosphate dehydrogenase (G6PD) enzymes are attractive enzymatic routes for cofactor reduction both in vitro [53–56] and in vivo [57,58]. Fortunately, F_{420}-dependent G6PDs and FDHs have been identified and characterized. The F_{420}-dependent G6PD from *Mycobacteria* (FGD) is one potential cofactor F_{420}-recycling enzyme. FGD is the only enzyme in these bacteria known to reduce oxidized cofactor F_{420}. The intracellular concentration of G6P in *Mycobacteria* is up to 100-fold higher than it is in *E. coli*, which provides a ready source of reducing power for F_{420}-dependent reduction reactions [59]. FGD from *Rhodococcus jostii* and *Mycobacterium smegmatis* have been studied and expressed in *E. coli*, both the enzymes were stable in in vitro assays [26,39,60]. Both FGDs have been expressed in engineered *E. coli* producing cofactor F_{420} together with FDORs [38,59] FGDs have been shown to efficiently regenerate reduced cofactor F_{420} both in vivo and in vitro. However, the cost of the glucose-6-phosphate and the need to separate reaction products from the accumulated FGD byproduct (6-phosphoglucono-ᴅ-lactone) may prove to be impediments for the adoption of FGD as a recycling system for cofactor F_{420} in the in vitro biotransformations.
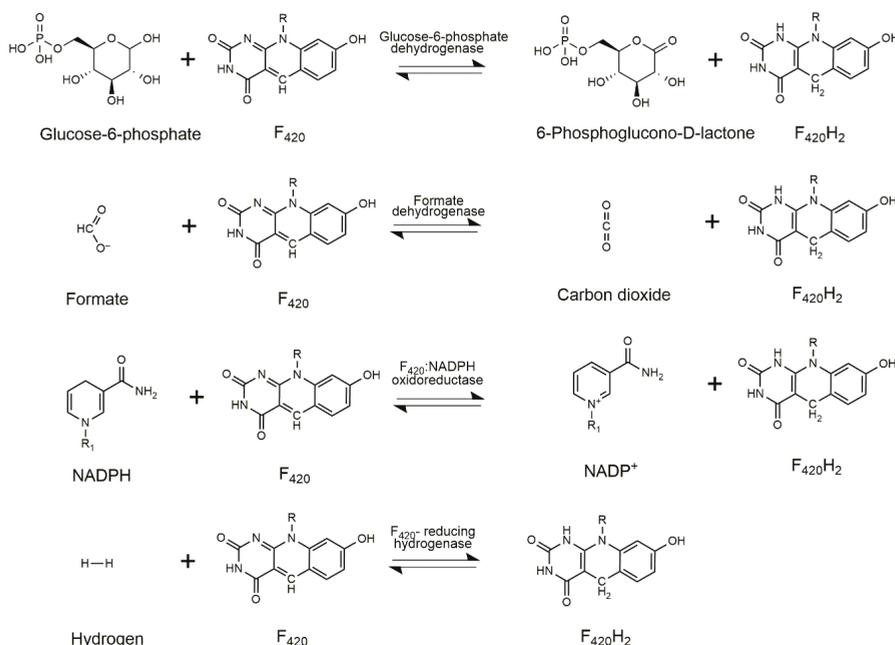
Figure 7. Cofactor F_{420}-dependent oxidation reactions that could be exploited to produce reduced cofactor F_{420}.

Formate is an excellent reductant for cofactor recycling, with FDH-dependent cofactor reduction yielding carbon dioxide, a volatile byproduct that can be easily removed from the reaction mixture, thereby simplifying the downstream processing of the product of interest. Additionally, formate is a low-cost reagent, leading to favorable process economics. Most methanogens have the capability to use formate as sole electron donor using F_{420}-dependent formate dehydrogenase [61]. The soluble F_{420}-dependent FDH from *Methanobacterium formicium* has been expressed in *E. coli* [62], purified and studied in vitro with the reduction of 41.2 µmol of F_{420} min^{-1} mg^{-1} of FDH, with non-covalently bound FAD required for optimal activity [8]. *Methanobacterium ruminantium* FDH reduces cofactor F_{420} at a much slower rate than *M. formicium*: 0.11 µmol of F_{420} min^{-1} mg^{-1} of FDH [8]. As yet, the use of F_{420}-dependnt FDHs for in vitro cofactor recycling has been sparsely studied. However, as these enzymes are soluble and can be heterologously expressed, they represent a promising system for use in cofactor F_{420}-dependent biocatalytic processes.

Another potential recycling system for cofactor F_{420} is the F_{420}:NADPH oxidoreductase (Fno), which couples the reduction of cofactor F_{420} with oxidation of NADPH. Methanogenic archaea use this enzyme to transfer reducing equivalents from hydrogenases to produce NADPH via F_{420}, while in bacteria it functions in the opposite direction, that is, to provide the cell with reduced F_{420} via NADPH [63]. Fno is also required for the production of reduced F_{420} for tetracycline production in *Streptomyces* [63]. The Fno enzymes from the thermophilic bacteria *Thermobifida fusca* and the thermophilic archaeon *Archeoglobus fulgidus* have been expressed in *E. coli* [64,65]. These enzymes are thermostable, with their highest activity observed at 65 °C. As the redox midpoint potentials of NADP and cofactor F_{420} are very similar, it is perhaps unsurprising that pH has a significant influence on the equilibrium of the reaction, with the reduction of NADP$^+$ favored at high pH (8–10) and the reduction of F_{420} favored at low pH (4–6) [64,65]. The Fno *Streptomyces griseus* has also been purified and characterized, and also displayed pH-dependent reaction directionality [66]. Fno may be an excellent enzyme for the in vivo reduction of cofactor F_{420}, where NADPH would be provided from

central metabolism. However, for its use as a cofactor F_{420} recycling enzyme in vitro, Fno would need to be coupled with an NADPH regenerating enzyme, such as an NADPH-dependent formate dehydrogenase [67]. This added complexity and cost may limit the use of Fno-dependent cofactor F_{420} recycling in vivo.

Hydrogenotrophic archaea, including methanogens and sulfate-reducing archaea, possess an essential, cofactor F_{420}-dependent hydrogenase (FhrAGB) [68–71]. These nickel/iron enzymes could potentially be used in vivo to allow the direct H_2-dependent reduction of cofactor F_{420}. However, as these heterododecameric enzymes have complex cofactor requirements (four [4Fe 4S] clusters, and NiFe center and FAD), are oxygen-sensitive and tend to aggregate [71], it is unclear if they can be made suitable for in vitro use.

4. Cofactor Production

The lack of a scalable production system for cofactor F_{420} has been noted as a major impediment to the adoption of F_{420}-dependent enzymes by industry [5]. Cofactor F_{420} is available as a research reagent (http://www.gecco-biotech.com/), but its production at scale is not yet economic. In fact, most research laboratories with an interest in cofactor F_{420}-dependent enzymes synthesize and purify the cofactor themselves using slow-growing F_{420} producing microorganisms, most commonly methanogens and actinobacteria (Table 2). The economic production of cofactor F_{420} at large scale is not feasible using natural producers as they are ill-suited to industrial fermentation and generally lack the genetic tools required to improve cofactor F_{420} yield.

Table 2. Published production systems for cofactor F_{420}.

Source	F_{420} Yield (μmol/g Cell Weight)	Growth Conditions	Ref
Methanobacterium thermoautotrophicum	0.42 [a,c]	Grown at 60 °C using complex media in fermenter, under pressurized hydrogen	[9]
Methanobacterium formicium	0.27 [a,c]	Grown at 37 °C using complex media in fermenters	[9]
Methanospirillum hungatii	0.41 [a,c]	Grown at 37 °C using complex media in fermenters	[9]
Methanobacterium strain M.o.H	0.53 [a,c]	Grown at 40 °C using complex media in fermenters	[9]
Methanobacterium thermoautotrophicum	1.7 [e]	Grown using complex media in fermenters, under pressurized hydrogen gas	[73]
Streptomyces flocculus	0.62 [e]	Grown using complex media in fermenters	[73]
Streptomyces coelicolor	0.04 [e]	Grown using complex media in fermenters	[73]
Streptomyces griseus	0.008 [a,c]	Growth conditions not mentioned in the publication	[74]
Rhodococcus rhodochrous	0.11 [e]	Grown using complex media in fermenters	[73]
Mycobacterium smegmatis	0.30 [e]	Grown using complex media in fermenters	[73]
Mycobacterium smegmatis	3.0 [d]	Overexpression of F_{420} pathway genes, cultivation in complex media at 37 °C in shake flasks	[72]
Escherichia coli	0.38 [b]	Overexpressing F_{420} pathway genes, grown in minimal media at 30 °C in shake flasks.	[59]

[a] Mol weight of F_{420} with 1 glutamate tail is 773.6 Da, which was used to convert values published as μg of F_{420}, noting that micro-organisms produce mixture of F_{420} with different number of glutamates (1–9) attached. [b] Concentration estimated through absorbance at 420 nm and using extinction coefficient of 41.4 mM^{-1} cm^{-1} [73]. [c] F_{420} concentration per g of wet cell weight. [d] Concentration of F_{420} not mentioned in the publication, but F_{420} yield was stated to be 10 times higher than wild-type *M. smegmatis*. [e] Concentration estimated through absorbance at 400 nm and using extinction coefficient of 25.7 mM^{-1} cm^{-1} [74].

Recently, there have been significant advances towards the scalable production of the cofactor for F_{420}-dependent enzymes. *M. smegmatis* has been engineered to overexpress the biosynthetic genes for cofactor F_{420} production, leading to a substantial improvement in yields (Table 2) [72]. However, *M. smegmatis* is not ideally suited as a fermentation organism as it is slow growing, forms clumps

during cultivation and is not recognized as GRAS (generally regarded as safe). More recently, the biosynthetic pathway for cofactor F_{420} has been successfully transplanted to *E. coli* [59], allowing the heterologous production of the cofactor at levels similar to those of the natural F_{420} producers (Table 2) [59], accumulated to 0.38 μmol of F_{420} per gram of dry cells.

There is scope to further improve the production of F_{420} in *E. coli*. Cofactor F_{420} does not appear to be toxic to *E. coli* [59], which suggests that there is little interaction between F_{420} and the enzymes *E. coli* (although this is yet to be confirmed experimentally). The thermodynamics of cofactor F_{420} production are favorable (Appendix A), suggesting that there are no major thermodynamic impediments to improving yield. Interestingly, the first dedicated step of cofactor F_{420} production (catalyzed by CofC/FbiD) is not energetically favorable and may consequently be sensitive to intracellular metabolite concentrations. In addition to the engineering considerations that this may impose, it may also be responsible for the biochemical diversity of this step in different microorganisms. In different microbes, the CofC/FbiD-dependent step uses 2-phospholactate [75], 3-phosphoglycerate [76] or phosphoenolpyruvate [59] as a substrate, which may reflect the relative abundance of those metabolites in various bacteria and archaea and the thermodynamic constraints on this step.

Another recent advance is the production of a synthetic analog of cofactor F_{420}, called F_O-5′-phosphate (F_OP). F_OP was derived from F_O, the metabolic precursor of cofactor F_{420}, which is phosphorylated using an engineered riboflavin kinase [38]. F_OP has also been shown to function as an active cofactor for cofactor F_{420}-dependent enzymes activities, albeit there is a penalty in the rates of these reactions [38]. Drenth and coworkers prepared F_O by chemical synthesis, using a method developed by Hossain et al. [77]. However, it is likely that the engineered kinase for the phosphorylation of F_O could be introduced into an organism that over-produces F_O allowing for the production of F_OP by fermentation. This semisynthetic pathway would have the advantage that it needs only two biosynthetic steps, instead of the four steps needed for cofactor F_{420} production, and demands less metabolic input from the native host metabolism (e.g., no glutamate is required) [38]. The production of F_OP also opens the possibility of making deazaflavin analogs of FMN and FAD, which would be electrochemically more like F_{420} than flavins, but may still bind FMN and FAD- dependent enzymes and potentially allow access to new chemistry with already well-characterized enzymes.

5. Prospects

Reduced cofactor F_{420} is electrochemically well suited for biocatalytic applications, and the small number of F_{420}-dependent enzymes characterized to date show promise as potential biocatalysts (as discussed above). However, before these enzymes can be widely and effectively used as biocatalysts, further research is needed to better characterize them as the biochemistry of cofactor F_{420}-dependent enzymes remains under-explored. The LLHT and FDOR families are a rich source of highly diverse enzymes with considerable potential for biocatalysis, albeit much of the research to date has focused on the physiological roles of these enzymes, rather than their in vitro enzymology. Although some of these enzymes have been shown to have small molecule substrates, those involved with secondary metabolite biosynthesis tend to act on high molecular weight substrates and it is not yet clear whether they will accept lower molecular weight molecules.

To be cost competitive, cofactor F_{420} needs to have effective recycling systems. The enzymes for cofactor recycling have already been identified, although there have been a few studies investigating their performance in this role. Moreover, alternative cofactor recycling strategies, such electrochemical or photochemical recycling, have not yet been investigated for cofactor F_{420}. The production of cofactor F_{420} at scale and at low cost remains a roadblock for the use of these enzymes by industry. However, considerable progress has been made on this front in the last few years and it is likely that low cost cofactor F_{420}, or F_{420} surrogates, will soon be available. Additionally, the availability of F_{420}-producing bacteria with tools for facile genetic manipulation, along with a growing number of empirically determined protein structures, opens up the prospect of improving this class of enzymes using in vitro evolution and rational design. It is notable that there is still some uncertainty concerning

the mechanistic detail of F_{420}-dependent reactions, which need to be addressed through a detailed structure/function analysis to enable a rational design of these enzymes.

Abbreviations

5AD: 5′-Deoxyadenosine; 5ARPD: 5-amino-6-(D-ribitylamino)uracil; 5ARPD4HB: 5-amino-5-(4-hydroxybenzyl)-6-(D-ribitylimino)-5,6-dihydrouracil; dF_{420}-0: Dehydro coenzyme F_{420}-0 (oxidized); EPPG: Enolpyruvyl-diphospho-5′-guanosine; F_{o}: 7,8-Didemethyl-8-hydroxy-5-deazariboflavin; F_{420}-0: Coenzyme F_{420}-0 (oxidized); F_{420}-1: Coenzyme F_{420}-1 (oxidized); F_{420}-2: Coenzyme F_{420}-2 (oxidized); F_{420}-3: Coenzyme F_{420}-3 (oxidized); FMN: Flavin mononucleotide (oxidized); FMNH$_2$: Flavin mononucleotide (reduced); GDP: Guanosine diphosphate; GMP: Guanosine monophosphate; Glu: L-Glutamate; GTP: Guanosine triphosphate; H$^+$: Proton; ImiAce: 2-iminoacetate or Dehydroglycine; Met: L-Methionine; NH$_4$: Ammonium; PEP: Phosphoenolpyruvate; Pi: Phosphate; PPi: Diphosphate; SAMe: S-Adenosyl-L-methionine; Tyr: L-Tyrosine.

Appendix A Thermodynamics of F420 Biosynthesis

The thermodynamic properties of each of the steps in cofactor F_{420} biosynthesis were estimated to evaluate the feasibility of increasing the production of the cofactor in an engineered microorganism. The pathway assembled by Bashiri et al. [59] in *E. coli* was used (i.e., PEP was used as substrate for CofC). The standard transformed Gibbs free energy ($\Delta_r G^t$) of each step were calculated under the physiological conditions (25 °C, pH 7, and ionic concentration of 0.25 M) as described elsewhere [78,79]. The overall Gibbs free energy (ΔG^t) was then calculated by summing up all individual $\Delta_r G^t$ (Table A1). The Gibbs free energy of metabolite formation ($\Delta_f G$) for each metabolite in the pathway was obtained (Supplementary Information) from comprehensive lists of metabolites whose $\Delta_f G$ were estimated using a group contribution method [80,81]. The $\Delta_f G$ for each metabolite was then converted into its transformed type ($\Delta_f G^t$) method of Alberty [78]. The data were collected from relevant biochemical databases and the literature for any metabolite with missing $\Delta_f G$ [82–84]. Owing to possessing different protonation states, the inconsistencies in $\Delta_f G$ of certain metabolites such as the glutamates in F_{420}-n among databases and the literature are inevitable. Thus, $\Delta_r G^t$ for reactions containing metabolites with varying $\Delta_f G$ were calculated considering the differences in their $\Delta_f G$ leading to the generation of a total of four sets of $\Delta_r G^t$. Finally, the mean and standard deviations were calculated for these sets to yield the variation in each reaction as well as in the overall pathway (Table A1).

The data shown in Table A1 confirms that the overall cofactor F_{420} biosynthesis pathway is thermodynamically feasible under the given conditions. However, certain steps in this pathway impose a thermodynamic barrier with respect to the physiological conditions examined. For example, CofC seems to be one of the major thermodynamically unfavorable steps in the whole pathway possibly due to the energy-dependent synthesis of EPPG, one of the precursors for making F_{420}. CofG/H combined appears to be the most thermodynamically favorable step in the whole pathway driving the biosynthesis of F_0, the other key precursor for F_{420} biosynthesis. Interestingly, the formation of F_{420}-2 molecule seems to be the most favorable step among other F_{420} molecules downstream of the pathway. It should be noted that the thermodynamic calculations were only performed up to three steps of F_{420} molecule production (i.e., F_{420}-3) largely because of the high levels of inconsistencies of the data available for $\Delta_f G$ of higher F_{420} molecules.

Table A1. Standard transformed Gibbs free energy of reaction ($\Delta_r G^t$), for the F_{420} biosynthesis pathway, calculated based on Gibbs free energy of metabolite formation ($\Delta_f G^t$) calculated at 25 °C, pH of 7, and ionic concentration of 0.25 M.

Enzyme	Reaction [a]	$\Delta_r G^t$ (kJ) [b]
CofC/FbiD	PEP + GTP → EPPG + PP$_i$ [d]	+71.27(±67)
CofG/FbiC	5ARPD + Tyr + SAMe → 5ARPD4HB + ImAcet + Met + 5AD	−1192.39(±0) [c]
CofH/FbiC	5ARPD4HB + SAMe → F$_O$ + NH$_4^+$ + Met + 5AD	+71.90(±36) [c]
CofD/FbiA	F$_O$ + EPPG → d$_{F420}$-0 + GMP	−31.3(±128)
CofX/FbiB	d$_{F420}$-0 + FMNH$_2$ → F$_{420}$-0 + FMN	−74.59(±87)
CofE/FbiB	F$_{420}$-0 + GTP + Glu → F$_{420}$-1 + GDP + P$_i$	−7.50(±24)
CofE/FbiB	F$_{420}$-1 + GTP + Glu → F$_{420}$-2 + GDP + P$_i$	−39.44(±35)
CofE/FbiB	F$_{420}$-2 + GTP + Glu → F$_{420}$-3 + GDP + P$_i$	−21.99(±38)
Overall	PEP + 5ARPD + Tyr + (2) SAMe + FMNH$_2$ + (3) Glu + (4) GTP → F$_{420}$-3 + (2) Met + (2) 5AD + ImAcet + NH$_4^+$ + FMN + (3) GDP + (3) P$_i$ + GMP + PP$_i$	−1224.05(±82)

[a] For simplicity, protons were omitted in these equations and subsequent calculations as the $\Delta_f G^t$ of a proton under the set conditions is ~0.08 kJ. However, all $\Delta_r G^t$ calculations are based on a balanced equation. [b] The mean values of four sets and their standard deviations in parenthesis shown for each reaction. [c] $\Delta_f G$ of 5ARPD4HB has only been reported in MetaCyc inferred by computational analysis. Including it in the calculations of $\Delta_r G^t$ for CofG and CofH results in −225.88(±0) and −894.62(±36), respectively. [d] Hydrolysis of PP$_i$ ($H_3P_2O_7^{3-}$ + H_2O → 2 HPO_4^{2-} + H^+) yields a $\Delta_r G^t$ of ~17 kJ/mole, resulting in less than 2% change in the overall $\Delta_r G^t$.

References

1. Patil, M.D.; Grogan, G.; Bommarius, A.; Yun, H. Oxidoreductase-catalyzed synthesis of chiral amines. *ACS Catal.* **2018**, *8*, 10985–11015. [CrossRef]
2. Toogood, H.S.; Scrutton, N.S. New developments in 'ene'-reductase catalysed biological hydrogenations. *Curr. Opin. Chem. Biol.* **2014**, *19*, 107–115. [CrossRef] [PubMed]
3. Cosgrove, S.C.; Brzezniak, A.; France, S.P.; Ramsden, J.I.; Mangas-Sanchez, J.; Montgomery, S.L.; Heath, R.S.; Turner, N.J. Imine reductases, reductive aminases, and amine oxidases for the synthesis of chiral amines: Discovery, characterization, and synthetic applications. In *Enzymes in Synthetic Biology*; Scrutton, N., Ed.; Elsevier Academic Press Inc.: San Diego, CA, USA, 2018; Volume 608, pp. 131–149.
4. Bai, D.Y.; He, J.Y.; Ouyang, B.; Huang, J.; Wang, P. Biocatalytic asymmetric synthesis of chiral aryl alcohols. *Prog. Chem.* **2017**, *29*, 491–501.
5. Taylor, M.; Scott, C.; Grogan, G. F-420-dependent enzymes-potential for applications in biotechnology. *Trends Biotechnol.* **2013**, *31*, 63–64. [CrossRef]
6. Mathew, S.; Trajkovic, M.; Kumar, H.; Nguyen, Q.-T.; Fraaije, M.W. Enantio- and regioselective *ene*-reductions using F$_{420}$H$_2$-dependent enzymes. *Chem. Commun.* **2018**, *54*, 11208–11211. [CrossRef]
7. Greening, C.; Ahmed, F.H.; Mohamed, A.E.; Lee, B.M.; Pandey, G.; Warden, A.C.; Scott, C.; Oakeshott, J.G.; Taylor, M.C.; Jackson, C.J. Physiology, biochemistry, and applications of F$_{420}$- and Fo-dependent redox reactions. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 451–493. [CrossRef]
8. Tzing, S.F.; Bryant, M.P.; Wolfe, R.S. Factor 420-dependent pyridine nucleotide-linked formate metabolism of *Methanobacterium ruminantium*. *J. Bacteriol.* **1975**, *121*, 192–196.
9. Eirich, L.D.; Vogels, G.D.; Wolfe, R.S. Distribution of coenzyme F$_{420}$ and properties of its hydrolytic fragments. *J. Bacteriol.* **1979**, *140*, 20–27.
10. Wang, P.; Bashiri, G.; Gao, X.; Sawaya, M.R.; Tang, Y. Uncovering the enzymes that catalyze the final steps in oxytetracycline biosynthesis. *J. Am. Chem. Soc.* **2013**, *135*, 7138–7141. [CrossRef]
11. Li, W.; Chou, S.C.; Khullar, A.; Gerratana, B. Cloning and characterization of the biosynthetic gene cluster for tomaymycin, an SJG-136 monomeric analog. *Appl. Environ. Microbiol.* **2009**, *75*, 2958–2963. [CrossRef]
12. Ahmed, F.H.; Carr, P.D.; Lee, B.M.; Afriat-Jurnou, L.; Mohamed, A.E.; Hong, N.-S.; Flanagan, J.; Taylor, M.C.; Greening, C.; Jackson, C.J. Sequence–structure–function classification of a catalytically diverse oxidoreductase superfamily in *mycobacteria*. *J. Mol. Biol.* **2015**, *427*, 3554–3571. [CrossRef] [PubMed]
13. Selengut, J.D.; Haft, D.H. Unexpected abundance of coenzyme F$_{420}$-dependent enzymes in *Mycobacterium tuberculosis* and other actinobacteria. *J. Bacteriol.* **2010**, *192*, 5788–5798. [CrossRef] [PubMed]

14. Lapalikar, G.V.; Taylor, M.C.; Warden, A.C.; Onagi, H.; Hennessy, J.E.; Mulder, R.J.; Scott, C.; Brown, S.E.; Russell, R.J.; Easton, C.J.; et al. Cofactor promiscuity among F_{420}-dependent reductases enables them to catalyse both oxidation and reduction of the same substrate. *Catal. Sci. Technol.* **2012**, *2*, 1560–1567. [CrossRef]

15. Harold, L.K.; Antoney, J.; Ahmed, F.H.; Hards, K.; Carr, P.D.; Rapson, T.; Greening, C.; Jackson, C.J.; Cook, G.M. FAD-sequestering proteins protect Mycobacteria against hypoxic and oxidative stress. *J. Biol. Chem.* **2019**, *294*, 2903–2912. [CrossRef] [PubMed]

16. Ahmed, F.H.; Mohamed, A.E.; Carr, P.D.; Lee, B.M.; Condic-Jurkic, K.; O'Mara, M.L.; Jackson, C.J. Rv2074 is a novel $F_{420}H_2$-dependent biliverdin reductase in *Mycobacterium tuberculosis*. *Protein Sci.* **2016**, *25*, 1692–1709. [CrossRef] [PubMed]

17. Mashalidis, E.H.; Mukherjee, T.; Śledź, P.; Matak-Vinković, D.; Boshoff, H.I.; Abell, C.; Barry, C.E. Rv2607 from *Mycobacterium tuberculosis* is a pyridoxine 5′-phosphate oxidase with unusual substrate specificity. *PLoS ONE* **2011**, *6*, e27643. [CrossRef] [PubMed]

18. Taylor, M.C.; Jackson, C.J.; Tattersall, D.B.; French, N.; Peat, T.S.; Newman, J.; Briggs, L.J.; Lapalikar, G.V.; Campbell, P.M.; Scott, C.; et al. Identification and characterization of two families of $F_{420}H_2$-dependent reductases from *Mycobacteria* that catalyse aflatoxin degradation. *Mol. Microbiol.* **2010**, *78*, 561–575. [CrossRef]

19. Cellitti, S.E.; Shaffer, J.; Jones, D.H.; Mukherjee, T.; Gurumurthy, M.; Bursulaya, B.; Boshoff, H.I.; Choi, I.; Nayyar, A.; Lee, Y.S.; et al. Structure of DDN, the deazaflavin-dependent nitroreductase from *Mycobacterium tuberculosis* involved in bioreductive activation of PA-824. *Structure* **2012**, *20*, 101–112. [CrossRef]

20. De Souza, G.A.; Leversen, N.A.; Målen, H.; Wiker, H.G. Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *J. Proteom.* **2011**, *75*, 502–510. [CrossRef]

21. He, Z.; De Buck, J. Cell wall proteome analysis of *Mycobacterium smegmatis* strain mc2 155. *BMC Microbiol.* **2010**, *10*, 121. [CrossRef]

22. Sinha, S.; Kosalai, K.; Arora, S.; Namane, A.; Sharma, P.; Gaikwad, A.N.; Brodin, P.; Cole, S.T. Immunogenic membrane-associated proteins of *Mycobacterium tuberculosis* revealed by proteomics. *Microbiology* **2005**, *151*, 2411–2419. [CrossRef] [PubMed]

23. Aufhammer, S.W.; Warkentin, E.; Ermler, U.; Hagemeier, C.H.; Thauer, R.K.; Shima, S. Crystal structure of methylenetetrahydromethanopterin reductase (MER) in complex with coenzyme F_{420}: Architecture of the F_{420}/FMN binding site of enzymes within the nonprolyl *cis*-peptide containing bacterial luciferase family. *Protein Sci.* **2005**, *14*, 1840–1849. [CrossRef] [PubMed]

24. Aufhammer, S.W.; Warkentin, E.; Berk, H.; Shima, S.; Thauer, R.K.; Ermler, U. Coenzyme binding in F_{420}-dependent secondary alcohol dehydrogenase, a member of the bacterial luciferase family. *Structure* **2004**, *12*, 361–370. [CrossRef] [PubMed]

25. Bashiri, G.; Squire, C.J.; Moreland, N.J.; Baker, E.N. Crystal structures of F_{420}-dependent glucose-6-phosphate dehydrogenase FGD1 involved in the activation of the anti-tuberculosis drug candidate PA-824 reveal the basis of coenzyme and substrate binding. *J. Biol. Chem.* **2008**, *283*, 17531–17541. [CrossRef] [PubMed]

26. Nguyen, Q.T.; Trinco, G.; Binda, C.; Mattevi, A.; Fraaije, M.W. Discovery and characterization of an F_{420}-dependent glucose-6-phosphate dehydrogenase (Rh-FGD1) from *rhodococcus jostii* rha1. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 2831–2842. [CrossRef] [PubMed]

27. Mascotti, M.L.; Kumar, H.; Nguyen, Q.-T.; Ayub, M.J.; Fraaije, M.W. Reconstructing the evolutionary history of F_{420}-dependent dehydrogenases. *Sci. Rep.* **2018**, *8*, 17571. [CrossRef] [PubMed]

28. Ceh, K.; Demmer, U.; Warkentin, E.; Moll, J.; Thauer, R.K.; Shima, S.; Ermler, U. Structural basis of the hydride transfer mechanism in F_{420}-dependent methylenetetrahydromethanopterin dehydrogenase. *Biochemistry* **2009**, *48*, 10098–10105. [CrossRef]

29. Shima, S.; Warkentin, E.; Grabarse, W.; Sordel, M.; Wicke, M.; Thauer, R.K.; Ermler, U. Structure of coenzyme F_{420} dependent methylenetetrahydromethanopterin reductase from two methanogenic archaea. *J. Mol. Biol.* **2000**, *300*, 935–950. [CrossRef]

30. Vaupel, M.; Thauer, R.K. Coenzyme F_{420} dependent N-5, N-10-methylenetetrahydromethanopterin reductase (MER) from *Methanobacterium thermautotrophicum* strain marburg: Cloning, sequencing, transcriptional analysis and functional expression in *Escherichia coli* of the *mer* gene. *Eur. J. Biochem.* **1995**, *231*, 773–778.

31. Purwantini, E.; Mukhopadhyay, B. Rv0132c of *Mycobacterium tuberculosis* encodes a coenzyme F_{420}-dependent hydroxymycolic acid dehydrogenase. *PLoS ONE* **2013**, *8*, e81985. [CrossRef]

32. Fida, T.T.; Palamuru, S.; Pandey, G.; Spain, J.C. Aerobic biodegradation of 2,4-dinitroanisole by *Nocardioides* sp. Strain js1661. *Appl. Environ. Microbiol.* **2014**, *80*, 7725–7731. [CrossRef] [PubMed]

33. Ebert, S.; Rieger, P.G.; Knackmuss, H.J. Function of coenzyme f$_{420}$ in aerobic catabolism of 2,4,6-trinitrophenol and 2,4-dinitrophenol by *Nocardioides simplex* FJ2-1A. *J. Bacteriol.* **1999**, *181*, 2669–2674. [PubMed]

34. Lapalikar, G.V.; Taylor, M.C.; Warden, A.C.; Scott, C.; Russell, R.J.; Oakeshott, J.G. F$_{420}$H$_2$-dependent degradation of aflatoxin and other furanocoumarins is widespread throughout the *Actinomycetales*. *PLoS ONE* **2012**, *7*, e30114. [CrossRef] [PubMed]

35. Greening, C.; Jirapanjawat, T.; Afroze, S.; Ney, B.; Scott, C.; Pandey, G.; Lee, B.M.; Russell, R.J.; Jackson, C.J.; Oakeshott, J.G.; et al. Mycobacterial F$_{420}$H$_2$-dependent reductases promiscuously reduce diverse compounds through a common mechanism. *Front. Microbiol.* **2017**, *8*. [CrossRef]

36. Gurumurthy, M.; Rao, M.; Mukherjee, T.; Rao, S.P.S.; Boshoff, H.I.; Dick, T.; Barry, C.E.; Manjunatha, U.H. A novel F$_{420}$-dependent anti-oxidant mechanism protects *Mycobacterium tuberculosis* against oxidative stress and bactericidal agents. *Mol. Microbiol.* **2013**, *87*, 744–755. [CrossRef]

37. Mohamed, A.E.; Ahmed, F.H.; Arulmozhiraja, S.; Lin, C.Y.; Taylor, M.C.; Krausz, E.R.; Jackson, C.J.; Coote, M.L. Protonation state of F$_{420}$H$_2$ in the prodrug-activating deazaflavin dependent nitroreductase (DDN) from *Mycobacterium tuberculosis*. *Mol. BioSys.* **2016**, *12*, 1110–1113. [CrossRef]

38. Drenth, J.; Trajkovic, M.; Fraaije, M.W. Chemoenzymatic synthesis of an unnatural deazaflavin cofactor that can fuel F$_{420}$-dependent enzymes. *ACS Catal.* **2019**, *9*, 6435–6443. [CrossRef]

39. Purwantini, E.; Daniels, L. Molecular analysis of the gene encoding F$_{420}$-dependent glucose-6-phosphate dehydrogenase from *Mycobacterium smegmatis*. *J. Bacteriol.* **1998**, *180*, 2212–2219.

40. Bleicher, K.; Winter, J. Purification and properties of F$_{420}$-and NADP$^+$-dependent alcohol dehydrogenases of *Methanogenium liminatans* and *Methanobacterium palustre*, specific for secondary alcohols. *Europ. J. Biochem.* **1991**, *200*, 43–51. [CrossRef]

41. Knaus, T.; Cariati, L.; Masman, M.F.; Mutti, F.G. In vitro biocatalytic pathway design: Orthogonal network for the quantitative and stereospecific amination of alcohols. *Org. Biomol. Chem.* **2017**, *15*, 8313–8325.

42. Guo, F.; Berglund, P. Transaminase biocatalysis: Optimization and application. *Green Chem.* **2017**, *19*, 333–360. [CrossRef]

43. Adams, J.P.; Brown, M.J.B.; Diaz-Rodriguez, A.; Lloyd, R.C.; Roiban, G.D. Biocatalysis: A pharma perspective. *Adv. Synth. Catal.* **2019**, *361*, 2421–2432. [CrossRef]

44. Musa, M.M.; Hollmann, F.; Mutti, F.G. Synthesis of enantiomerically pure alcohols and amines *via* biocatalytic deracemisation methods. *Cat. Sci. Technol.* **2019**, *9*, 10–1039. [CrossRef]

45. Ichikawa, H.; Bashiri, G.; Kelly, W.L. Biosynthesis of the thiopeptins and identification of an F$_{420}$H$_2$-dependent dehydropiperidine reductase. *J. Am. Chem. Soc.* **2018**, *140*, 10749–10756. [CrossRef]

46. Miller, A.F.; Park, J.T.; Ferguson, K.L.; Pitsawong, W.; Bommarius, A.S. Informing efforts to develop nitroreductase for amine production. *Molecules* **2018**, *23*, 22. [CrossRef]

47. Heiss, G.; Hofmann, K.W.; Trachtmann, N.; Walters, D.M.; Rouvière, P.; Knackmuss, H.J. *Npd* gene functions of *Rhodococcus (opacus) erythropolis* Hl PM-1 in the initial steps of 2,4,6-trinitrophenol degradation. *Microbiology* **2002**, *148*, 799–806. [CrossRef]

48. Xu, J.; Green, A.P.; Turner, N.J. Chemo-enzymatic synthesis of pyrazines and pyrroles. *Angew. Chem.-Int. Ed.* **2018**, *57*, 16760–16763. [CrossRef]

49. Busacca, C.A.; Fandrick, D.R.; Song, J.J.; Senanayake, C.H. The growing impact of catalysis in the pharmaceutical industry. *Adv. Synth. Catal.* **2011**, *353*, 1825–1864. [CrossRef]

50. Heiss, G.; Trachtmann, N.; Abe, Y.; Takeo, M.; Knackmuss, H.J. Homologous *npdgi* genes in 2,4-dinitrophenol- and 4-nitrophenol-degrading *Rhodococcus* spp. *Appl. Environ. Microbiol.* **2003**, *69*, 2748–2754. [CrossRef]

51. Purwantini, E.; Daniels, L.; Mukhopadhyay, B. F$_{420}$H$_2$ is required for phthiocerol dimycocerosate synthesis in *Mycobacteria*. *J. Bacteriol.* **2016**, *198*, 2020–2028. [CrossRef]

52. Wichmann, R.; Vasic-Racki, D. Cofactor regeneration at the lab scale. In *Technology Transfer in Biotechnology*; Springer: Berlin, Germany, 2005; pp. 225–260.

53. Tishkov, V.I.; Popov, V.O. Catalytic mechanism and application of formate dehydrogenase. *Biochemistry* **2004**, *69*, 1252. [CrossRef] [PubMed]

54. Eguchi, T.; Kuge, Y.; Inoue, K.; Yoshikawa, N.; Mochida, K.; Uwajima, T. NADPH regeneration by glucose dehydrogenase from *Gluconobacter scleroides* for L-leucovorin synthesis. *Biosci. Biotechnol. Biochem.* **1992**, *56*, 701–703. [CrossRef] [PubMed]

55. Demir, A.S.; Talpur, F.N.; Sopaci, B.; Kohring, G.-W.; Celik, A. Selective oxidation and reduction reactions with cofactor regeneration mediated by galactitol-, lactate-, and formate dehydrogenases immobilized on magnetic nanoparticles. *J. Biotechnol.* **2011**, *152*, 176–183. [CrossRef] [PubMed]

56. Wong, C.-H.; Whitesides, G.M. Enzyme-catalyzed organic synthesis: NAD(P)H cofactor regeneration by using glucose-6-phosphate and the glucose-5-phosphate dehydrogenase from *Leuconostoc mesenteroides*. *J. Am. Chem. Soc.* **1981**, *103*, 4890–4899. [CrossRef]

57. Lee, W.-H.; Park, J.-B.; Park, K.; Kim, M.-D.; Seo, J.-H. Enhanced production of ϵ-caprolactone by overexpression of NADPH-regenerating glucose 6-phosphate dehydrogenase in recombinant *Escherichia coli* harboring cyclohexanone monooxygenase gene. *Appl. Microbiol. Biotechnol.* **2007**, *76*, 329–338. [CrossRef] [PubMed]

58. Berrios-Rivera, S.J.; Bennett, G.N.; San, K.Y. Metabolic engineering of *Escherichia coli*: Increase of NADH availability by overexpressing an NAD^+-dependent formate dehydrogenase. *Metab. Eng.* **2002**, *4*, 217–229. [CrossRef]

59. Bashiri, G.; Antoney, J.; Jirgis, E.N.M.; Shah, M.V.; Ney, B.; Copp, J.; Stuteley, S.M.; Sreebhavan, S.; Palmer, B.; Middleditch, M.; et al. A revised biosynthetic pathway for the cofactor F_{420} in prokaryotes. *Nat. Commun.* **2019**, *10*, 1558. [CrossRef]

60. Purwantini, E.; Daniels, L. Purification of a novel coenzyme F_{420}-dependent glucose-6-phosphate dehydrogenase from *Mycobacterium smegmatis*. *J. Bacteriol.* **1996**, *178*, 2861–2866. [CrossRef]

61. Costa, K.C.; Wong, P.M.; Wang, T.; Lie, T.J.; Dodsworth, J.A.; Swanson, I.; Burn, J.A.; Hackett, M.; Leigh, J.A. Protein complexing in a methanogen suggests electron bifurcation and electron delivery from formate to heterodisulfide reductase. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11050–11055. [CrossRef]

62. Shuber, A.P.; Orr, E.C.; Recny, M.A.; Schendel, P.F.; May, H.D.; Schauer, N.L.; Ferry, J.G. Cloning, expression, and nucleotide sequence of the formate dehydrogenase genes from *Methanobacterium formicicum*. *J. Biol. Chem.* **1986**, *261*, 12942–12947.

63. Novotná, J.; Neužil, J.; Hoš?álek, Z. Spectrophotometric identification of 8-hydroxy-5-deazaflavin: Nadph oxidoreductase activity in *Streptomycetes* producing tetracyclines. *FEMS Microbiol. Lett.* **1989**, *59*, 241–245. [CrossRef]

64. Kumar, H.; Nguyen, Q.T.; Binda, C.; Mattevi, A.; Fraaije, M.W. Isolation and characterization of a thermostable F_{420}:NADPH oxidoreductase from *Thermobifida fusca*. *J. Biol. Chem.* **2017**, *292*, 10123–10130. [CrossRef] [PubMed]

65. Kunow, J.; Schwörer, B.; Stetter, K.O.; Thauer, R.K. A F_{420}-dependent NADP reductase in the extremely thermophilic sulfate-reducing *Archaeoglobus fulgidus*. *Arch. Microbiol.* **1993**, *160*, 199–205.

66. Eker, A.P.M.; Hessels, J.K.C.; Meerwaldt, R. Characterization of an 8-hydroxy-5-deazaflavin: NADPH oxidoreductase from *Streptomyces griseus*. *Biochim. Biophys. Acta* **1989**, *990*, 80–86. [CrossRef]

67. Seelbach, K.; Riebel, B.; Hummel, W.; Kula, M.R.; Tishkov, V.I.; Egorov, A.M.; Wandrey, C.; Kragl, U. A novel, efficient regenerating method of NADPH using a new formate dehydrogenase. *Tetrahedron Lett.* **1996**, *37*, 1377–1380. [CrossRef]

68. Alex, L.A.; Reeve, J.N.; Ormejohnson, W.H.; Walsh, C.T. Cloning, sequence determination, and expression of the genes encoding the subunits of the nickel-containing 8-hydroxy-5-deazaflavin reducing hydrogenase from *Methanobacterium thermoautotrophicum* delta-H. *Biochemistry* **1990**, *29*, 7237–7244. [CrossRef]

69. Tersteegen, A.; Hedderich, R. *Methanobacterium thermoautotrophicum* encodes two multisubunit membrane-bound NiFe hydrogenases - transcription of the operons and sequence analysis of the deduced proteins. *Eur. J. Biochem.* **1999**, *264*, 930–943. [CrossRef]

70. Hocking, W.P.; Stokke, R.; Roalkvam, I.; Steen, I.H. Identification of key components in the energy metabolism of the hyperthermophilic sulfate-reducing archaeon *archaeoglobus fulgidus* by transcriptome analyses. *Front. Microbiol.* **2014**, *5*, 20. [CrossRef]

71. Vitt, S.; Ma, K.; Warkentin, E.; Moll, J.; Pierik, A.J.; Shima, S.; Ermler, U. The F-420-reducing NiFe -hydrogenase complex from *methanothermobacter marburgensis*, the first x-ray structure of a group 3 family member. *J. Mol. Biol.* **2014**, *426*, 2813–2826. [CrossRef]

72. Bashiri, G.; Rehan, A.M.; Greenwood, D.R.; Dickson, J.M.; Baker, E.N. Metabolic engineering of cofactor F_{420} production in *Mycobacterium smegmatis*. *PLoS ONE* **2010**, *5*, e15803. [CrossRef]

73. Isabelle, D.; Simpson, D.R.; Daniels, L. Large-scale production of coenzyme F_{420}-5,6 by using *Mycobacterium smegmatis*. *Appl. Environ. Microbiol.* **2002**, *68*, 5750–5755. [CrossRef] [PubMed]

74. Eker, A.P.M.; Pol, A.; van der Meyden, P.; Vogels, G.D. Purification and properties of 8-hydroxy-5-deazaflavin derivatives from *Streptomyces griseus*. *FEMS Microbiol. Lett.* **1980**, *8*, 161–165. [CrossRef]

75. Grochowski, L.L.; Xu, H.M.; White, R.H. Identification and characterization of the 2-phospho-L-lactate guanylyltransferase involved in coenzyme F-420 biosynthesis. *Biochemistry* **2008**, *47*, 3033–3037. [CrossRef] [PubMed]

76. Braga, D.; Lasta, D.; Hasan, M.; Guo, H.; Leichnitz, D.; Uzum, Z.; Richter, I.; Schalk, F.; Beemelmanns, C.; Hertweck, C.; et al. Metabolic pathway rerouting in *Paraburkholderia rhizoxinica* evolved long-overlooked derivatives of coenzyme F_{420}. *ACS Chem. Biol.* **2019**, 2088–2094. [CrossRef]

77. Hossain, M.S.; Le, C.Q.; Joseph, E.; Nguyen, T.Q.; Johnson-Winters, K.; Foss, F.W. Convenient synthesis of deazaflavin cofactor F_o and its activity in F_{420}-dependent NADP reductase. *Organ. Biomol. Chem.* **2015**, *13*, 5082–5085. [CrossRef]

78. Alberty, R.A. Calculation of standard transformed formation properties of biochemical reactants and standard apparent reduction potentials of half reactions. *Arch. Biochem. Biophys.* **1998**, *358*, 25–39. [CrossRef]

79. Alberty, R.A. Calculation of standard transformed gibbs energies and standard transformed enthalpies of biochemical reactants. *Arch. Biochem. Biophys.* **1998**, *353*, 116–130. [CrossRef]

80. Benedict, M.N.; Gonnerman, M.C.; Metcalf, W.W.; Price, N.D. Genome-scale metabolic reconstruction and hypothesis testing in the methanogenic archaeon *Methanosarcina acetivorans* C2A. *J. Bacteriol.* **2012**, *194*, 855–865. [CrossRef]

81. Jankowski, M.D.; Henry, C.S.; Broadbelt, L.J.; Hatzimanikatis, V. Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys. J.* **2008**, *95*, 1487–1499. [CrossRef]

82. Henry, C.S.; DeJongh, M.; Best, A.A.; Frybarger, P.M.; Linsay, B.; Stevens, R.L. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* **2010**, *28*, 977–982. [CrossRef]

83. Nazem-Bokaee, H.; Gopalakrishnan, S.; Ferry, J.G.; Wood, T.K.; Maranas, C.D. Assessing methanotrophy and carbon fixation for biofuel production by *Methanosarcina acetivorans*. *Microb. Cell Fact.* **2016**, *15*, 10. [CrossRef] [PubMed]

84. Caspi, R.; Billington, R.; Fulcher, C.A.; Keseler, I.M.; Kothari, A.; Krummenacker, M.; Latendresse, M.; Midford, P.E.; Ong, Q.; Ong, W.K.; et al. The metacyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* **2018**, *46*, D633–D639. [CrossRef] [PubMed]

A Machine Learning Approach for Efficient Selection of Enzyme Concentrations and Its Application for Flux Optimization

Anamya Ajjolli Nagaraja [1,2,3,4]**, Philippe Charton** [2,3,4]**, Xavier F. Cadet** [5]**, Nicolas Fontaine** [5]**, Mathieu Delsaut** [1]**, Birgit Wiltschi** [6]**, Alena Voit** [6]**, Bernard Offmann** [7]**, Cedric Damour** [1]**, Brigitte Grondin-Perez** [1] **and Frederic Cadet** [2,3,4,]*

[1]   Laboratory of Energy, Electronics and Processes (LE2P-EnergyLab), EA 4079, Faculty of Sciences and Technology, University of La Reunion, 97444 St Denis CEDEX, France; anamya.ajjolli-nagaraja@univ-reunion.fr (A.A.N.); mathieu.delsaut@univ-reunion.fr (M.D.); cedric.damour@univ-reunion.fr (C.D.); Brigitte.Grondin@univ-reunion.fr (B.G.-P.)

[2]   University of Paris, BIGR—Biologie Intégrée du Globule Rouge, Inserm, UMR_S1134, F-75015 Paris, France; philippe.charton@univ-reunion.fr

[3]   Laboratory of Excellence GR-Ex, Boulevard du Montparnasse, F-75015 Paris, France

[4]   DSIMB—Dynamics of Structures and Interactions of Biological Macromolecules, UMR_S1134, BIGR, Inserm, Faculty of Sciences and Technology, University of La Reunion, F-97715 Saint-Denis, France

[5]   PEACCEL—Protein Engineering Accelerator, 6 Square Albin Cachot, box 42, 75013 Paris, France; xavier.cadet.fjf@gmail.com (X.F.C.); nicolas.fontaine@peaccel.com (N.F.)

[6]   ACIB—Austrian Centre of Industrial Biotechnology, Synthetic Biology Group, Petersgasse 14, 8010 Graz, Austria; birgit.wiltschi@acib.at (B.W.); alenavoit@acib.at (A.V.)

[7]   Unité Fonctionnalité et Ingénierie des Protéines (UFIP), UFR Sciences et Techniques, Université de Nantes, UMR 6286 CNRS, 2, Chemin de la Houssinière, 03 44322 Nantes CEDEX, France; bernard.offmann@univ-nantes.fr

*   Correspondence: frederic.cadet.run@gmail.com; Tel.: +33-69-591-8108

Abstract: The metabolic engineering of pathways has been used extensively to produce molecules of interest on an industrial scale. Methods like gene regulation or substrate channeling helped to improve the desired product yield. Cell-free systems are used to overcome the weaknesses of engineered strains. One of the challenges in a cell-free system is selecting the optimized enzyme concentration for optimal yield. Here, a machine learning approach is used to select the enzyme concentration for the upper part of glycolysis. The artificial neural network approach (ANN) is known to be inefficient in extrapolating predictions outside the box: high predicted values will bump into a sort of "glass ceiling". In order to explore this "glass ceiling" space, we developed a new methodology named glass ceiling ANN (GC-ANN). Principal component analysis (PCA) and data classification methods are used to derive a rule for a high flux, and ANN to predict the flux through the pathway using the input data of 121 balances of four enzymes in the upper part of glycolysis. The outcomes of this study are i. in silico selection of optimum enzyme concentrations for a maximum flux through the pathway and ii. experimental in vitro validation of the "out-of-the-box" fluxes predicted using this new approach. Surprisingly, flux improvements of up to 63% were obtained. Gratifyingly, these improvements are coupled with a cost decrease of up to 25% for the assay.

Keywords: machine learning; flux optimization; artificial neural network; synthetic biology; glycolysis; metabolic pathways optimization; cell-free systems

1. Introduction

Many chemical molecules like peptides, organic acids, etc., are synthesized by different methods such as chemical reactions [1–5] and fermentation process for their application in everyday life. Due to the depletion of non-renewable resources, synthesis of these molecules through a biological system is essential on an industrial scale [6,7]. For decades, scientists have been successful in producing different chemical molecules through microbial fermentation by optimizing the process [7–10]. The costs of microbial fermentation are low, for instance, in comparison to mammalian cell cultures. Microbial systems are easily scalable, use inexpensive synthetic media and have lower batch-to-batch variability [11]. However, microbial systems such as *Escherichia coli* or yeasts have no or only limited capacity for post-translational modifications. Microbial biosynthesis may show low productivity and the coproduction of by-products is possible, which make product recovery complex and protracted [12]. With the advancement of science and technology, there is a continuous effort to improve productivity through novel techniques like gene regulation, which helps to channel the pathway in particular directions, substrate channeling where reactants are directed to the active site of enzymes [13,14], quorum sensing [15], enzyme engineering, etc. However, even after numerous studies, synthesizing some molecules on an industrial scale through microbial fermentation is not cost-effective.

Nobel laureate Eduard Buchner laid the foundation for the cell-free system (CFS) of biomolecule production by converting sugar into ethanol in 1897. It has been successfully used in the synthesis of many products like bio-hydrogen [16,17], bio-ethanol [18,19], antibodies [20], vaccines [21], proteins [22], etc. The CFS is classified into two broad categories: i) cell-extract based: in which the host cells are lysed [23,24] and ii) purified-enzyme based: a mixture of purified enzymes and cofactors are in the system [25]. The CFS has high toxic tolerance, rapid development timeline, easy incorporation of unnatural amino acids and easy purification of the product. The disadvantages of CFS include poor scalability, and post-translation modification of proteins is challenging [22]. The selection of enzymes is crucial in metabolic engineering since low performing enzymes result in poor titer and yield. Homology based methodologies like Selenzyme [26] have been developed to select better performing enzymes. One of the main challenges of purified enzyme-based CFS is the selection of optimum enzyme concentrations for maximum product formation. The experimental selection of optimum enzyme concentrations is expensive and tedious.

Researchers became more interested in the mathematical modeling of biological systems due to the availability of data from omics studies [27]. The modeling helps to organize the system information, to simulate and hence optimize the experiment and to understand system characteristics. Out of many different kinds of modeling methods, constraints-based and statics-based models, as well as kinetics-based or dynamic models have been used extensively to study metabolic pathways. The constraints-based methods such as the flux balance analysis [28] depend on physicochemical constraints like mass and energy balance [29]. However, the constraint-based method does not provide information about the concentration of metabolites. Kinetic modeling depends on the kinetic parameters of the enzymes involved in the pathway and provides information about their concentrations [30]. Kinetic modeling of pathways helps to better understand their behavior and replicate the system. Since the kinetic parameters are essential for this kind of modeling, it is not always easy to replicate the system. Finding the kinetic parameters is expensive, tedious [31], and some parameters are difficult to estimate experimentally [32]. For example, phosphofructokinase requires more than ten parameters to model [33]. Hence, the development of a computational method for selecting optimum enzyme concentrations without detailed knowledge of their kinetic parameters, using other existing experimental data, is helpful.

Machine learning methods help to predict the outcome based on the existing experimental data. The artificial neural network (ANN) is one such method inspired by brain architecture [34]. The neural network consists of connections between three layers: input, hidden and output layer. An activation function for the hidden layer is used to define the output. The neural network has been widely used in different fields of science for system identification and control, pattern recognition, medical diagnosis,

weather prediction, etc. In particular, the ANN has been used for the selection of optimized medium components in the fermentation process for producing different molecules such as lipids from *Chlorella vulgaris* [35] and Spinosyns from *Saccharopolyspora spinose* [36]. ANN was employed, for instance, for the prediction of the flux through mammalian gluconeogenesis, using the simulated data from metabolite isotopic labeling [37]. Glycolysis, one of the central carbon metabolism pathways, is not only important for organisms, but also in biotechnology for producing different biomolecules [38]. Many chemicals such as organic acids [39,40] and biofuels [41,42] have been successfully produced with high titer using engineered microorganisms including *Saccharomyces cerevisiae* or *Escherichia coli*. Glycolysis is widely studied from various perspectives. The availability of data from Fievet et al. [43] for flux prediction with different enzyme concentrations makes it a good candidate for developing a new approach to select optimum enzyme concentrations.

Previously, ANN was used to predict the flux through the upper part of glycolysis using enzyme concentrations, i.e., phosphogluco isomerase (PGI), phosphofructokinase (PFK), fructose biphosphate aldolase (FBA), and triosephosphate isomerase (TPI) as the input to the model [44]. The predicted flux has a root mean square error (RMSE) of 0.84 and an R^2 of 0.93, with 13 hidden units. Since the ANN is a training-based method, the new prediction depends on the training dataset. Since ANN is not efficient in extrapolating predictions [45,46], the new predictions will always lie in the range of the known output predictions; in other words, we could say that they will remain "in-the-box". High predicted output values will bump into a sort of "glass ceiling". Our working hypothesis was that, in reality, actual flux values could be higher than the predicted ones. So, in order to explore this "glass ceiling" space, we developed a new methodology (GC-ANN, for glass ceiling ANN) to predict the flux for the upper part of glycolysis, given enzyme concentrations using an artificial neural network. The outcomes of this study are i. in silico selection of optimum enzyme concentrations for maximum flux through the pathway and ii. experimental in vitro validation of the "out-of-the-box" flux predicted using this new approach. Initially, we expected to obtain slight improvements, i.e., improved flux values close to the highest one that we fed into the model. Surprisingly, improvements up to 63% were obtained. Moreover, these improvements are coupled with a cost decrease of up to 25% for the assay.

2. Methodology

2.1. Data for New Methodology

The data from Fievet et al. [43] were used to develop the new methodology for selecting optimum enzyme concentrations using ANN. The dataset consisted of 121 combinations of four enzymes (PGI, PFK, FBA and TPI) for the upper stage of glycolysis for a flux value of 0.74–12.9 μM/s. The total enzyme concentration was kept constant for four enzymes of 101.9 mg/L. The flux was measured as NADH consumption through G3PDH. For more details about the data, please refer to the experimental section of the research article by Fievet et al. [43].

2.2. ANN-Based Flux Prediction Workflow

The new GC-ANN methodology is explained in three steps: i.) the preparation stage: the data dimension is reduced to find the possibly correlated variable, the rule for obtaining higher flux (>12 μM/s) is derived from the data, and a neural network model is built to predict the flux using the enzyme concentrations; ii.) execution stage: new enzyme concentrations are generated using the rule obtained and the flux is predicted for the new concentration using ANN; and iii.) validation of the methodology: the new methodology of predicting flux using ANN is validated through simulation and experiment.

2.2.1. Preparation stage

Reduction of Data Dimensionality

Principal component analysis (PCA) is one of the methods for the reduction of dimensionality of the dataset [47,48]. For datasets with a high degree of freedom, PCA is very useful to find possible correlations between the variables. PCA is performed using the R (V 3.4.3; R Development Core Team (2008)) package FactoMineR [49].

Visualization of Data

Three-dimensional viewing of data could provide insight into the distribution of flux in the space. Therefore, the fluxes in the 3D space of concentrations PGI, PFK, and TPI were visualized using R statistical packages plot3D [50] and plot3Drgl [51].

Classification of Data for Higher Flux (> 12 μM/s)

Data classification is the process of categorizing data into various homogeneous groups or types based on common characteristics. Decision tree analysis is a method of data classification helping to search for possible associations within the dataset. The decision tree is a simple tree-like graph method to understand and interpret the observations. The discriminant analysis helps to discriminate between the groups of data. The classification is supported by a discriminant analysis.

The data were classified into 5 groups, i.e., flux value from 0.728–3.17, 3.17–5.6, 5.6–8.04, 8.04–10.5 and 10.5–12.9. Approximately, 40% of the data are in the final group, which consists of higher flux concentrations (greater than 10.5 μM/s). The R packages klaR [52] and rpart [53] were used for discriminant analysis and decision tree respectively. The results from the decision tree and discriminant analysis were used to derive the concentration rule for higher flux values (> 12 μM/s) through the pathway.

Neural Network Model

The artificial neural network for predicting the flux through the upper part of glycolysis is built using the data described earlier in the section "Data for new methodology". The model predicts flux as an NADH consumption through the pathway. The model is built using the R package neuralnet [54], which gives us the freedom to choose two different activation functions: logistic and tanh [44].

2.2.2. Execution Stage

Generation of New Enzyme Concentration

The new enzyme concentrations were generated between the highest (PGI = 70, PFK = 70, FBA = 86.1, TPI = 66.1 mg/L) and lowest (PGI = 1, PFK = 1, FBA = 2, TPI = 1.66 mg/L) concentrations of the data from Fievet et al. [43], with a step size of 1 mg/L using R script. The total enzyme concentration of four enzymes was kept constant at 101.9 mg/L as in Fievet et al. [43]. The newly generated concentrations were used in the additional analysis.

Flux Prediction Using ANN

Newly generated enzyme concentrations were fed to the ANN model to predict the flux. The data consisted of flux values ranging from 0.74 to 12.9 μM/s. Since ANN is not good for extrapolation, these values limit the prediction to this range. Nevertheless, it is likely that new enzyme concentrations could provide higher flux. However, ANN prediction will remain in the glass ceiling space. Hence, we decided to explore this space with squeezed flux, i.e., the flux that lies in this particular space. Thus, for our study, fluxes > 12 μM/s predicted by ANN and the concentrations that obeyed the rule derived to obtain possible higher flux values from data classification were retained.

2.2.3. Validation of Methodology

The artificial neural network-based methodology for flux prediction was validated in two steps. In the first step of validation, the available kinetic parameters from Fievet et al. [43] helped us to build the model and replicate the experimental conditions. In the second step, the methodology was experimentally validated.

Simulation of Upper Part of Glycolysis

In CellDesigner (ver4.4) [55,56], the kinetic model of the upper part of glycolysis was built using the kinetic parameters from Fievet et al. [43] and the parameters for cofactors chosen from the BRENDA [57] database (Table 1). The model was built to replicate the experimental condition with the Michaelis-Menten equation (Table 1). ATP is regenerated using the creatine kinase system. The hexokinase concentration was kept constant at 0.1 μM and flux was measured as NADH consumption, as catalyzed by 1 μM of G3PDH. The concentrations of PGI, PFK, FBA and TPI were varied according to the selected balance from Section 2.2.2. (i.e., with concentrations that provide a flux ≥ 12 μM/s as predicted by the ANN model). The concentrations were converted from mg/L to μM using the molecular weight as suggested by Fievet et al.

The model was simulated for 120 s using COPASI [58] to measure NADH consumption. The slope of NADH decay between 60 and 120 s was estimated as flux through the pathway 182 enzyme balances yielding flux ≥ 15 μM/s from simulation using an in silico model were selected as the potential higher flux balances.

Experimental Validation

The upper part of glycolysis was reconstructed as described in Fievet et al. [43] (Figure 1). The in vitro system consisted of varied concentrations of PGI, PFK, FBA and TPI. The HK and G3PDH were kept constant and creatine kinases were used to regenerate ATP in the system. The NADH decay was measured as flux through the pathway. The slope of the linear NADH decay was used to calculate the flux in μM/s.

2.2.4. The Workflow of the Proposed Methodology

Based on the data listed in Fievet et al. [43], the ANN model was built to predict the flux using enzyme balances, and the rule for enzyme balance for higher flux was obtained by data classification. The fluxes for newly generated enzyme balances were predicted using the ANN model. The balances with a flux value > 12 μM/s (balances from the glass-ceiling) and the balances obeying the derived rule for higher flux were selected as potential higher flux balances. These selected balances were validated using the kinetic model and by experiments. The methodology that followed for exploring the glass-ceiling of ANN (GC-ANN) is represented diagrammatically in Figure 2.

Table 1. The kinetic equations and parameters used to build the kinetic model of the upper part of glycolysis. Glu: glucose; G6P: glucose-6-phosphate; F6P: fructose-6-phosphate; FBP: fructose bisphosphate; DHAP: dihydroxyacetone phosphate.

| Reaction catalyzed by	Kinetic Equation	Kinetic Parameters
Hexokinase (HK)	$$v = \frac{kcat_{HK} \times HK \times Glu \times ATP}{(Glu + KmGlucose) \times (ATP + KmATP)}$$	$kcat_{HK} = 72\ s^{-1}$; $Km_{Glucose} = 120\ \mu M$; $Km_{ATP} = 100\ \mu M$
Glucose-6-phosphate Isomerase (PGI)	$$v = \frac{\left(kcat_{PGI_f} \times PGI \times \left(\dfrac{G6P}{Kmg6p}\right) - kcat_{PGI_R} \times PGI \times \left(\dfrac{F6P}{KeqPGI \times Kmf6p}\right)\right)}{\left(1 + \dfrac{G6P}{Kmg6p} + \dfrac{F6P}{Kmf6p}\right)}$$	$kcat_{PGIF} = 1410\ s^{-1}$; $kcat_{PGIR} = 3720\ s^{-1}$; $Kmg6p = 1650\ \mu M$; $Kmf6p = 4100\ \mu M$; $KeqPGI = 31$
Phosphofructokinase (PFK)	$$v = \frac{kcat_{PFK} \times PFK \times F6P^{nH} \times ATP}{\left((Kmf6p^{nH} + F6P^{nH}) \times (Km_{atp} + ATP)\right)}$$	$kcat = 41.7\ s^{-1}$; $Km\_F6P = 33\ \mu M$; $nH = 1.1$; $Kmatp = 120\ \mu M$
Aldolase (ALD)	$$v = \frac{\left(kcat_{ALDF} \times FBA \times \left(\dfrac{FBP}{Km_{FrucBPhosp}}\right) - kcat_{ALDR} \times FBA \times \left(\dfrac{glyc3pho \times DHAP}{Km_{gap} \times Km_{dhap}}\right)\right)}{\left(1 + \dfrac{FBP}{Km_{FrucBPhosp}} + \dfrac{glyc3pho}{Km_{gap}} + \dfrac{DHAP}{Km_{dhap}} + \dfrac{FBP \times glyc3pho}{(Km_{FrucBPhosp} \times Ki_{g3p})} + \dfrac{glyc3pho \times DHAP}{(Km_{gap} \times Km_{dhap})}\right)}$$	$kcat_{ALDF} = 7.59\ s^{-1}$; $kcat_{ALDR} = 720\ s^{-1}$; $Km_{FrucBPhosp} = 12\ \mu M$; $Km_{gap} = 2000\ \mu M$; $Km_{dhap} = 2400\ \mu M$; $Ki_{g3p} = 10,000\ \mu M$;
Triose-phosphate Isomerase (TPI)	$$V = \frac{kcat_{TPI} \times TPI \times glyc3pho}{Km_{gap} + glyc3pho}$$	$kcat_{TPI} = 6680\ s^{-1}$; $Km_{gap} = 2380\ \mu M$
Glycerol-3-phosphate dehydrogenase (G3PDH)	$$v = \frac{kcat_{G3PDH} \times G3PDH \times \dfrac{DHAP}{Km_{DHAP}} \times \dfrac{NADH}{Km_{NADH}}}{\left(1 + \dfrac{DHAP}{Km_{DHAP}} + \dfrac{g3p}{Km_{G3P}}\right) \times \left(1 + \dfrac{NADH}{Km_{NADH}} + \dfrac{NAD}{Km_{NAD}}\right)}$$	$kcat_{G3PDH} = 189.1\ s^{-1}$; $Km_{DHAP} = 75\ \mu M$; $Km_{G3P} = 909\ \mu M$; $Km_{NADH} = 22\ \mu M$; $Km_{NAD} = 83\ \mu M$
Creatine kinase (CK)	$$v = \frac{kcat_{CK} \times CK \times phosphocreatine \times ADP}{\left(\left(1 + \dfrac{phosphocreatine}{Km_{PhosphoCrea}} + \dfrac{Creatine}{Km_{Creatine}}\right) \times \left(1 + \dfrac{ADP}{Km_{ADP}} + \dfrac{ATP}{Km_{ATP}}\right)\right)}$$	$kcat_{CK} = 148\ s^{-1}$; $Km_{PhosphoCrea} = 5000\ \mu M$; $Km_{Creatine} = 16,000\ \mu M$; $Km_{ADP} = 800\ \mu M$; $Km_{ATP} = 500\ \mu M$
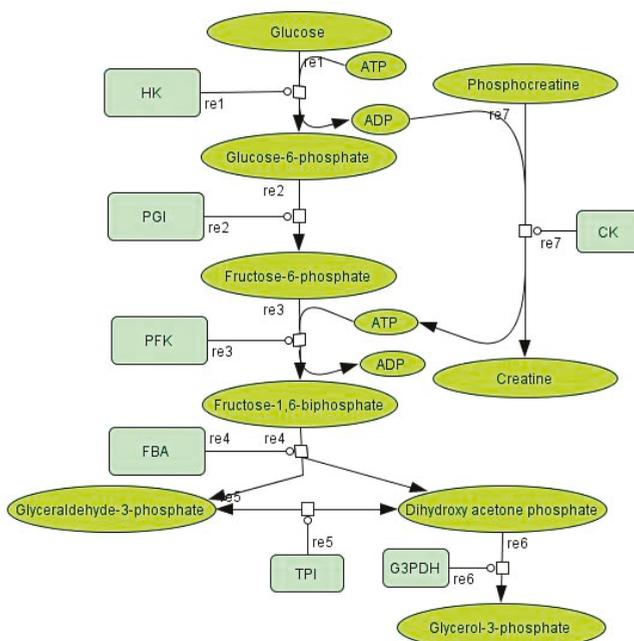
Figure 1. CellDesigner diagram for the upper part of glycolysis, which replicates the experimental conditions described by Fievet et al. [41]. HK: hexokinase, PGI: glucose 6- phosphate isomerase, PFK: phosphofructokinase, FBA: aldolase, TPI: triose-phosphate Isomerase, G3PDH: glycerol-3-phosphate dehydrogenase, CK: creatine kinase, re: reaction.
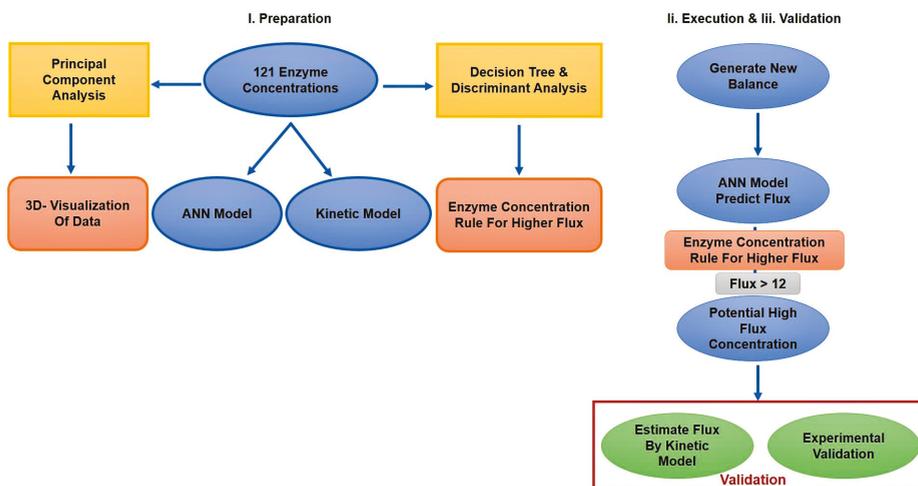

Figure 2. The methodology followed to obtain the new flux values from the generated enzyme concentration.

3. Application and Results

3.1. Preparation

3.1.1. Data Dimension Reduction

In our study, PCA did not provide much information regarding the data. The total four-enzyme concentration was constant in the system, which reduced the degree of freedom to limit the enzyme concentrations to three. If the total enzyme concentration is not constant or the dataset presents a high degree of freedom, PCA will be more useful for obtaining uncorrelated variables: this is why we mentioned PCA as a useful tool in the framework of this methodology.

3.1.2. Visualization of Data

After the PCA, data was visualized in 3D (Figure 3). We could observe on the plot that the higher flux (red dots) was quite distinct. This is a good indication that a quantitative method could be applied and should provide good results. Indeed, this is verified in the section "Flux prediction using ANN" (Figure 3). In this methodology, we were exploring the space around those higher flux concentrations to obtain new concentrations of PGI, PFK and TPI.
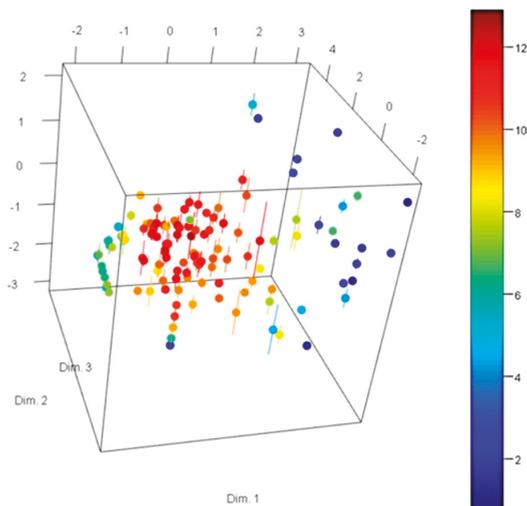

Figure 3. Three-dimensional visualization of Fievet et al. [43] enzyme balances after PCA (Dim1: 43.55%, Dim2: 23.78% and Dim 3: 17.56%). The change from blue to red indicates the gradient from low to high fluxes, respectively. Standard deviation of experimental flux is represented on the third-dimension.

3.1.3. Enzyme Concentration Rule

Decision tree analysis was performed using the R package rpart by dividing the data into five groups; this provides the best compromise on the gain in inter-class inertia. The five groups were determined using kmeans clustering.

Figure 4 represents the classification of data where the percentage of data belongs to the branch of tree and fraction represents the distribution into different groups. For example, 89% of the data had FBA concentration > 11 and is distributed in five groups as a fraction of 0.01, 0.09, 0.17, 0.29 and 0.44 (Figure 4, node 3).
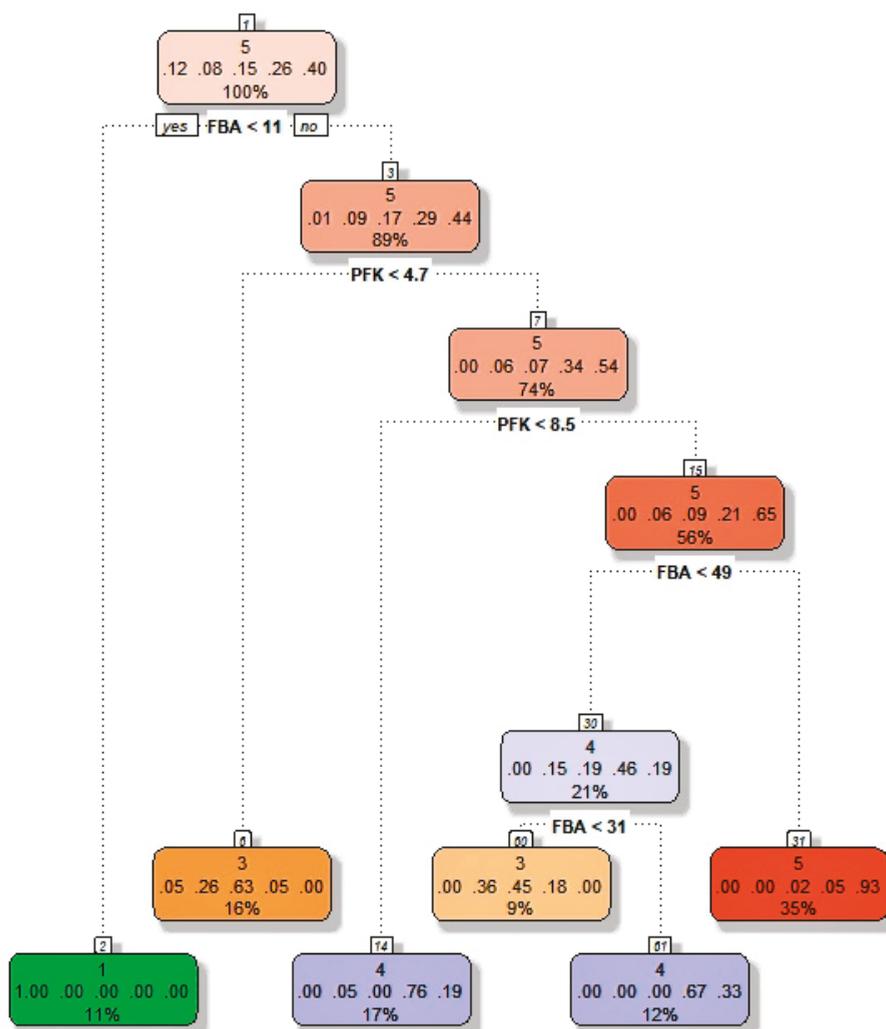
Figure 4. Decision tree analysis for Fievet et al. [43] data to obtain the rule for higher flux (≥ 12 µM/s). The data is classified into 5 groups (i.e., flux value from (0.728–3.17), (3.17–5.6), (5.6–8.04), (8.04–10.5) and (10.5–12.9).

Among the different methods of discriminant analysis studied, rpart performed the best with an approximate error rate of 0.1. The different methods studied were LDA (linear discriminant analysis), QDA (quadratic discriminant analysis), SKNN (simple k nearest neighbors), RDA (regularized discriminant analysis) and naïve Bayes (under R package). For the SKNN method, the error rate was low but it led to an over-classification (data not shown). Figure 5 represents the discriminant analysis for the classification of data from Fievet et al. [43] using the rpart [53] method from R.
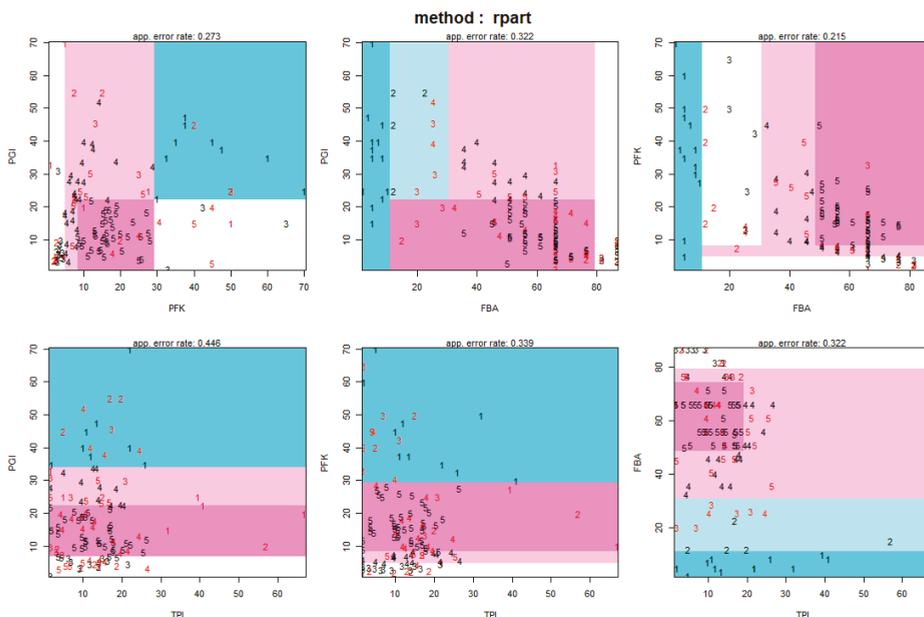
Figure 5. Discriminant analysis for the classification of data from Fievet et al. [43] using the rpart50 method from R. Color code according to the feature space of data, where group 1 (flux: 0.728–3.17 µM/s) is shown in light blue, group 2 (flux: 3.17–5.6 µM/s) in dark blue, group 3 (flux: 5.6–8.04 µM/s) in white, group 4 (flux: 8.04–10.5 µM/s) in light pink and group 5 (flux: 10.5–12.9 µM/s) in dark pink. Numbers in black represent the data classified to the same group, and in red represent data misclassified into the other groups.

After using the decision tree (Figure 4) and discriminant analysis (Figure 5), the following rule was derived to obtain a flux ≥ 12 µM/s:

PGI < 11; 10 < PFK < 16; TPI < 18; 59 >FBA (mg/L), which corresponds to PGI < 15.07 U/mL; 0.7 U/mL < PFK < 1.12 U/mL; TPI < 264.42 U/mL; 2.48 U/mL > FBA.

The conversion from mg/L to U/mL is given in Methods S1 in Supplementary Materials. The derived rule is applied for the selection of the best concentrations of the enzymes PFK, PGI, TPI, and FBA to obtain a high flux through the pathway.

3.1.4. Neural Network Model

ANN is a training-based method, the structure of the neural network needs to be chosen carefully since it depends on the number of inputs, sampling in the training dataset and the outputs. The structure was determined based on our previous study [44]. The neuralnet package from R statistical tool with the logistic activation function was used. It has 13 hidden units in a single layer. The ANN model used has an RMSE value of 0.84 and an R^2 value of 0.93, using leave-one-out cross-validation [44].

3.2. Execution

3.2.1. Generation of New Enzyme Concentrations

The new concentrations of PFK, PGI, TPI and FBA were generated as explained in the methodology section. These new balances were used for further analysis to predict the flux.

3.2.2. Flux Prediction Using ANN

The new balances were fed into the previously built neural network to predict the flux. The ANN predicted flux from the newly generated data was visualized in 3-dimensions (Figure 6).
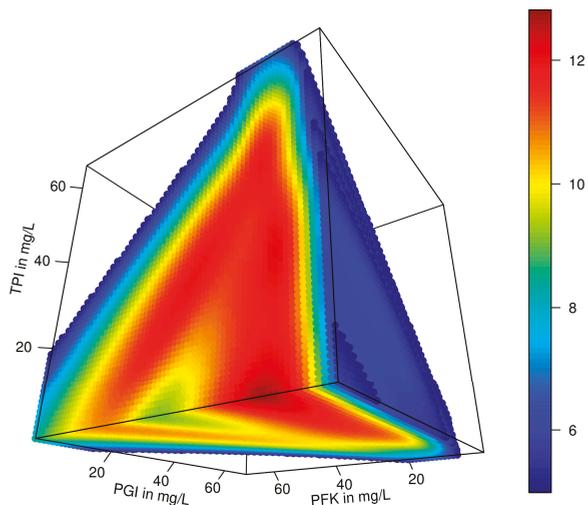

Figure 6. Three-dimensional visualization of flux predicted by an artificial neural network (ANN) for newly generated enzyme concentrations. The color gradient is from the lowest (blue) to the highest (red) predicted flux.

As expected, the new prediction remained in the box (see the maximum value of the color gradient bar in Figure 6) since ANN is a training-based method that depends on the training dataset. The high predicted values bump into the "glass ceiling". Our hypothesis was that even though they remain in the roof of the "glass ceiling", the experimental values could be higher than the predicted ones. By exploring this space, we could obtain new balances with higher flux values.

In order to explore the "glass ceiling" space, we developed this new methodology (named GC-ANN) using the artificial neural network to predict the flux through the upper part of glycolysis for given enzyme concentrations. In this study, we showed (see below in the section validation) that by careful selection of enzyme concentrations from the "glass-ceiling" space, it is possible to obtain higher flux values "out-of-the-box".

For all the enzyme concentrations generated between minimum and maximum of experimental data, only flux values above 12 µM/s predicted by neural network, and only enzyme balances (total of 335 balances, a balance being defined as a mixture of the four enzymes PGI, PFK, FBA and TPI) obeying the enzyme concentration rule were selected as potential high-flux balances.

3.3. Validation

The methodology for exploring the glass-ceiling using ANN (GC-ANN) was validated in two steps: first using the kinetic model and second, in vitro.

3.3.1. Simulation of Upper Part of Glycolysis

The kinetic model is built using CellDesigner [55,56] (Figure 1) and validated with COPASI [58] using the 121 concentrations from Fievet et al. [43]. The model has an RMSE value of 1.58 and R^2 of 0.84 in a cross-validation procedure, compared to the experimentally determined flux (Figure 7). Figure 7 proves that the kinetic model was good and could be used for the validation of the new

approach. The highest flux predicted by the kinetic model of the reconstituted upper part of glycolysis was 14.93 µM/s, where the highest experimentally observed flux was 12.9 µM/s. The flux predicted by ANN for new enzyme balances from the section "Flux prediction using ANN" was compared with the simulated flux for each enzyme (Figure 8). Figure 8 shows that the balances that were predicted with higher flux through GC-ANN were also estimated to have higher flux using the kinetic model. This validates the good quality of the kinetic model.
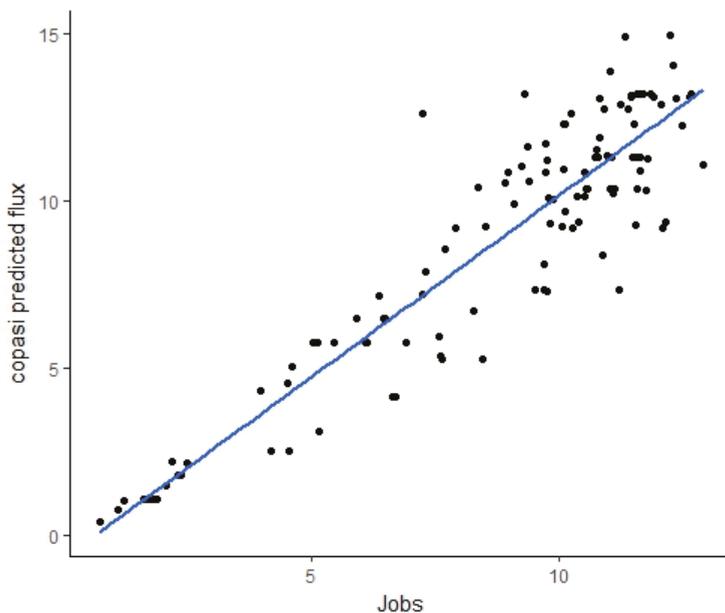

Figure 7. Relationship between experimental flux (J_{Fievet}) estimated by Fievet et al. [43] and COPASI [58] estimated by the kinetic model.

3.3.2. Experimental Validation of the Methodology

To validate this new approach to exploring the glass-ceiling (GC-ANN), the new enzyme balances generated were assayed in vitro. For the control experiment, 10 enzyme balances from previously used Fievet et al. [43] enzyme concentrations (Figure 9) were selected (Figure 10; Table S1). These selected balances have a correlation R^2 of 0.99 and an RMSE of 0.17 between the predicted flux from our kinetic model and the experimental flux assessed by Fievet et al. [43]. Figure 9 shows that balances selected for the control study are an appropriate choice. Two of these selected Fievet's balances were tested experimentally. The resulting fluxes for these two balances were 0.59 (±0.10) µM/s and 8.03 (±0.56) µM/s (see Table S2 in the Supporting Information) while Fievet et al. had determined 1.22 (±0.08) µM/s and 11.05 (±0.29) µM/s, respectively.
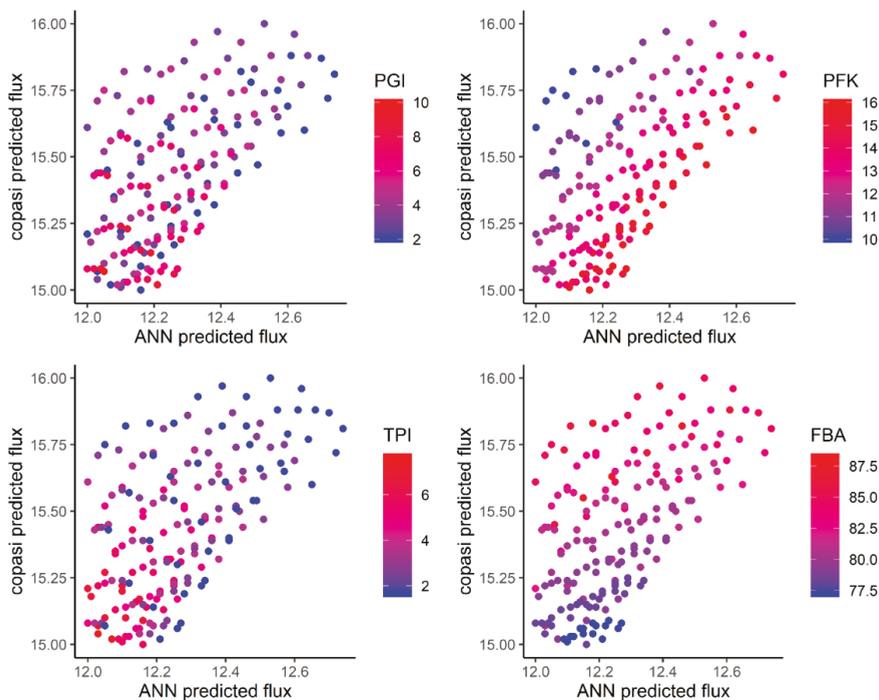
Figure 8. The relationship between flux values predicted by ANN vs COPASI for newly generated enzyme balances. The enzymes considered are: upper, left (PGI), right (PFK), lower left (TPI), right (FBA). The color gradient from blue to red represents the particular enzyme concentration from low to high, respectively.

Figure 9. Correlation between Fievet et al. [43] experimental flux and Copasi predicted flux. The balances corresponding to these flux values are selected as experimental control.

Figure 10. Comparison between glass ceiling ANN (GC-ANN) predicted flux and simulated flux. The enzyme balance corresponding to these flux values are selected for experimental validation of the methodology.
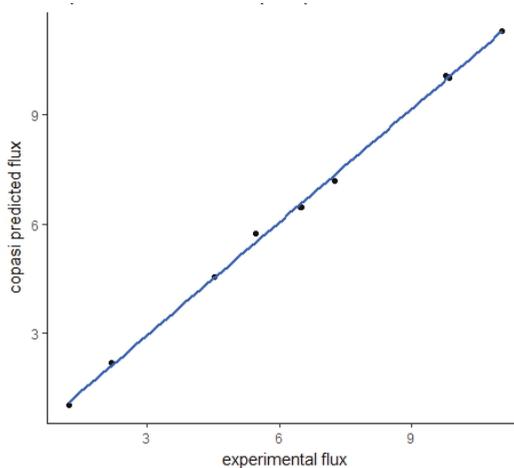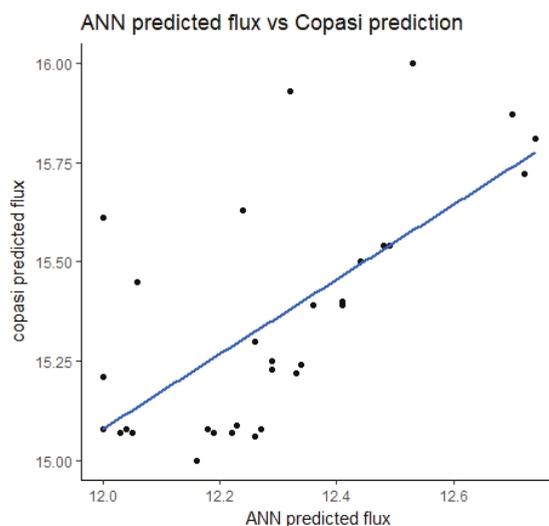
From the GC-ANN approach, 31 new balances were selected (Figure 10; Table S1) for experimental validation. The flux values associated with the selected balances had a coefficient of determination R^2 of 0.44, between GC-ANN predictions and simulated flux. This low R^2 between ANN and Copasi prediction is due to the glass-ceiling effect: the underestimation of the flux due to the inability to obtain "out-of-the-box" values for the ANN was expected.

Enzyme Assays for Measurement of Kinetic Parameters

HK activity was assessed using glucose-6-phosphate dehydrogenase (G6PDH) in a coupled reaction. The substrate glucose was converted to 6-phosphogluconate, the formation of NADPH was followed spectrophotometrically at 340 nm (Figure 11A).

We assessed the activities of PGI, PFK and FBA using a coupled NADH assay applied to the upper part of glycolysis (Figure 11B). To determine the activity of PGI, we started the assay with glucose-6-P (Figure 11B, reaction 1); for the measurement of the activities of PFK and FBA, fructose 6-P and fructose 1,6-bisP were used as the substrates (Figure 11B, reactions 2 and 3). All reactions were monitored by reading the absorbance of NADH at 340 nm and the initial rates were used to calculate the Michaelis constant K_m and the maximal velocity V_{max}. The kinetic parameters K_m for HK, PGI, PFK and FBA corresponded well to the values listed by the manufacturer (Sigma) or by the Enzyme Database Brenda (Table 2). Nevertheless, some enzymes, particularly HK and FBA, showed lower specific activity compared to the Sigma reference. The loss of activity could have occurred during delivery and/or storage of the enzymes or could be attributed to a different enzyme assay.
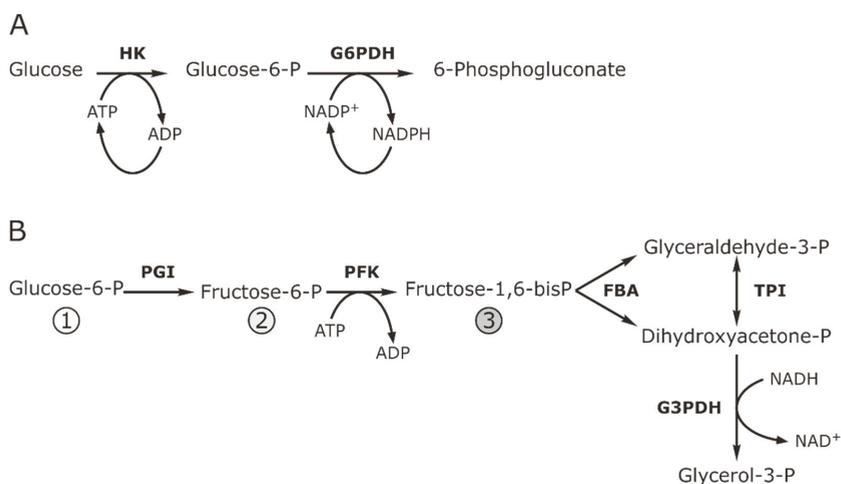
Figure 11. (**A**) Coupled HK/G6PDH assay to assess the HK activity. (**B**) Coupled NADH assay to assess the activities of PGI, PFK, and FBA. The individual reactions were started with substrates indicated by the numbers in circles.

Table 2. Summary of the kinetic parameters of HK, PGI, PFK, and FBA. The experimentally assessed values were deduced from Lineweaver-Burk and Eadie-Hofstee plots. Reference values for K_m and V_{max} from Brenda and Sigma's product data sheets are indicated, respectively. Lot No., lot number; sp. act., specific activity.

Enzyme	Lot No.	Reference Sigma sp. act. (U/mg)	This Study sp. act. (U/mg)	Reference Brenda K_m (mM)	Lineweaver-Burk * K_m (mM)	V_{max} (U/mL)	k_{cat} s^{-1}	Eadie-Hofstee * K_m (mM)	V_{max} (U/mL)	k_{cat} s^{-1}
HK	SLBT5451	472	163	0.12–0.5 [59]	0.28	225.5	299	0.30	248.7	330
PGI	SLBW8689	618	556	0.084–1.5 [60]	1.1	7409	1107	0.9	7685	1147
PFK	SLBW6641	72	73	0.023–0.15 [61]	0.13	196	166	0.11	206	175
FBA	SLBR7752V	11.5	6.4	0.00084–2 [62]	0.14	19.6	17	0.12	18.7	16
	SLBV7445	12.4	10	0.00084–2 [62]	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

* measured in this study, n.d.: not determined in this study.

Flux Determinations

The reaction mixtures for the measurements of the flux through the upper part of glycolysis were based on Fievet et al. [43] (Table 3). In contrast to Fievet et al., we based our mixtures on relative enzyme activities rather than enzyme concentrations. Calculations are explained in Method S1, in the Supplementary Materials.

Table 3. Comparison of ANN predicted flux (J_{ANN} in μM/s), simulated flux (J_{Copasi} in μM/s) and experimentally assessed flux (J_{Exp} in μM/s). The four enzymes PGI, PFK, FBA and TPI were used at the indicated concentrations for the experimental assessment of the flux with mean deviation (M.D) of triplicates.

Index	U/mL				μM/s			
	PGI	PFK	FBA	TPI	J_{ANN}	J_{Copasi}	J_{Exp}	M.D
11	2.74	0.7	3.71	24.39	12.24	15.63	15.7	2.5
12	2.74	0.7	3.62	53.77	12.06	15.45	16.3	2.7
13	2.74	0.77	3.45	97.84	12	15.21	12.1	4.2
14	2.74	0.84	3.37	112.53	12.03	15.07	16.6	0.1
15	2.74	0.91	3.58	24.39	12.7	15.87	13.9	3.9
16	2.74	0.98	3.54	24.39	12.74	15.81	18.3	1.2
17	2.74	1.05	3.50	24.39	12.72	15.72	17.1	0.2
18	2.74	1.12	3.29	83.15	12.16	15	20.1	0.3
19	4.11	0.7	3.58	53.77	12	15.61	14.4	0.1
20	4.11	0.84	3.58	24.39	12.53	16	15.8	0.2
21	4.11	1.12	3.37	39.08	12.44	15.5	20.6	0.2
22	5.48	0.77	3.58	24.39	12.32	15.93	15.4	0.2
23	5.48	1.12	3.37	24.39	12.49	15.54	16.1	2.3
24	5.48	1.12	3.33	39.08	12.36	15.39	19.3	0.6
25	6.85	1.05	3.37	24.39	12.48	15.54	18.5	0.6
26	6.85	1.12	3.33	24.39	12.41	15.4	17.8	0.1
27	6.85	1.12	3.29	39.08	12.29	15.25	16.3	0.3
28	6.85	1.12	3.24	53.77	12.18	15.08	19.7	2.5
29	8.22	1.05	3.33	24.39	12.41	15.39	17.8	1
30	8.22	1.05	3.29	39.08	12.29	15.23	19	0.6
31	8.22	1.05	3.24	53.77	12.19	15.07	21	0.6
32	8.22	1.12	3.29	24.39	12.34	15.24	15.6	3.1
33	8.22	1.12	3.24	39.08	12.23	15.09	17.8	2.2
34	9.59	0.84	3.29	68.46	12	15.08	17.1	0.7
35	9.59	1.05	3.29	24.39	12.33	15.22	17.7	1
36	9.59	1.05	3.24	39.08	12.22	15.07	18.8	1.8
37	9.59	1.12	3.24	24.39	12.27	15.08	20.4	0.6
38	10.96	0.91	3.33	24.39	12.26	15.3	15.9	0.9
39	10.96	1.05	3.24	24.39	12.26	15.06	17.9	0.8
40	12.33	0.84	3.29	39.08	12.04	15.08	15.8	0.9
41	13.7	0.84	3.29	24.39	12.05	15.07	13.6	2.4

Out of 41 selected balances, 31 newly predicted enzyme concentrations were tested experimentally to estimate flux. All 31 new enzyme balances experimentally tested were estimated with flux values greater than 12 μM/s (Table 3). Table 3 shows that 28 out of 31, i.e., 90.3%, had a value above 15.0 μM/s, as expected according to the kinetic model. Moreover, 31 out of 31, i.e., 100%, had a value above 12.0 μM/s, as expected according to our methodology.

3.4. Application: Selection of Cost-Efficient Enzyme Balances

For industrial-scale production, the selection of best enzyme concentrations in terms of cost is essential. Therefore, we estimated the cost per μM of NADH consumed per second for all the enzyme balances generated (Figure 12) and for those selected balances from ANN prediction that obey the enzyme concentration rule (flux greater than 12 μM/s), i.e., 335 balances from the section "Flux prediction using ANN" (Figure 13). The calculations were described in Method S2 in the Supplementary Materials. The cost calculation for each reaction observed in the selection of enzymes could help to reduce cost. Figures 12 and 13 show the variation in cost according to each balance and its flux and allow the selection of balances with higher flux at low cost.
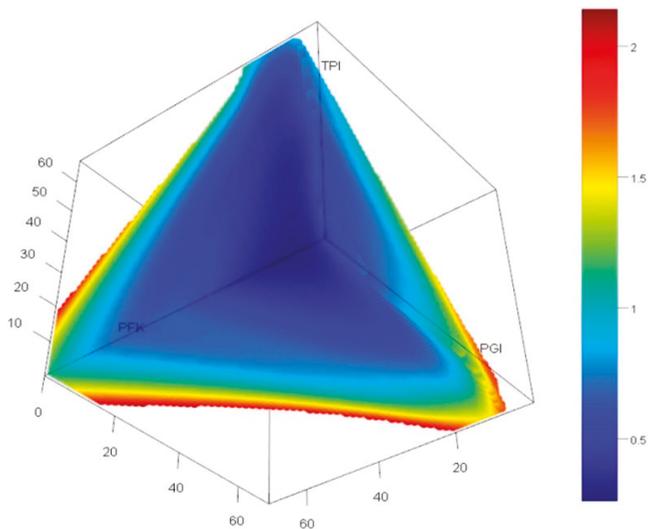
Figure 12. 3D-representation of cost estimated for all the enzyme concentrations generated. The color gradient is according to the cost required for each balance: blue is the lowest and red is the highest cost for a selected balance of the four enzymes PGI, PFK, FBA and TPI.

Figure 13. 3D-representation of the cost estimated for the enzyme concentration that obeys the rule obtained for higher flux values. The color gradient is according to the cost required for each balance, blue is the lowest and red is the highest cost for a selected balance of the four enzymes PGI, PFK, FBA and TPI.
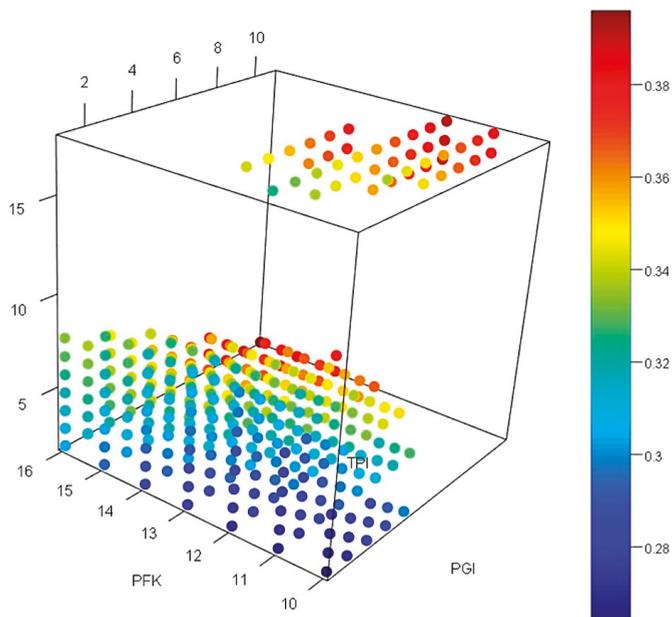
As an example: the enzyme balance (in mg/L) with PGI = 2, PFK = 12, FBA = 81.24 and TPI = 4.66 (index 13 in Table S6 of the Supporting Information) could give a flux of 12.1 μM/s with a cost of 3.79 EUR.

4. Discussion

Traditionally, chemical molecules are synthesized by the chemical reaction of petroleum-based products. Due to the depletion of petroleum products, in-vivo biosynthesis has gained a lot of attention. Limitations of the cellular production system, such as low productivity, by-product formation, and low host cell tolerance to toxins moved the focus towards development of cell-free systems. Compared to cell systems, cell-free systems have high productivity and high toxin tolerance [22]. The selection of optimal enzyme concentrations for maximal productivity is a crucial step for industrial scale, cell-free production of biomolecules. The modeling of metabolic pathways helps to study and predict the behavior of the biological system. Constraint-based methods facilitate the understanding of the system but do not provide information about the concentration of the individual metabolites. In contrast, kinetic models provide information about individual metabolite concentrations but require kinetic parameters of enzymes, which are tedious and expensive to determine [32]. Design of experiment (DOE) is a systematic approach to optimize the conditions for biomolecule production in the field of biotechnology [63]. In DOE, multiple variables are studied to find the correlation between the variables and the final outcome. The main objective of DOE is to reduce the number of experiments, time and cost; our study has the same objective. The benefit of GC-ANN is that the objective optimum can be "out-of-the-box" but will nevertheless be found without additional experiments.

4.1. GC-ANN Approach Could be Used to Predict "Out-of-the-Box" Values

In this study, a new methodology, GC-ANN, to select the optimum enzyme balances for industrial biotechnology is devised. This approach aims to see beyond the "glass ceiling", using an artificial neural network and different statistical methods like PCA and data classification. The method was designed and validated for the upper part of glycolysis but could be applied to any other natural or reconstituted biosynthesis pathway.

The workflow of the methodology used in the upper part of glycolysis is summarized in Figure 2. In the first step, for selecting the optimum concentrations of the four relevant enzymes PGI, PFK, FBA and TPI, a rule was devised for high flux values (supported by Figures 3–5). We generated all possible balances using a step of 1 mg/L in terms of variation for each enzyme concentration. The balances newly generated in the present study have higher and lower limits than those in Fievet et al. [43]. These new enzyme balances were used to predict the flux through the upper glycolysis using ANN, and the predicted fluxes were depicted in 3D representation (Figure 6); we observed a zone (Figure 6, brown zone) with predicted flux > 12 μM/s. To explore this space in order to obtain even higher fluxes, the high-flux-rule was applied, i.e., 10 < PFK < 16; PGI < 11; TPI < 18; 59 < FBA (in mg/L), and 335 enzyme balances were scrutinized. The main idea behind our approach is based on the fact that: *i.* ANN is known to be a good tool for predicting class and/or quantitative values inside the box (i.e., prediction close to training data), *ii.* the brown region in Figure 6 contains values that are all very close to 12 μM/s (from 12 to 12.9 μM/s) because ANN is not useful for extrapolation and new predictions remain inside the box; and *iii.* we postulate that among these flux values, in fact, some could be higher than predicted.

In the second step, to validate our hypothesis we conducted in silico and in vitro experiments.

4.1.1. In-Silico Validation

Due to the availability of kinetic parameters, to avoid unnecessary expenses linked to in vitro assays:

First, we built a kinetic model. Figure 7 shows good agreement (R^2 = 0.84) between the fluxes predicted by the kinetic model and all the flux values experimentally assessed by Fievet et al. [41]. Then, we selected 10 balances associated with experimental values between 0.74 and 12.9 μM/s of

Fievet's data for the benchmark study. Figure 9 shows excellent correlation with R^2 of 0.99 and an RMSE of 0.17 between the predicted flux from our kinetic model and the experimental flux assessed by Fievet et al. Taken together, these first results were a good validation of our kinetic model.

Second, we intended to validate our in vitro assay by reproducing the results obtained by Fievet et al. [43]. We decided to carry out in vitro experiments for the balances that had a good correlation between simulated and experimental flux. The experimentally determined fluxes using the balances selected from the Fievet data were lower than those previously determined by these authors (Table S3). Nevertheless, the fold-increase was comparable (approximately 9-fold, this study vs. 13-fold, Fievet et al. [43]). The deviation of the absolute flux values could be attributed to experimental settings, i.e., NADH depletion assay in cuvettes at 390 nm (Fievet et al. [43]) vs. in 96 well plates at 365 nm, in this study; or to differences in the assays performed to measure kinetic parameters of the individual enzymes.

Finally, as our kinetic model has been validated, we used it to conduct the first verification, in silico, of our hypothesis. For 31 new balances selected according to the methodology described above (Section 3.3.2), Figure 10 shows how flux values predicted by the kinetic model fit with the simulated values. All the balances selected from the brown zone (Figure 6) were indeed superior to 12.0 μM/s. Moreover, the flux should be above 15.0 μM/s. So, this is a first, in silico, validation of our hypothesis, i.e., the ANN-based approach could be used to predict "out-of-the-box" values.

At this point, we had to keep in mind that this preliminary verification was conducted because the kinetic model was possible to establish, but this step is not mandatory in the proposed methodology. Indeed, the 31 balances were chosen first, based only on the outcome of GC-ANN methodology that combines ANN and different statistical methods like PCA and data classification.

4.1.2. In Vitro Validation

The 31 new enzyme balances were assessed experimentally. Table 3 proves our hypothesis: with careful selection of enzyme concentrations from the glass ceiling, it is possible to obtain higher flux values. For the 27 best enzyme balances, the improvement of flux ranged from 20% (observed flux: 15.4, original flux: 12.9) to 63% (observed flux: 21.0, original flux: 12.9). This clearly demonstrates that exploring the predicted values, which hit the "glass ceiling" using the GC-ANN approach is a good way to select the optimum enzyme concentration.

Since artificial neural networks do not require much information regarding experimental conditions, and particularly, in our case, kinetic parameters hard to obtain, they are easy to apply in different fields of science. Our GC-ANN approach could be applied to any pathway provided the experimental data are available. Currently, we are looking for other experimental datasets to which this methodology can be applied.

4.2. The Proposed Methodology Is Cost-Efficient

From an industrial perspective, production costs per quantity of product are very important. Choosing an enzyme balance that results in maximum flux at a very low cost per given quantity of product is essential. The ANN-based methodology makes it easy to estimate the total cost. The approximate price for each reaction was calculated using the details provided by the manufacturer, such as specific activity and units of enzyme in the sample. We could calculate the approximate cost required for 1 μM of product formation per second through the pathway. This would help us to decide which is the most suitable enzyme balance for maximum flux in terms of cost minimization, which is important for industrial-scale production. For example, to obtain a flux of 12.1 μM/s, the approximate cost should be 6.28 EUR, whereas we could achieve the same flux value with a cheaper rate of 3.79 EUR (40%). Figure 12 clearly shows how costs vary. Details are provided in Table S6 and Figure S1. Among the enzyme combinations selected for the validation of our methodology, PGI = 3, PFK = 16, FBA = 80.24 and TPI = 2.66 (mg/L) had an estimated flux value of 20.6 μM/s with the lowest cost of 0.197 EUR per μM of NADH consumed per second using GC-ANN methodology for the selection of

enzyme balances (Figure S2). In contrast, the lowest price in Fievet et al. [43] with the selected balance PGI = 7, PFK = 12, FBA = 66.23 and TPI = 16.66 (mg/L) was 0.349 EUR per μM/s with an experimentally estimated flux value of 12.35 μM/s (Figure S2). This method, therefore, makes it possible to identify the production costs of 1 μM of product from 0.197 to 6.28 € in order to choose the best compromise between the cost and speed of the reaction.

Lastly and interestingly, the validated kinetic model makes it possible to generate a huge amount of data so as to feed our ANN-based model with more flux values from the newly predicted enzyme balances. This should be explored in future studies.

5. Materials and Methods

All enzymes as well as phosphocreatine, glucose-6-phosphate, fructose-6-phosphate and fructose-1,6-bisphosphate were purchased from Sigma-Aldrich (St. Louis, MO, USA). D-Glucose, ATP, NADH, and NADP were obtained from Carl Roth GmbH (Karlsruhe, Germany). Hexokinase (HK), phosphoglucoisomerase (PGI), triose-phosphate isomerase (TPI), and glucose-6-phosphate dehydrogenase (G6PDH) originated from baker's yeast; fructose biphosphate aldolase (FBA), glycerol-3-phosphate dehydrogenase (G3PDH), and creatine kinase (CK) were obtained from rabbit muscle and phosphofructokinase (PFK) originated from *Bacillus stearothermophilus*. The enzymes were obtained as lyophilized powder except for PGI and TPI, which were ammonium sulphate suspensions. Detailed information on the enzymes used is provided in Table S1 of Supplementary Materials.

5.1. Determination of Protein Concentration

Protein concentrations were determined using the Bradford protein assay [64] from Bio-Rad Laboratories (Hercules, CA). Of the protein solutions 10 μL was mixed with 200 μL of Bio-Rad Protein Assay Dye Reagent, incubated for 5 minutes at room temperature, and the absorbance was measured spectrophotometrically at 595 nm. A dilution series of 0.06–0.5 mg/mL BSA (Carl Roth GmbH) was used for calibration.

5.2. Enzyme Assays for the Determination of Kinetic Parameters

Enzyme assays were performed in 96-well UV-STAR®microplates (Greiner Bio-One GmbH, Kremsmünster, Austria) in a total volume of 100 μL at 25 °C. The reaction buffer contained 50 mM PIPES (pH 7.5), 100 mM KCl, and 5 mM magnesium acetate. The cofactors for the reactions were 1 mM ATP and 1 mM NADH or NADP.

HK activity was measured with 0.05 U HK, 2.5 U G6PDH, and glucose concentrations from 10 to 0.01 mM. PGI activity was measured with 0.02–0.01 U PGI, 1–0.5 U PFK, 0.5 U FBA, 2 U G3PDH, 5 U TPI, and glucose 6-phosphate concentrations ranging from 30 to 0.03 mM. PFK activity was measured with 0.02 U PFK, 0.5 U FBA, 2 U G3PDH, 5 U TPI, and fructose 6-phosphate concentrations from 10 to 0.01 mM. FBA activity was measured with 0.01–0.05 U FBA, 2 U G3PDH, 5 U TPI, and fructose 1,6-phosphate concentrations from 10 to 0.01 mM. All reactions were monitored by recording the absorption at a wavelength of 340 nm (molar extinction coefficient $\varepsilon_{340 \text{ nm}, 25\,°C}$ 6.22 L mmol^{-1} cm^{-1}). For calculation of the kinetic parameters V_{max}, K_m, and k_{cat} we used Lineweaver-Burk as well as Eadie-Hofstee representations.

5.3. Flux Measurements

The total reaction volume of 100 μL contained fixed concentrations of 3 mM NADH, 20 mM phosphocreatine, 1 μM CK, 0.1 μM HK, and 1 μM G3PDH. The concentrations of PGI, PFK, FBA, and TPI were varied as indicated (Section 3.3.2). The reactions were started with 1 mM ATP and 100 mM glucose. Blank reactions contained all ingredients except ATP and glucose. Each condition was measured in triplicates. The NADH decay was monitored every 3 s at 365 nm using a SynergyMxSMATBLD(+) Gen5 SW plater reader (SZABO-SCANDIC, Vienna, Austria). The slope of NADH decay was measured as the flux through the pathway (molar extinction coefficient $\varepsilon_{365 \text{ nm}, 25\,°C}$ 3.4 L mmol^{-1} cm^{-1}).

6. Conclusions

The selection of enzymes is an important step in the production of biomolecules. Methods based on homology are widely used to select the best performing enzymes. In addition, the selection of optimum enzyme balances is also crucial. Most methods use kinetic information for concentration selection via modeling. However, the determination of kinetic parameters is not always easy; therefore, developing new methodologies for selecting the optimum enzyme balances is of great interest.

In this study, we developed a new approach, GC-ANN, which uses an artificial neural network along with different statistical methods (PCA and data classification) to select enzyme balances that improve the flux as well as the costs. The selected balances might not be the balances with the highest flux, but they would be among the best. This approach allows cost-efficient selection of enzyme balances using a small existing dataset, and it opens the door for rapid optimization of cell-free systems in an industrial environment.

References

1. Borgia, J.A.; Fields, G.B. Chemical synthesis of proteins. *Trends Biotechnol.* **2000**, *18*, 243–251. [CrossRef]
2. Hojo, H. Recent progress in the chemical synthesis of proteins. *Curr. Opin. Struct. Biol.* **2014**, *26*, 16–23. [CrossRef]
3. Liu, F.; Zaykov, A.N.; Levy, J.J.; Dimarchi, R.D.; Mayer, J.P. Chemical synthesis of peptides within the insulin superfamily. *J. Pept. Sci.* **2016**, *22*, 260–270. [CrossRef] [PubMed]
4. Graf, M.; Mardirossian, M.; Nguyen, F.; Seefeldt, A.C.; Guichard, G.; Scocchi, M.; Innis, C.A.; Wilson, D.N. Proline-rich antimicrobial peptides targeting protein synthesis. *Nat. Prod. Rep.* **2017**, *34*, 702–711. [CrossRef] [PubMed]
5. Arora, K.; Program, B.; Arbor, A. *Total Synthesis of Glycosylated Proteins Alberto*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 200, pp. 165–187.
6. Zhang, Y.H.P. Renewable carbohydrates are a potential high-density hydrogen carrier. *Int. J. Hydrogen Energy* **2010**, *35*, 10334–10342. [CrossRef]

7. Yim, H.; Haselbeck, R.; Niu, W.; Pujol-Baxley, C.; Burgard, A.; Boldt, J.; Khandurina, J.; Trawick, J.D.; Osterhout, R.E.; Stephen, R.; et al. Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. *Nat. Chem. Biol.* **2011**, *7*, 445–452. [CrossRef]

8. Martínez, J.A.; Bolívar, F.; Escalante, A. Shikimic Acid Production in Escherichia coli: From Classical Metabolic Engineering Strategies to Omics Applied to Improve Its Production. *Front. Bioeng. Biotechnol.* **2015**, *3*, 1–16. [CrossRef]

9. Lee, J.W.; Na, D.; Park, J.M.; Lee, J.; Choi, S.; Lee, S.Y. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* **2012**, *8*, 536–546. [CrossRef] [PubMed]

10. Chen, X.; Wang, Y.; Dong, X.; Hu, G.; Liu, L. Engineering rTCA pathway and C4-dicarboxylate transporter for l-malic acid production. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 4041–4052. [CrossRef]

11. Stanton, D. Microbial or Mammalian? Biosilta Backs the Former Licensing E. Coli platform. Biopharma Reporter. Available online: https://www.biopharma-reporter.com/Article/2016/04/08/Microbial-or-mammalian-BioSilta-licenses-E.-Coli-tech (accessed on 20 February 2020).

12. Theisen, M.; Liao, J.C. Industrial Biotechnology: Escherichia coli as a Host. *Ind. Biotechnol.* **2016**, *1*, 149–181.

13. Zhang, Y.H.P. Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* **2011**, *29*, 715–725. [CrossRef]

14. Wheeldon, I.; Minteer, S.D.; Banta, S.; Barton, S.C.; Atanassov, P.; Sigman, M. Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **2016**, *8*, 299–309. [CrossRef]

15. Tan, S.Z.; Prather, K.L. Dynamic pathway regulation: Recent advances and methods of construction. *Curr. Opin. Chem. Biol.* **2017**, *41*, 28–35. [CrossRef]

16. Fontaine, N.; Grondin-Perez, B.; Cadet, F.; Offmann, B. Modeling of a Cell-Free Synthetic System for Biohydrogen Production. *J. Comput. Sci. Syst. Biol.* **2015**, *8*, 132–139.

17. Ye, X.; Wang, Y.; Hopkins, R.C.; Adams, M.W.W.; Evans, B.R.; Mielenz, J.R.; Zhang, Y.H.P. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. *ChemSusChem* **2009**, *2*, 149–152. [CrossRef]

18. Khattak, W.A.; Ul-Islam, M.; Ullah, M.W.; Yu, B.; Khan, S.; Park, J.K. Yeast cell-free enzyme system for bio-ethanol production at elevated temperatures. *Process. Biochem.* **2014**, *49*, 357–364. [CrossRef]

19. Zhang, Y.H.P. Production of biofuels and biochemicals by in vitro synthetic biosystems: Opportunities and challenges. *Biotechnol. Adv.* **2015**, *33*, 1467–1483. [CrossRef]

20. Huang, L.; Sheng, J.; Xu, Z.; Zhu, X.; Cai, J. Reconstitution of the peptidoglycan cytoplasmic precursor biosynthetic pathway in cell-free system and rapid screening of antisense oligonucleotides for Mur enzymes. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 1785–1794.

21. Yang, J.; Voloshin, A.; Swartz, J.R.; Velkeen, H.; Levy, R.; Michel-Reydellet, N. Rapid expression of vaccine proteins for B-cell lymphoma in a cell-free system. *Biotechnol. Bioeng.* **2005**, *89*, 503–511. [CrossRef]

22. Lu, Y. Cell-free synthetic biology: Engineering in an open world. *Synth. Syst. Biotechnol.* **2017**, *2*, 23–27. [CrossRef]

23. Schoborg, J.A.; Hodgman, C.E.; Anderson, M.J.; Jewett, M.C. Substrate replenishment and byproduct removal improve yeast cell-free protein synthesis. *Biotechnol. J.* **2014**, *9*, 630–640. [CrossRef]

24. Shrestha, P.; Holland, T.M.; Bundy, B.C. Streamlined extract preparation for Escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing. *Biotechniques* **2012**, *53*, 163–174. [CrossRef]

25. Zhang, Y.-H.P. Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: Challenges and opportunities. *Biotechnol. Bioeng.* **2009**, *105*, 663–677. [CrossRef]

26. Carbonell, P.; Wong, J.; Swainston, N.; Takano, E.; Turner, N.J.; Scrutton, N.S.; Kell, D.B.; Breitling, R.; Faulon, J.-L. Selenzyme: Enzyme selection tool for pathway design. *Bioinformatics* **2018**, *34*, 2153–2154. [CrossRef]

27. Stelling, J. Mathematical models in microbial systems biology. *Curr. Opin. Microbiol.* **2004**, *7*, 513–518. [CrossRef]

28. Orth, J.D.; Thiele, I.; Palsson, B.O. What is flux balance analysis? *Nat. Biotechnol.* **2010**, *28*, 245–248. [CrossRef]

29. Covert, M.W.; Famili, I.; Palsson, B.O. Identifying Constraints that Govern Cell Behavior: A Key to Converting Conceptual to Computational Models in Biology? *Biotechnol. Bioeng.* **2003**, *84*, 763–772. [CrossRef]

30. Smallbone, K.; Simeonidis, E.; Broomhead, D.S.; Kell, D.B. Something from nothing-Bridging the gap between constraint-based and kinetic modelling. *FEBS J.* **2007**, *274*, 5576–5585. [CrossRef]

31. Schmeier, S.; Hakenberg, J.; Klipp, E.; Leser, U.; Kowald, A. Finding Kinetic Parameters Using Text Mining. *Omi. A J. Integr. Biol.* **2004**, *8*, 131–152.

32. Bisswanger, H. Enzyme assays. *Perspect. Sci.* **2014**, *1*, 41–55. [CrossRef]

33. Teusink, B.; Passarge, J.; Reijenga, C.A.; Esgalhado, E.; van der Weijden, C.C.; Schepper, M.; Walsh, M.C.; Bakker, B.M.; van Dam, K.; Westerhoff, H.V.; et al. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* **2000**, *267*, 5313–5329. [CrossRef] [PubMed]

34. Basheer, I.A.; Hajmeer, M. Artificial neural networks: Fundamentals, computing, design, and application. *J. Microbiol. Methods.* **2000**, *43*, 3–31. [CrossRef]

35. Morowvat, M.H.; Ghasemi, Y. Medium optimization by artificial neural networks for maximizing the triglycerides-rich lipids from biomass of Chlorella vulgaris. *Int. J. Pharm. Clin. Res.* **2016**, *8*, 1414–1417.

36. Lan, Z.; Zhao, C.; Guo, W.; Guan, X.; Zhang, X. Optimization of culture medium for maximal production of spinosad using an artificial neural network-genetic algorithm modeling. *J. Mol. Microbiol. Biotechnol.* **2015**, *25*, 253–261. [CrossRef]

37. Antoniewicz, M.R.; Stephanopoulos, G.; Kelleher, J.K. Evaluation of regression models in metabolic physiology: Predicting fluxes from isotopic data without knowledge of the pathway. *Metabolomics* **2006**, *2*, 41–52. [CrossRef]

38. Liu, J.; Li, J.; Shin, H.D.; Liu, L.; Du, G.; Chen, J. Protein and metabolic engineering for the production of organic acids. *Bioresour. Technol.* **2017**, *239*, 412–421. [CrossRef]

39. Song, C.W.; Kim, D.I.; Choi, S.; Jang, J.W.; Lee, S.Y. Metabolic engineering of Escherichia coli for the production of fumaric acid. *Biotechnol. Bioeng.* **2013**, *110*, 2025–2034. [CrossRef]

40. Yang, J.; Wang, Z.; Zhu, N.; Wang, B.; Chen, T.; Zhao, X. Metabolic engineering of Escherichia coli and in silico comparing of carboxylation pathways for high succinate productivity under aerobic conditions. *Microbiol. Res.* **2014**, *169*, 432–440. [CrossRef]

41. Clomburg, J.M.; Gonzalez, R. Biofuel production in Escherichia coli: The role of metabolic engineering and synthetic biology. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 419–434. [CrossRef]

42. Yang, X.; Xu, M.; Yang, S.T. Metabolic and process engineering of Clostridium cellulovorans for biofuel production from cellulose. *Metab. Eng.* **2015**, *32*, 39–48. [CrossRef]

43. Fiévet, J.B.; Dillmann, C.; Curien, G.; de Vienne, D. Simplified modelling of metabolic pathways for flux prediction and optimization: Lessons from an in vitro reconstruction of the upper part of glycolysis. *Biochem. J.* **2006**, *396*, 317–326. [CrossRef]

44. Ajjolli Nagaraja, A.; Fontaine, N.; Delsaut, M.; Charton, P.; Damour, C.; Offmann, B.; Grondin-Perez, B.; Cadet, F. Flux prediction using artificial neural network (ANN) for the upper part of glycolysis. *PLoS ONE* **2019**, *14*, 1–15. [CrossRef]

45. Minns, A.W.; Hall, M.J. Artificial neural networks as rainfall- runoff models. *Hydrol. Sci. J.* **1996**, *41*, 399–417. [CrossRef]

46. Balabin, R.M.; Smirnov, S.V. Interpolation and extrapolation problems of multivariate regression in analytical chemistry: Benchmarking the robustness on near-infrared (NIR) spectroscopy data. *Analyst* **2012**, *137*, 1604–1610. [CrossRef]

47. Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [CrossRef]

48. Ringnér, M. What is principal component analysis? *Nat. Biotechnol.* **2008**, *26*, 303–304. [CrossRef]

49. Le, S.; Josse, J.; Husson, F. FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Softw.* **2008**, *25*, 1–18. [CrossRef]

50. Soetaert, K. plot3D: Plotting Multi-Dimensional Data, 2017. Available online: https://CRAN.R-project.org/package=plot3D (accessed on 2 March 2020).

51. Soetaert, K. plot3Drgl: Plotting Multi-Dimensional Data-Using "rgl". 2016. Available online: https://CRAN.R-project.org/package=plot3Drgl (accessed on 2 March 2020).

52. Weihs, C.; Ligges, U.; Luebke, K.; Raabe, N. klaR Analyzing German Business Cycles. In *Data Analysis and Decision Support*; Baier, D., Decker, R., Schmidt-Thieme, L., Eds.; Springer-Verlag: Berlin, Germany, 2005; pp. 335–343.

53. Therneau, T.; Atkinson, B.; Ripley, B. rpart: Recursive Partitioning and Regression Trees. *R Package* **2018**, *4*, 1–9.

54. Günther, F.; Fritsch, S. Neuralnet: Training of Neural Networks. *R. J.* **2010**, *2*, 30–38. [CrossRef]

55. Funahashi, A.; Morohashi, M.; Kitano, H.; Tanimura, N. CellDesigner: A process diagram editor for gene-regulatory and biochemical networks. *Biosilico* **2003**, *1*, 159–162. [CrossRef]

56. Funahashi, A.; Matsuoka, Y.; Jouraku, A.; Morohashi, M.; Kikuchi, N.; Kitano, H. CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks. *Proc IEEE.* **2008**, *96*, 1254–1265. [CrossRef]

57. Schomburg, I.; Chang, A.; Schomburg, D. BRENDA, enzyme data and metabolic information. *Nucleic. Acids Res.* **2002**, *30*, 47–49. [CrossRef]

58. Lee, C.; Pahle, J.; Singhal, M.; Gauges, R.; Sahle, S.; Mendes, P.; Kummer, U.; Hoops, S. COPASI–a COmplex PAthway SImulator. *Bioinformatics* **2006**, *22*, 3067–3074.

59. BRENDA-Information on EC 2.7.1.1-hexokinase. Available online: https://www.brenda-enzymes.org/enzyme.php?ecno=2.7.1.1 (accessed on 22 May 2019).

60. BRENDA-Information on EC 5.3.1.9-glucose-6-phosphate isomerase. Available online: https://www.brenda-enzymes.org/enzyme.php?ecno=5.3.1.9 (accessed on 22 May 2019).

61. BRENDA-Information on EC 2.7.1.11-6-phosphofructokinase. Available online: https://www.brenda-enzymes.org/enzyme.php?ecno=2.7.1.11 (accessed on 22 May 2019).

62. BRENDA-Information on EC 4.1.2.13-fructose-bisphosphate aldolase. Available online: https://www.brenda-enzymes.org/enzyme.php?ecno=4.1.2.13 (accessed on 22 May 2019).

63. Kumar, V.; Bhalla, A.; Rathore, A.S. Design of experiments applications in bioprocessing: Concepts and approach. *Biotechnol. Prog.* **2014**, *30*, 86–99. [CrossRef]

64. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]

Article

Immobilization of β-Galactosidases on the *Lactobacillus* Cell Surface Using the Peptidoglycan-Binding Motif LysM

Mai-Lan Pham [1], Anh-Minh Tran [1,2], Suwapat Kittibunchakul [1], Tien-Thanh Nguyen [3], Geir Mathiesen [4] and Thu-Ha Nguyen [1,*]

[1] Food Biotechnology Laboratory, Department of Food Science and Technology, BOKU-University of Natural Resources and Life Sciences, A-1190 Vienna, Austria; mailanpham.22@gmail.com (M.-L.P.); tranminh1703@gmail.com (A.-M.T.); suwapatkt@gmail.com (S.K.)

[2] Department of Biology, Faculty of Fundamental Sciences, Ho Chi Minh City University of Medicine and Pharmacy, 217 Hong Bang, Ho Chi Minh City, Vietnam

[3] School of Biotechnology and Food Technology, Hanoi University of Science and Technology, 1 Dai Co Viet, Hanoi, Vietnam; nguyenthanh.ibft@gmail.com

[4] Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), N-1432 Ås, Norway; geir.mathiesen@nmbu.no

[*] Correspondence: thu-ha.nguyen@boku.ac.at; Tel.: +43-1-47654-75215; Fax: +43-1-47654-75039

Abstract: Lysin motif (LysM) domains are found in many bacterial peptidoglycan hydrolases. They can bind non-covalently to peptidoglycan and have been employed to display heterologous proteins on the bacterial cell surface. In this study, we aimed to use a single LysM domain derived from a putative extracellular transglycosylase Lp_3014 of *Lactobacillus plantarum* WCFS1 to display two different lactobacillal β-galactosidases, the heterodimeric LacLM-type from *Lactobacillus reuteri* and the homodimeric LacZ-type from *Lactobacillus delbrueckii* subsp. *bulgaricus*, on the cell surface of different *Lactobacillus* spp. The β-galactosidases were fused with the LysM domain and the fusion proteins, LysM-LacLMLreu and LysM-LacZLbul, were successfully expressed in *Escherichia coli* and subsequently displayed on the cell surface of *L. plantarum* WCFS1. β-Galactosidase activities obtained for *L. plantarum* displaying cells were 179 and 1153 U per g dry cell weight, or the amounts of active surface-anchored β-galactosidase were 0.99 and 4.61 mg per g dry cell weight for LysM-LacLMLreu and LysM-LacZLbul, respectively. LysM-LacZLbul was also displayed on the cell surface of other *Lactobacillus* spp. including *L. delbrueckii* subsp. *bulgaricus*, *L. casei* and *L. helveticus*, however *L. plantarum* is shown to be the best among *Lactobacillus* spp. tested for surface display of fusion LysM-LacZLbul, both with respect to the immobilization yield as well as the amount of active surface-anchored enzyme. The immobilized fusion LysM-β-galactosidases are catalytically efficient and can be reused for several repeated rounds of lactose conversion. This approach, with the β-galactosidases being displayed on the cell surface of non-genetically modified food-grade organisms, shows potential for applications of these immobilized enzymes in the synthesis of prebiotic galacto-oligosaccharides.

Keywords: *Lactobacillus*; β-galactosidase; immobilization; cell surface display; LysM domains

---

1. Introduction

β-Galactosidases catalyze the hydrolysis and transgalactosylation of β-D-galactopyranosides (such as lactose) [1–3] and are found widespread in nature. They catalyze the cleavage of lactose (or related compounds) in their hydrolysis mode and are thus used in the dairy industry to remove lactose from various products. An attractive biocatalytic application is found in the transgalactosylation

potential of these enzymes, which is based on their catalytic mechanism [1,4]. β-Galactosidases can be obtained from different sources including microorganisms, plants and animals, however microbial sources of β-galactosidase are of great biotechnological interest because of easier handling, higher multiplication rates, and production yield. Recently, a number of studies have focused on the use of the genus *Lactobacillus* for the production and characterization of β-galactosidases, including the enzymes from *L. reuteri, L. acidophilus, L. helveticus, L. plantarum, L. sakei, L. pentosus, L. bulgaricus, L. fermentum, L. crispatus* [5–15]. β-Galactosidases from *Lactobacillus* species are different at molecular organization [6,8,10,12,16]. The predominant glycoside hydrolase family 2 (GH2) β-galactosidases found in lactobacilli are of the LacLM type, which are heterodimeric proteins encoded by the two overlapping genes, *lacL* and *lacM*, including *lacLM* from *L. reuteri* [16], *L. acidophilus* [6], *L. helveticus* [7], *L. pentosus* [11], *L. plantarum* [8], and *L. sakei* [10]. Di- or oligomeric GH2 β-galactosidases of the LacZ type, encoded by the single *lacZ* gene, are sometimes, but not often found in lactobacilli such as in *L. bulgaricus* [12]. Lactobacilli have been studied intensively with respect to their enzymes for various different reasons, one of which is their 'generally recognized as safe' (GRAS) status and their safe use in food applications. It is anticipated that galacto-oligosaccharides (GOS) produced by these β-galactosidases will have better selectivity for growth and metabolic activity of this bacterial genus in the gut.

An economical, sustainable and intelligent use of biocatalysts can be achieved through immobilization, where the enzyme is bound onto a suitable food-grade carrier. Efforts have been made to immobilize β-galactosidases from *L. reuteri*, a LacLM-type, and *Lactobacillus bulgaricus*, a LacZ-type, on chitin using the chitin binding domain (ChBD) of *Bacillus circulans* WL-12 chitinase A1 [17]. Cell surface display has been shown as a new strategy for enzyme immobilization, which involves the use of food-grade organism *L. plantarum* both as a cell factory for the production of enzymes useful for food applications and as the carrier for the immobilization of the over-expressed enzyme by anchoring the enzyme on the cell surface [18,19]. This enables the direct use of the microbial cells straight after the fermentation step as an immobilized biocatalysts, offering the known advantages of immobilization (reuse of enzyme, stabilization, etc.) together with a significant simplification of the production process since costly downstream processing of the cells producing the enzyme (cell disruption, protein purification, etc.) as well as the use of carrier material will not be necessary. We recently reported cell surface display of mannanolytic and chitinolytic enzymes in *L. plantarum* using two anchors from *L. plantarum*, a lipoprotein-anchor derived from the Lp_1261 protein and a cell wall anchor (cwa2) derived from the Lp_2578 protein [19]. However, this approach works less efficient with dimeric and oligomeric enzymes, such as β-galactosidases from lactobacilli, due to low secretion efficiency of target proteins. Therefore, it is of our interest to find another strategy to display lactobacillal β-galactosidases on *Lactobacillus* cell surface for use as immobilized biocatalysts for applications in lactose conversion and GOS formation processes.

There are two principally different ways of anchoring a secreted protein to the bacterial cell wall: covalently, via the sortase pathway, or non-covalently, via a protein domain that interacts strongly with cell wall components. In sortase-mediated anchoring, the secreted protein carries a C-terminal anchor containing the so-called LP × TG motif followed by a hydrophobic domain and a positively charged tail [20]. The hydrophobic domain and the charged tail keep the protein from being released to the medium, thereby allowing recognition of the LP × TG motif by a membrane-associated transpeptidase called sortase [20–22]. The sortase cleaves the peptide bond between threonine and glycine in the LP × TG motif and links the now C-terminal threonine of the surface protein to a pentaglycine in the cell wall [21–25]. One of the non-covalent cell display systems exploits so-called LysM domains, the peptidoglycan binding motifs, that are known to promote cell wall association of several natural proteins [23,26]. These domains have been used to display proteins in lactic acid bacteria (LAB) by fusing the LysM domain N- or C-terminally to the target protein [27–30]. In *L. plantarum* WCFS1 ten proteins are predicted to be displayed at the cell wall through LysM domains [31].

In this present study, we exploit a single LysM domain derived from the Lp_3014 protein in *L. plantarum* WCFS1 for external attachment of two lactobacillal β-galactosidases, a LacLM-type from *L. reuteri* and a LacZ-type from *L. bulgaricus*, on the cell surface of four *Lactobacillus* species. The immobilization of active β-galactosidases through cell-surface display can be utilized as safe and stable non-GMO food-grade biocatalysts that can be used in the production of prebiotic GOS.

2. Results

2.1. Expression of Recombinant Lactobacillal β-Galactosidases in E. coli

The overlapping *lacLM* genes from *L. reuteri* L103 and the *lacZ* gene from *L. bulgaricus* DSM20081, both encoding β-galactosidases, were fused N-terminally to the LysM motif for expression and later attachment of the hybrid proteins to the peptidoglycan layer of *Lactobacillus* spp. An 88 residue fragment of the LysM motif from the 204-residue-Lp_3014 protein of an extracellular transglycosylase of *L. plantarum* WCFS1 [31,32] was fused to two β-galactosidases for production in *E. coli*. The two hybrid sequences were then cloned into the expression vector pBAD containing an N-terminal 7 × Histidine tag for immunodetection, yielding pBAD3014LacLMLreu and pBAD3014LacZLbul (Figure 1).
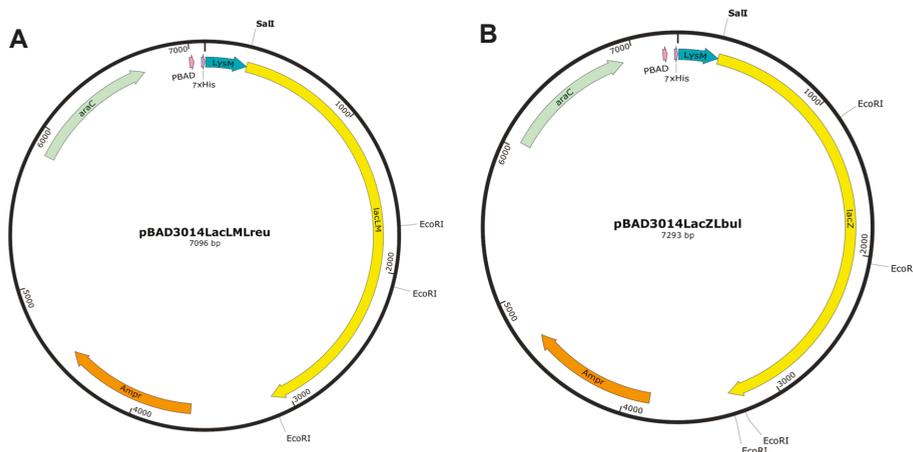
Figure 1. The expression vectors for LysM-LacLMLreu (**A**) and LysM-LacZLbul (**B**) in *E. coli*. The vectors are the derivatives of the pBAD vector (Invitrogen, Carlsbad, CA, USA) containing a 7 × His tag sequence fused to a single LysM domain from Lp_3014, *L. plantarum* WCFS1. LacLMLreu encoded by two overlapping genes *lacLM* and LacZLbul encoded by the *lacZ* gene are the β-galactosidases from *L. reuteri* and *L. delbrueckii* subsp. *bulgaricus* DSM 20081, respectively. See text for more details.

The *E. coli* strains were cultivated in Luria-Bertani (LB) medium, induced for gene expression (as described in Materials and Methods), and the SDS-PAGE and Western blot analyses of cell-free extracts (Figure 2) showed the production of the two recombinant β-galactosidases, LysM-LacLMLreu and LysM-LacZLbul. As judged by SDS-PAGE (Figure 2A), LysM-LacLMLreu shows two bands with apparent molecular masses corresponding to a large subunit (LacL) and a small subunit (LacM) at ~90 kDa and ~35 kDa. These values are in agreement with reported molecular masses of 73 and 35 kDa for these two subunits of β-galactosidase from *L. reutei* [5,16]. The increase in molecular mass of a larger subunit in LysM-LacLMLreu is due to the added His-LysM fragment (~18 kDa). On the other hand, β-galactosidase from *L. bulgaricus* was reported to be a homodimer, consisting of two identical subunits of ~115 kDa [12]. A unique band of ~130 kDa corresponding to the molecular mass of a single subunit of LacZ fused with the 18 kDa-fragment of the histidine-tag and the LysM domain was shown on SDS-PAGE analysis of a cell-free extract of LysM-LacZLbul as expected (Figure 2A). Western blot

analysis of the crude, cell-free extracts was performed using anti-His antibody for detection. Figure 2B shows that the recombinant bacteria produced the expected proteins, LysMLacL (lane 2) and LysMLacZ (lane 4). LacM was not detected as it does not contain the histidine-tag.
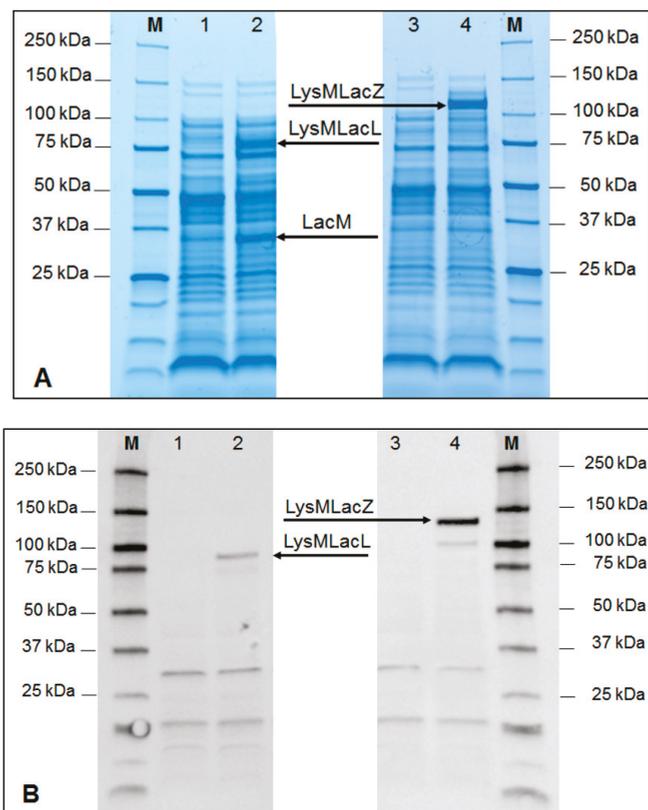
Figure 2. SDS-PAGE analysis (**A**) and Western blot analysis (**B**) of a cell-free extract of crude β-galactosidase fusion proteins, LysM-LacLMLreu (non-induced: lane 1, induced: lane 2) and LysM-LacZLbul (non-induced: lane 3, induced: lane 4), overexpressed in *E. coli* HST08. LacLMLreu encoded by two overlapping genes *lacLM* and LacZLbul encoded by *lacZ* gene are the β-galactosidases from *L. reuteri* and *L. delbrueckii* subsp. *bulgaricus* DSM 20081, respectively. The cultivation and induction conditions are as described in Materials and Methods and samples were taken at different time points after induction during cultivations. The arrows indicate the subunits of the recombinant β-galactosidases. M denotes the Precision protein ladder (Biorad, CA, USA).

To check if the heterologously produced enzymes were functionally active, β-galactosidase activities of cell-free lysates of *E. coli* cells carrying different expression vectors were measured. The highest yields obtained for the two recombinant enzymes were 11.1 ± 1.6 k·U_{oNPG} per L of medium with a specific activity of 6.04 ± 0.03 U·mg^{-1} for LysM-LacLMLreu and 46.9 ± 2.7 kU_{oNPG} per L of medium with a specific activity of 41.1 ± 0.9 U·mg^{-1} for LysM-LacZLbul, respectively (Table 1). The β-galactosidase activities in non-induced *E. coli* cells were negligible for both LysM-LacLMLreu and LysM-LacZLbul showing that the activity is from the overproduced β-galactosidases (Table 1).

Table 1. β-Galactosidase activities in cell-free lysates of *E. coli* cells carrying different expression vectors.

Expression Vector	Volumetric Activity (k·U/L Culture Medium)		Specific Activity (U/mg Protein)	
	Non-Induced	Induced	Non-Induced	Induced
pBAD3014LacLMLreu	n.d.	11.1 ± 1.5	n.d.	6.04 ± 0.03
pBAD3014LacZLbul	n.d.	46.9 ± 2.7	n.d.	41.1 ± 0.9

n.d.: not detected.

2.2. Display of Lactobacillal β-Galactosidases on Lactobacillus Cell Surface

To investigate the attachment of the two hybrid proteins, LysM-LacLMLreu and LysM-LacZLbul, to the cell wall of *L. plantarum,* cell-free crude extracts from *E. coli* harboring β-galactosidases corresponding to 50 U$_{oNPG}$ (~5–6 mg protein) were incubated with *L. plantarum* cells collected from one mL cultures at OD$_{600}$ ~4.0. The enzymes and *L. plantarum* were incubated at 37 °C with gentle agitation, and after 24 h of incubation, the residual activities in the supernatant as well as on the cell surface were determined for both enzymes (Table 2A). The immobilization yield (IY) is a measure of how much of the applied protein bound to the surface of *Lactobacillus* cells. Immobilizations yields for LysM-LacLMLreu and LysM-LacZLbul were 6.5% and 31.9%, respectively. SDS-PAGE analysis of the samples after the immobilization procedure showed strong bands of LysM-LacL and LacM or LysM-LacZ in the residual supernatants (Figure 3A, lane 2; Figure 3B, lane 2), indicating relatively high amounts of non-anchored proteins in the supernatants. Two successive washing steps with 50 mM sodium phosphate buffer (NaPB, pH 6.5) did not release the enzymes showing that the immobilization is both effective and stable (Figure 3A, lanes 4, 5; Figure 3B, lanes 3, 4). The low immobilization yield for LysM-LacLMLreu was confirmed by the SDS-PAGE analysis (Figure 3A, lane 3). Western blot analysis of the crude, cell-free extracts of *L. plantarum* LacZLbul-displaying cells was performed using an anti-His antibody for detection showing the presence of LacZLbul (Figure 3C; lane 3). Flow cytometry confirmed the surface localization of both enzymes LysM-LacLMLreu and LysM-LacZLbul as clear shifts in the fluorescence signals for *L. plantarum* LacLMLreu- and LacZLbul-displaying cells in comparison to the control strain were observed (Figure 4A,B). The surface-displayed enzymes were shown to be functionally active. β-Galactosidase activities obtained for *L. plantarum* displaying cells were 179 and 1153 U per g dry cell weight, corresponding to approximately 0.99 and 4.61 mg of active, surface-anchored β-galactosidase per g dry cell mass for LysM-LacLMLreu and LysM-LacZLbul (Table 2A), respectivel.
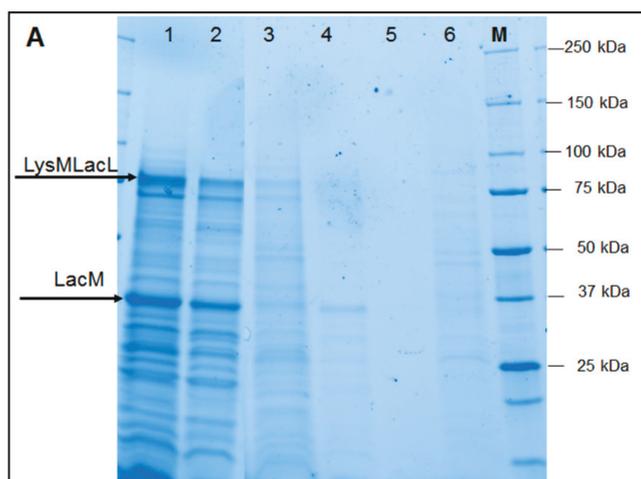
Figure 3. *Cont.*
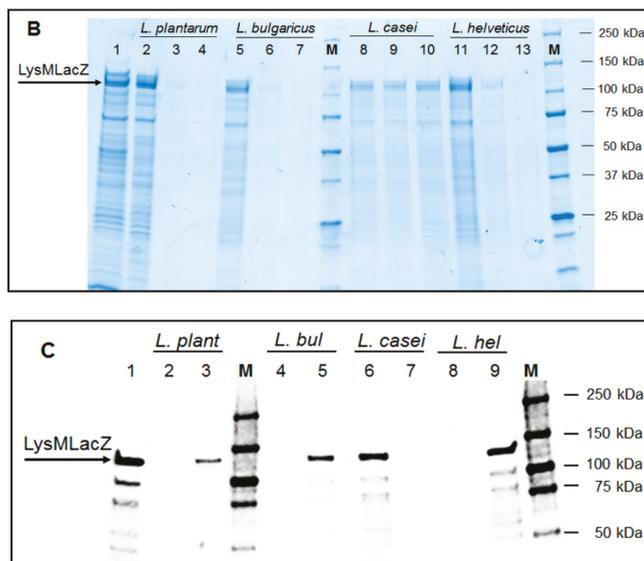
Figure 3. SDS-PAGE analysis (**A**,**B**) and Western blot analysis (**C**) of immobilization of recombinant enzymes. LacLMLreu encoded by two overlapping genes *lacLM* and LacZLbul encoded by *lacZ* gene are the β-galactosidases from *L. reuteri* and *L. delbrueckii* subsp. *bulgaricus* DSM 20081, respectively. The arrows indicate the subunits of the recombinant β-galactosidases. M denotes the Precision protein ladder (Biorad, CA, USA). (**A**) Cell-free crude extracts of *E. coli* HST08 harboring pBAD3014LacLMLreu (containing LysM-LacLMLreu) at 18 h of induction (lane1); flow through during immobilization (lane 2); surface anchored-LysM-LacLMLreu in *L. plantarum* WCFS1 (lane 3) and washing fractions (lanes 4, 5); non-displaying *L. plantarum* WCFS1 cells, negative control (lane 6). (**B**) Cell-free crude extracts of *E. coli* HST08 harboring pBAD3014LacZLbul (containing LysM-LacZLbul) at 18 h of induction (lane1); flow through during immobilization on the cell surface of *L. plantarum* WCFS1 (lane 2) and washing fractions (lanes 3, 4); flow through during immobilization on the cell surface of *L. delbrueckii* subsp. *bulgaricus* DSM 20081 (lane 5) and washing fractions (lanes 6, 7); flow through during immobilization on cell surface of *L. casei* (lane 8) and washing fractions (lanes 9, 10); flow through during immobilization on cell surface of *L. helveticus* DSM 20075 (lane 11) and washing fractions (lanes 12, 13). (**C**) Cell-free crude extracts of *E. coli* HST08 harboring pBAD3014LacZLbul (containing LysM-LacZLbul) at 18 h of induction (lane 1); non-displaying *L. plantarum* WCFS1 cells (lane 2) and surface anchored-LysM-LacZLbul in *L. plantarum* WCFS1 (lane 3); non-displaying *L. delbrueckii* subsp. *bulgaricus* DSM 20081 cells (lane 4) and surface anchored-LysM-LacZLbul in *L. delbrueckii* subsp. *bulgaricus* DSM 20081 (lane 5); surface anchored-LysM-LacZLbul in *L. casei* (lane 6) and non-displaying *L. casei* cells (lane 7); non-displaying *L. helveticus* DSM 20075 cells (lane 8) and surface anchored-LysM-LacZLbul in *L. helveticus* DSM 20075 (lane 9).
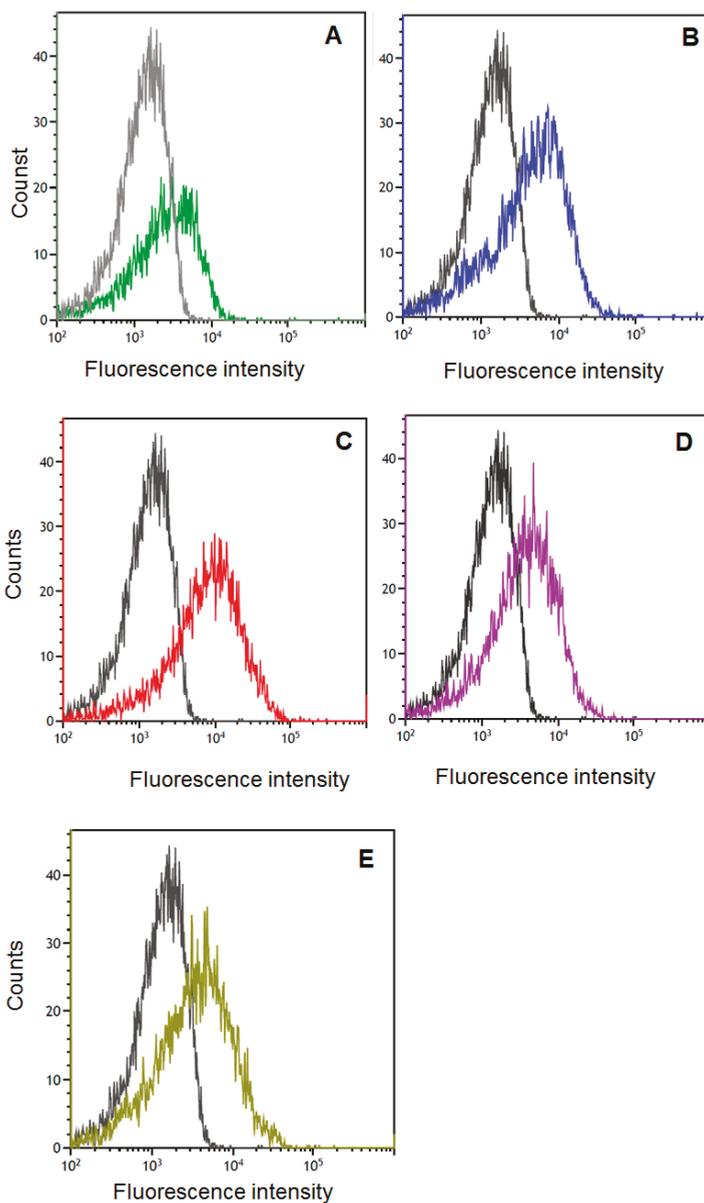
Figure 4. Analysis of surface localization of LysM-LacLMLreu and LysM-LacZLbul in *Lactobacillus* cells by using flow cytometry: surface anchored-LysM-LacLMLreu in *L. plantarum* WCFS1 (**A**, green line); surface anchored-LysM-LacZLbul in *L. plantarum* WCFS1 (**B**, blue line), in *L. delbrueckii* subsp. *bulgaricus* DSM 20081 (**C**, red line), in *L. casei* (**D**, purple line) and in *L. helveticus* DSM 20075 (**E**, olive line). Non-displaying *Lactobacillus* cells were used as negative controls (A–E, black line).

Table 2. Immobilization of (**A**) recombinant lactobacillal β-galactosidases on *L. plantarum* WCFS1 cell surface and (**B**) recombinant β-galactosidase from *L. bulgaricus* DSM 20081 (LysM-LacZLbul) on the cell surface of different *Lactobacillus* spp.

	Residual Activities in Supernatant	Immobilization Yield [a] (IY)	Activity on Cell Surface [b]		Activity Retention [c] (AR)	Amount of Active Surface Anchored β-gal [d]
	(%)	(%)	(%)	U/g DCW	(%)	mg/g DCW
(A) Enzyme (on *L. plantarum* WCFS1 cell surface)						
LysM-LacLMLreu	93.5 ± 1.2	6.53	3.06 ± 0.08	179 ± 5	46.9	0.99 ± 0.02
LysM-LacZLbul	68.1 ± 0.1	20.3 ± 0.2	20.3 ± 0.2	1153 ± 12	63.5	4.61 ± 0.05
(B) *Lactobacillus* spp. (with enzyme LysM-LacZLbul)						
L. plantarum WCFS1	68.1 ± 0.1	31.9	20.3 ± 0.2	1153 ± 12	63.5	4.61 ± 0.05
L. bulgaricus DSM 20081	71.3 ± 0.9	28.7	14.0 ± 0.9	795 ± 53	48.5	3.18 ± 0.11
L. casei	76.1 ± 0.9	23.9	15.1 ± 0.8	861 ± 48	63.2	3.44 ± 0.09
L. helveticus DSM20075	75.3 ± 0.9	24.7	14.3 ± 0.5	812 ± 27	57.7	3.25 ± 0.11

[a] IY (%) was calculated by subtraction of the residual enzyme activity (%) in the supernatant after immobilization from the total activity applied (100%). [b] Activity on the cell surface (%) is the percentage of enzyme activity measured on the cell surface to the total applied activity. Activity on the cell surface (U/g DCW) is calculated as the amount of enzyme (Units) per g dry cell weight. [c] Activity retention, AR (%), is the ratio of activity on the cell surface (%) to IY (%). [d] It was calculated based on specific activities of purified LacLMLreu of 180 U/mg protein [16] and of purified LacZLbul (His Tagged) of 250 U/mg protein [12]. Values given are the average value from at least two independent experiments, and the standard deviation was always less than 5%.

Due to higher immobilization yields and increased amounts of active surface-anchored protein in *L. plantarum*, LysM-LacZLbul was chosen for further analysis of its display on the cell surface of other food-relevant *Lactobacillus* spp. including *L. bulgaricus*, *L. casei* and *L. helveticus*. The parameters of residual activities in the supernatant after the anchoring experiment, activity on the cell surface, immobilization yields, activity retention and amounts of active surface-anchored LysM-LacZLbul were determined and are presented in Table 2B. It was shown that surface-anchored LysM-LacZLbul was released from the cell surface of *L. casei* during the subsequent washing steps (Figure 3B, lanes 9, 10). Western blot analysis of the crude, cell-free extracts of *Lactobacillus* LysM-LacZLbul-displaying cells indicated the binding of LysM-LacZLbul to all four *Lactobacillus* spp. tested (Figure 3C; lanes 3, 5, 6, 9) as was also confirmed by flow cytometry (Figure 4B–E). *L. plantarum* bound most efficiently among the tested *Lactobacillus* species shown by the highest immobilization yield and the highest amount of active, surface-anchored LysM-LacZLbul (Table 2B).

2.3. Enzymatic Stability of β-Galactosidase-Displaying Cells

Both temperature stability and reusability of β-galactosidase displaying cells were determined. For temperature stability, *L. plantarum* galactosidase-displaying cells were incubated in 50 mM sodium phosphate buffer (NaPB), pH 6.5 at different temperatures, and at certain time intervals, the residual β-galactosidase activities on *L. plantarum* cell surface were measured. Both LysM-LacLMLreu and LysM-LacZLbul-displaying cells are very stable at −20 °C with a half-life time of activity ($\tau_{\frac{1}{2}}$) of approximately 6 months (Table 3). The half-life time of activity of LysM-LacLMLreu-displaying cells at 30 °C is 55 h, whereas half-life times of activity of LysM-LacZLbul-displaying cells at 30 °C and 50 °C are 120 h and 30 h, respectively (Table 3).

Table 3. Stability of *L. plantarum* β-galactosidase-displaying cells at various temperatures [a].

LysM-LacLMLreu		LysM-LacLZLbul	
Temperature	$\tau_{\frac{1}{2}}$	Temperature	$\tau_{\frac{1}{2}}$
−20 °C	6 months	−20 °C	6 months
4 °C	3 months	4 °C	Nd [b]
30 °C	55 h	30 °C	120 h
50 °C	nd [b]	50 °C	30 h

[a] *L. plantarum* galactosidase-displaying cells were incubated in 50mM sodium phosphate buffer (NaPB), pH 6.5 at different temperatures. Experiments were performed at least in duplicates. [b] not determined.

To test the reusability of LysM-LacLMLreu- and LysM-LacZLbul-displaying cells, the enzyme activity was measured during several repeated rounds of lactose conversion with two washing steps between each cycle. The enzymatic activities of *L. plantarum* LysM-LacZLbul-displaying cells decreased by ~23% and 27% at 30 °C and 50 °C (Figure 5), respectively, after three conversion/washing cycles, indicating that these displaying cells can be reused for several rounds of biocatalysis at tested temperatures. LysM-LacLMLreu-displaying cells are less stable than LysM-LacZLbul-displaying cells as only 56% of the initial β-galactosidase activity are retained at 30 °C after the third cycle (Figure 5). LysM-LacZLbul-displaying cells retained 35% of β-galactosidase activity after the fourth cycle at 50 °C, 57% and 51% after the fourth and fifth cycle, respectively, at 30 °C (Figure 5). These observations indicate that immobilized fusion LysM-β-galactosidases can be reused for at least four to five repeated rounds of lactose conversion.
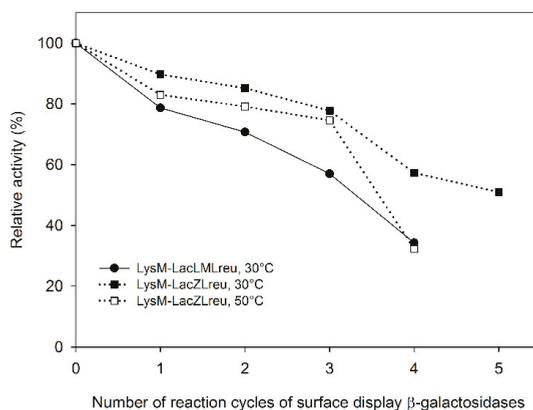
Figure 5. Enzymatic activity of surface display β-galactosidases, LysM-LacLMLreu- and LysM-LacZLbul, during several repeated rounds of lactose conversion using *L. plantarum* WCFS1 displaying cells. Experiments were performed in duplicates, and the standard deviation was always less than 5%.

2.4. Formation of Galacto-Oligosaccharides (GOS)

Figure 6 shows the formation of GOS using *L. plantarum* cells displaying β-galactosidase LacZ from *L. bulgaricus* (LysM-LacZLbul) with 1.0 U_{Lac} β-galactosidase activity per mL of the reaction mixture and 205 g/L initial lactose in 50 mM sodium phosphate buffer (pH 6.5) at 30°C. The maximal GOS yield was around 32% of total sugars obtained at 72% lactose conversion after 7 h of conversion. This observation shows that surface-displayed LacZ is able to convert lactose to form galacto-oligosaccharides. We could identify the main GOS products of transgalactosylation, which are β-D-Gal*p*-(1→6)-D-Glc, β-D-Gal*p*-(1→3)-D-Lac, β-D-Gal*p*-(1→3)-D-Glc, β-D-Gal*p*-(1→3)-D-Gal, β-D-Gal*p*-(1→6)-D-Gal, and

β-D-Gal*p*-(1→6)-D-Lac. This is similar to the product profile when performing the conversion reaction with the free enzyme as previously reported [12].
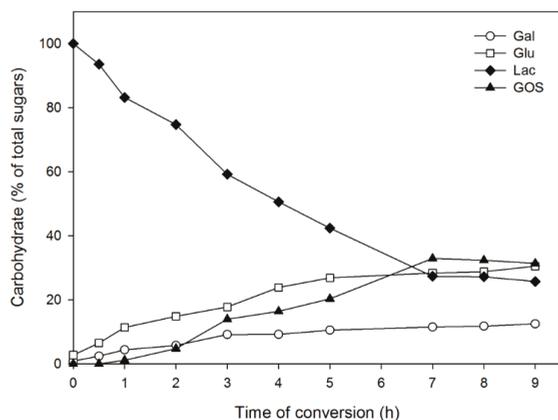
Figure 6. Course of reaction for lactose conversion by surface display β-galactosidase from *L. bulgaricus* (LysM-LacZLbul) in *L. plantarum* WCFS1 as determined by HPLC. The batch conversion was carried out at 30 °C using 205 g/L initial lactose concentration in 50 mM NaPB (pH 6.5) and constant agitation (500 rpm). *L. plantarum* LysM-LacZLbul displaying cells were added to equivalent concentrations of 1.0 U_{Lac}/mL of the reaction mixture. Experiments were performed in duplicates, and the standard deviation was always less than 5%.

3. Discussion

Surface display of proteins on cells of lactic acid bacteria (LAB) generally requires genetic modifications, which might have limitations in food and medical applications due to the sensitive issue of the use of genetically modified organisms (GMO). Anchoring heterologous proteins on the cell surface of non-genetically modified LAB (non-GMO) via mediated cell wall binding domains including surface layer domain (SLPs) [33,34], LysM domain [26,30,35–37], W × L domains [38] attracts increasing interest.

Lysin motif (LysM) domains are found in many bacterial peptidoglycan hydrolases [26,38,39]. Peptidoglycan contains sugar (glycan) chains, which consist of *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) units joined by glycosidic linkages. Proteins harboring LysM motifs have been shown to bind non-covalently to the peptidoglycan layer and have been employed to display heterologous proteins on the bacterial cell-surface [26,40,41]. These domains can contain single or multiple LysM motifs [41], and they have been used to display proteins in LAB by fusion either to the N- or C-terminus of a target protein [27–30]. Interestingly, the LysM motif derived from the *L. plantarum* Lp_3014 transglycosylase has been used successfully for surface display of invasin [36] and a chemokine fused to an HIV antigen [37] previously.

In this work, we used the single LysM domain derived from Lp_3014 to anchor two different lactobacillal β-galactosidases, a heterodimeric type from *L. reuteri* and a homodimeric type from *L. bulgaricus*, on the cell surface of four species of lactobacilli. Functional active fusion proteins, LysM-LacLMLreu and LysM-LacZLbul, were successfully expressed in *E. coli*. However, the expression yield of LysM-LacLMLreu was ten-fold lower than that of the β-galactosidase from *L. reuteri* (LacLMLreu) without LysM expressed previously in *E. coli*, which was reported to be 110 kU of β-galactosidase activity per liter of cultivation medium [16]. This may indicate that the fusion of the LysM domain has a negative effect on the expression level. Interestingly, the expression yields of LysM-LacZLbul were 4-fold and 7-fold higher in terms of volumetric and specific activities, respectively, than that of LysM-LacLMLreu using the same host, expression system and induction conditions.

β-Galactosidase from *L. bulgaricus* (LysM-LacZLbul) is a homodimer whereas β-galactosidase from *L. reuteri* (LysM-LacLMLreu) is a heterodimer, and hence the fusion of the LysM domain only to the LacL subunits might lead to the discrepancy between the yields of these two fusion proteins due to different folding mechanisms.

Not surprisingly, the affinity for peptidoglycan of homodimeric LysM-LacZLbul is significantly higher than LysM-LacLMLreu as shown by the immobilization yield (Table 2A). As aforementioned LacLMLreu from *L. reuteri* is a heterodimer and the LysM domain is fused N-terminally to only LacL, while LacZLbul from *L. bulgaricus* is a homodimer, hence each of the identical subunits will carry its own LysM domain leading to stronger attachment of LacZ on the *L. plantarum* cell wall. This could be a likely explanation for the higher immobilization yields observed for LysM-LacZLbul. Even though the immobilization yields obtained in this study were significantly lower than the immobilization yields for these same enzymes when a chitin binding domain (ChBD) together with chitin was used [17], the activity retention (AR) on the *L. plantarum* cell surface (46.9% and 63.5% for LysM-LacLMLreu and LysM-LacZLbul, respectively) were significantly higher. The AR values for ChBD-LacLM, LacLM-ChBD and LacZ-ChBD using chitin beads were 19%, 26% and 13%, respectively [17]. Notably, the amount of active surface anchored LysM-LacLMLreu (0.99 ± 0.02 mg per g dry cell weight) on the cell surface of *L. plantarum* WCFS1 is significantly lower than LysM-LacZLbul (4.61 ± 0.05 mg per g dry cell weight). This is mainly due to the low immobilization yield of LysM-LacLMLreu. *L. plantarum* collected from one mL cultures at OD_{600} ~4.0 was used in immobilization reactions, hence the amount of *L. plantarum* cells was estimated to be ~3.0×10^9 cfu/mL. Therefore, we calculated that 8.22 µg LysM-LacLMLreu and 38.3 µg LysM-LacZLbul anchored on 3.0×10^9 *L. plantarum* cells or 0.002 pg LysM-LacLMLreu and 0.012 pg LysM-LacZLbul per *L. plantarum* cell. Xu et al. (2011) reported the use of the putative muropeptidase MurO (Lp_2162) from *L. plantarum* containing two putative LysM repeat regions for displaying a green fluorescent protein (GFP) and a β-galactosidase from *Bifidobacterium bifidum* on the surface of *L. plantarum* cells [42]. They reported that 0.008 pg of GFP was displayed per cell on non-treated *L. plantarum* cells, while the amount of active surface anchored β-galactosidase from *B. bifidum* on the surface of *L. plantarum* cells was not reported in that study.

Further, we tested the capability of binding the fusion protein LysM-LacZLbul to the cell wall of three other *Lactobacillus* species. *L. plantarum* showed the best capacity among the tested *Lactobacillus* for surface anchoring of LysM-LacZLbul (Table 2B), whereas *L. bulgaricus*, *L. casei* and *L. helveticus* are comparable in term of the amount of active surface-anchored enzyme.

The highest GOS yield of 32% obtained with the surface-immobilized enzyme is lower than the yield obtained with the free enzyme LacZ from *L. bulgaricus* (Figure 6), which was previously reported to be approximately 50% [12]. This could be due to the binding of LysM-LacZLbul to the peptidoglycan and the attachment of the enzyme on *Lactobacillus* cell surface, which might hinder the access of the substrate lactose to the active site of the enzyme. Interestingly, the GOS yield obtained from lactose conversion using *L. plantarum* cells displaying β-galactosidase (LysM-LacZLbul) from *L. bulgaricus* is significantly higher than the yield obtained with immobilized β-galactosidase (LacZ-ChBD) on chitin, which was previously reported around 23%–24% [12]. It indicates that β-galactosidase from *L. bulgaricus* anchored on *L. plantarum* cell surface is more catalytically efficient than its immobilized form on chitin.

4. Materials and Methods

4.1. Bacterial Strains and Culture Conditions

The bacterial strains and plasmids used in this study are listed in Table 4. *Lactobacillus plantarum* WCFS1, isolated from human saliva as described by Kleerebezem et al. [32], was originally obtained from NIZO Food Research (Ede, The Netherlands) and maintained in the culture collection of the Norwegian University of Life Sciences, Ås, Norway. *L. helveticus* DSM 20075 (ATCC 15009) and *L. delbrueckii* subsp. *bulgaricus* DSM 20081 (ATCC 11842) were obtained from the German Collection

of Microorganisms and Cell Cultures (DSMZ; Braunschweig, Germany). *L. casei* was obtained from the culture collection of the Food Biotechnology Laboratory, BOKU-University of Natural Resources and Life Sciences Vienna. *Lactobacillus* strains were cultivated on MRS medium (*Lactobacillus* broth according to De Man, Rogosa and Shape [43]) (Carl Roth, Karlsruhe, Germany) at 37 °C without agitation. *E. coli* NEB5α (New England Biolabs, Frankfurt am Main, Germany) was used as cloning hosts in the transformation of DNA fragments; whereas *E. coli* HST08 (Clontech, Mountain View, CA, USA) was used as the expression host strain. *E. coli* strains were cultivated in Luria-Bertani (LB) medium (10g/L tryptone, 10 g/L NaCl, and 5 g/L yeast extract) at 37 °C with shaking at 140 rpm. Agar media were prepared by adding 1.5% agar to the respective media. When needed, ampicillin was supplemented to media to a final concentration of 100 µg/mL for *E. coli* cultivations.

Table 4. Strains and plasmids used in the study.

Strains or Plasmids	Relevant Characteristics	Reference Source
Strains		
L. plantarum WCFS1		[32]
L. delbrueckii subsp. *bulgaricus* DSM 20081		DSMZ
L. casei		BOKU
L. helveticus DSM 20075		DSMZ
E. coli HST08	Host strain	Clontech
Plasmids		
pBAD_3014_AgESAT_DC	Amp^r; pBAD derivate with the LysM domain sequence from Lp3014 fused to the hybrid antigen AgESAT_DC	[44]
pBAD3014LacLMLreu	Amp^r; pBAD3014_AgESAT_DC derivative with a fragment of *lacLM* genes instead of the gene fragment encoding AgESAT_DC	This study
pBAD3014LacZLbul	Amp^r; pBAD_3014_AgESAT_DC derivate with *lacZ* fragment instead of the gene fragment encoding AgESAT_DC	This study
pHA1032	Amp^r; pET21d derivative for expression of *lacLM* from *L. reuteri* in *E. coli*	[16]
pTH103	Erm^r; *spp*-based expression vector pSIP409 for expression of *lacZ* from *L. bulgaricus* DSM 20081 in *L. plantarum* WCFS1	[12]

4.2. Chemicals, Enzymes and Plasmids

All chemicals and enzymes were purchased from Sigma (St. Louis, MO, USA) unless stated otherwise and were of the highest quality available. All restriction enzymes, Phusion high-fidelity DNA polymerase, T4 DNA ligase, and corresponding buffers were from New England Biolabs (Frankfurt am Main, Germany). Staining dyes, DNA and protein standard ladders were from Bio-Rad (Hercules, CA, USA). All plasmids used in this study are listed in Table 4.

4.3. DNA Manipulation

Plasmids were isolated from *E. coli* strains using Monarch Plasmid Miniprep Kit (New England Biolabs, Frankfurt am Main, Germany) according to the manufacturer's instructions. PCR amplifications of DNA were done using Q5 High-Fidelity 2X Master Mix (New England Biolabs). The primers used in this study, which were supplied by VBC-Biotech Service (Vienna, Austria), are listed in Table 5. PCR products and DNA fragments obtained by digestion with restriction enzymes were purified using Monarch DNA Gel Extraction Kit (New England Biolabs); and the DNA amounts were estimated using Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA). The sequences of PCR-generated fragments were verified by DNA sequencing performed by a commercial provider (Microsynth, Vienna, Austria). The ligation of DNA fragments was performed using NEBuilder HiFi Assembly Cloning Kit (New England Biolabs). All plasmids were transformed into *E. coli* NEB5α chemical competent cells following the manufacturer's protocol for obtaining the plasmids in sufficient amounts. The constructed plasmids (Table 4) were chemically transformed into expression host strain *E. coli* HST08.

Table 5. Primers used in the study.

Primer	Sequence* 5'→3'	Restriction Site Underlined
Fwd1LreuSalI	GAGTTCAACT<u>GTCGAC</u>CAAGCAAATATAAA	*Sal*I
Rev1LreuEcoRI	AGCCAAGCTTC<u>GAATT</u>CTTATTTTGCATTC	*Eco*RI
Fwd2LbulSalI	GTTCAACT<u>GTCGAC</u>AGCAATAAGTTAGTAAAAGAAAAAGAG	*Sal*I
Rev2LbulEcoRI	CAGCCAAGCTTC<u>GAATT</u>CTTATTTTAGTAAAAGGGGCTGAATC	*Eco*RI

* The nucleotides in italics are the positions that anneal to the DNA of the target genes (*lacLM* or *lacZ*).

4.4. Plasmid Construction

Two recombinant fusion proteins were constructed. The first fusion protein was based on LacLM from *L. reuteri* and the LysM domain attached upstream of LacLM (termed LysM-LacLMLreu). The second fusion protein was based on LacZ from *L. delbrueckii* subsp. *bulgaricus* DSM 20081 and the LysM domain attached upstream of LacZ (termed LysM-LacZLbul). Plasmid pBAD_3014AgESAT_DC (Table 4) [44] (was used for the construction of the expression plasmids. This plasmid is a derivate of pBAD vector (Invitrogen, Carlsbad, CA, USA) containing a 7 × His tag sequence and a single LysM domain from Lp_3014, which is a putative extracellular transglycosylase with LysM peptidoglycan binding domain from *L. plantarum* WCFS1 (NCBI reference sequence no. NC_004567.2) [31,32], fused to the hybrid tuberculosis antigen AgESAT-DC [44]. The fragment of *lacLM* genes from *L. reuteri* was amplified from the plasmid pHA1032 (Table 4) [16] with the primer pair Fwd1LreuSalI and Rev1LreuEcoRI (Table 5), whereas the *lacZ* gene from *L. bulgaricus* was amplified from the plasmid pTH103 (Table 4) [12] with the primer pair Fwd2LbulSalI and Rev2LbulEcoRI (Table 5). The PCR-generated products were then cloned into *Sal*I and *Eco*RI cloning sites of the pBAD_3014AgESAT_DC vector using and NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs) following the manufacturer's instructions, resulting in two expression plasmids pBAD3014LacLMLreu and pBAD3014LacZLbul (Figure 1).

4.5. Gene Expression in E. coli

The constructed plasmids pBAD3014LacLMLreu and pBAD3014LacZLbul were chemically transformed into expression host *E. coli* HST08. For gene expression, overnight cultures of *E. coli* HST08 were diluted in 300 mL of fresh LB broth containing 100 μg/mL ampicillin to an OD_{600} of ~0.1 and incubated at 37 °C with shaking at 140 rpm to an OD_{600} ~0.6. Gene expression was then induced by L-arabinose to a final concentration of 0.7 mg/mL and the cultures were incubated further at 25 °C for 18 h with shaking at 140 rpm. Cells were harvested at an OD_{600} of ~3.0 by centrifugation at 4000× *g* for 30 min at 4 °C, washed twice, and resuspended in 50 mM sodium phosphate buffer (NaPB), pH 6.5. Cells were disrupted by using a French press (AMINCO, Maryland, USA). Debris was removed by centrifugation (10,000× *g* for 15 min at 4 °C) to obtain the crude extract.

4.6. Immobilization of β-Galactosidases on Lactobacillus Cell Surface

One mL of *Lactobacillus* cultures were collected at OD_{600} ~4.0 by centrifugation (4000× *g* for 5 min at 4 °C) and the cells were washed with 50 mM sodium phosphate buffer (NaPB), pH 6.5. The cell pellets were then mixed with one mL of diluted cell-free crude extracts of 50 U_{oNPG}/mL (~5–6 mg protein/mL) of fused LysM-β-galactosidases (LysM-LacLMLreu or LysM-LacZLbul) and incubated at 37 °C for 24 h with gentle agitation. *Lactobacillus* β-galactosidase displaying cells were separated from the supernatants by centrifugation (4000× *g* for 5 min at 4 °C). Cells were then washed with NaPB (pH 6.5) two times; the supernatants and wash solutions were collected for SDS-PAGE analysis and activity and protein measurements. *Lactobacillus* β-galactosidase displaying cells were resuspended in NaPB (pH 6.5) for further studies.

4.7. Protein Determination

Protein concentrations were determined using the method of Bradford [45] with bovine serum albumin (BSA) as standard.

4.8. β-Galactosidase Assays

β-Galactosidase activity was determined using *o*-nitrophenyl-β-D-galactopyranoside (*o*NPG) or lactose as the substrates as previously described [5] with modifications. When chromogenic substrate *o*NPG was used, the reaction was started by adding 20 μL of *Lactobacillus* β-galactosidase displaying cell suspension to 480 μL of 22 mM *o*NPG in 50 mM NaBP (pH 6.5) and stopped by adding 750 μL of 0.4 M Na_2CO_3 after 10 min of incubation at 30 °C. The release of *o*-nitrophenol (*o*NP) was measured by determining the absorbance at 420 nm. One unit of *o*NPG activity was defined as the amount of β-galactosidase releasing 1 μmol of *o*NP per minute under the defined conditions.

When lactose was used as the substrate, 20 μL of *Lactobacillus* β-galactosidase displaying cell suspension was added to 480 μL of a 600 mM lactose solution in 50 mM sodium phosphate buffer, pH 6.5. After 10 min of incubation at 30 °C, the reaction was stopped by heating the reaction mixture at 99 °C for 5 min. The reaction mixture was cooled to room temperature, and the release of D-glucose was determined using the test kit from Megazyme. One unit of lactase activity was defined as the amount of enzyme releasing 1 μmol of D-glucose per minute under the given conditions.

4.9. Gel Electrophoresis Analysis

For visual observation of the expression level of the two recombinant β-galactosidases (LysM-LacLMLreu and LysM-LacZLbul) in *E. coli* and the effectiveness of the immobilization, cell-free extracts, supernatants, and wash solutions were analyzed by Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE). Protein bands were visualized by staining with Bio-safe Coomassie (Bio-Rad). The determination of protein mass was carried out using Unstained Precision plus Protein Standard (Bio-Rad).

4.10. Western Blotting

Proteins in the cell-free extracts were separated by SDS-PAGE. Protein bands were then transferred to a nitrocellulose membrane using the Trans-Blot Turbo™ Transfer System (Biorad) following the manufacturer's instructions. Monoclonal mouse anti-His antibody (Penta His Antibody, BSA-free) was obtained from Qiagen (Hilden, Germany), diluted 1:5000 and used as recommended by the manufacturer. The protein bands were visualized by using polyclonal rabbit anti-mouse antibody conjugated with horseradish peroxidase (HRP) (Dako, Denmark) and the Clarity™ Western ECL Blotting Substrate from Bio-Rad (Hercules, CA, USA).

4.11. Flow Cytometry

Lactobacillus β-galactosidase displaying cells were resuspended in 50 μL of phosphate buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 2 mM KH_2PO_4, and 10 mM Na_2HPO_4, pH 7.4) containing 2% of BSA (PBS-B) and 0.1 μL of Penta His Antibody, BSA-free (Qiagen; diluted 1:500 in PBS-B). After incubation at RT for 40 min, the cells were centrifuged at 4000× *g* for 5 min at 4 °C and washed three times with 500 μL PBS-B. The cells were subsequently incubated with 50 μL PBS-B and 0.1 μL anti-mouse IgG H&L/Alexa Flour 488 conjugate (Cell Signaling Technology, Frankfurt am Main, Germany, diluted 1:750 in PBS-B) for 40 min in the dark at room temperature. After washing five times with 500 μL PBS-B, the stained cells were analyzed by flow cytometry using a CytoFLEX Flow Cytometer (Beckman Coulter, Brea, CA, USA) following the manufacturer's instructions.

4.12. Temperature Stability and Reusability of Immobilized Enzymes

The temperature stability of immobilized enzymes was studied by incubating *L. plantarum* LysM-LacLMLreu- and LysM-LacZLbul-displaying cells in 50 mM NaPB (pH 6.5) at various temperatures (−20, 4, 30, 50 °C). At certain time intervals, samples were withdrawn, the residual activity was measured using oNPG as the substrate under standard assay conditions and the $\tau_{1/2}$ value was determined.

To test the reusability of immobilized enzymes, several repeated rounds of lactose conversion at 30 °C using LysM-LacLMLreu- and LysM-LacZLbul-displaying cells and at 50 °C using LysM-LacZLbul-displaying cells were carried out with 600 mM initial lactose in 50mM NaBP (pH 6.5) and constant agitation (500 rpm). The enzyme activity during these repeated cycles with intermediate two washing steps was measured using oNPG as the substrate under standard assay conditions.

4.13. Lactose Conversion and Formation of Galacto-Oligosaccharides (GOS)

The conversion of lactose was carried out in discontinuous mode using *L. plantarum* cells displaying β-galactosidase LacZ from *L. bulgaricus* (LysM-LacZLbul). The conversion was performed at 30 °C using 205 g/L initial lactose concentration in 50 mM NaPB (pH 6.5) and constant agitation (500 rpm). *L. plantarum* LysM-LacZLbul displaying cells were added to equivalent concentrations of 1.0 U_{Lac}/mL of reaction mixture. Samples were withdrawn at intervals, heated at 99 °C for 5 min and further analyzed for lactose, galactose, glucose and GOS present in the samples.

4.14. Analysis of Carbohydrate Composition

The carbohydrate composition in the reaction mixture was analyzed by high-performance liquid chromatography (HPLC) equipped with a Dionex ICS-5000+ system (Thermo Fisher Scientific) consisting of an ICS-5000+ dual pump (DP) and an electrochemical detector (ED). Separations were performed at room temperature on CarboPac PA-1 column (4 × 250 mm) connected to a CarboPac PA-1 guard column (4 × 50 mm) (Thermo Fisher Scientific) with flow rate 1 mL/min. All eluents A (150 mM NaOH), B (150 mM NaOH and 500 mM sodium acetate) and C (deionized water) were degassed by flushing with helium for 30 min. Separation of D-glucose, D-galactose, lactose and allolactose was carried out with a run with the following gradient: 90% C with 10% A for 45 min at 1.0 mL/min, followed by 5 min with 100% B. The concentration of saccharides was calculated by interpolation from external standards. Total GOS concentration was calculated by subtraction of the quantified saccharides (lactose, glucose, galactose) from the initial lactose concentration. The GOS yield (%) was defined as the percentage of GOS produced in the samples compared to initial lactose.

4.15. Statistical Analysis

All experiments and measurements were conducted at least in duplicate, and the standard deviation (SD) was always less than 5%. The data are expressed as the mean ± SD when appropriate.

5. Conclusions

This work describes the immobilization of two lactobacillal β-galactosidases, a β-galactosidase from *L. reuteri* of the heterodimeric LacLM-type and one from *L. bulgaricus* of the homodimeric LacZ-type, on the *Lactobacillus* cell surface using a peptidoglycan-binding motif as an anchor, in this case, the single LysM domain Lp_3014 from *L. plantarum* WCFS1. The immobilized fusion LysM-β-galactosidases are catalytically efficient and can be reused for several repeated rounds of lactose conversion. Surface anchoring of β-galactosidases in *Lactobacillus* results in safe, non-GMO and stable biocatalysts that can be used in the applications for lactose conversion and production of prebiotic galacto-oligosaccharides.

Author Contributions: Conceptualization, G.M. and T.-H.N.; Data Curation, M.-L.P., G.M. and T.-H.N.; Investigation, M.-L.P., A.-M.T., S.K. and T.-T.N.; Methodology, M.-L.P. and T.-H.N.; Supervision, T.-H.N.; Writing-Original Draft Preparation, M.-L.P.; Writing-Review & Editing, T.-H.N.

References

1. Nakayama, T.; Amachi, T. β-Galactosidase, enzymology. In *Encyclopedia of Bioprocess Technology, Fermentation, Biocatalysis, and Bioseparation*; Flickinger, M.C., Drew, S.W., Eds.; John Willey and Sons: New York, NY, USA, 1999; Volume 3, pp. 1291–1305.

2. Pivarnik, L.F.; Senegal, A.G.; Rand, A.G. Hydrolytic and transgalactosylic activities of commercial β-galactosidase (lactase) in food processing. *Adv. Food Nutr. Res.* **1995**, *38*, 1–102.

3. Prenosil, J.E.; Stuker, E.; Bourne, J.R. Formation of oligosaccharises during enzymatic lactose hydrolysis: Part I: State of art. *Biotechnol. Bioeng.* **1987**, *30*, 1019–1025. [CrossRef] [PubMed]

4. Petzelbauer, I.; Zeleny, R.; Reiter, A.; Kulbe, K.D.; Nidetzky, B. Development of an ultra-high-temperature process for the enzymatic hydrolysis of lactose: II. Oligosaccharide formation by two thermostable β-glycosidases. *Biotechnol. Bioeng.* **2000**, *69*, 140–149. [CrossRef]

5. Nguyen, T.H.; Splechtna, B.; Steinböck, M.; Kneifel, W.; Lettner, H.P.; Kulbe, K.D.; Haltrich, D. Purification and characterization of two novel β-galactosidases from *Lactobacillus reuteri*. *J. Agric. Food Chem.* **2006**, *54*, 4989–4998. [CrossRef]

6. Nguyen, T.H.; Splechtna, B.; Krasteva, S.; Kneifel, W.; Kulbe, K.D.; Divne, C.; Haltrich, D. Characterization and molecular cloning of a heterodimeric β-galactosidase from the probiotic strain *Lactobacillus acidophilus* R22. *FEMS Microbiol. Lett.* **2007**, *269*, 136–144. [CrossRef] [PubMed]

7. Kittibunchakul, S.; Pham, M.-L.; Tran, A.-M.; Nguyen, T.-H. β-Galactosidase from *Lactobacillus helveticus* DSM 20075: Biochemical characterization and recombinant expression for applications in dairy industry. *Int. J. Mol. Sci.* **2019**, *20*, 947. [CrossRef] [PubMed]

8. Iqbal, S.; Nguyen, T.H.; Nguyen, T.T.; Maischberger, T.; Haltrich, D. β-galactosidase from *Lactobacillus plantarum* WCFS1: Biochemical characterization and formation of prebiotic galacto-oligosaccharides. *Carbohydr. Res.* **2010**, *345*, 1408–1416. [CrossRef]

9. Gobinath, D.; Prapulla, S.G. Permeabilized probiotic *Lactobacillus plantarum* as a source of β-galactosidase for the synthesis of prebiotic galactooligosaccharides. *Biotechnol. Lett.* **2013**, *36*, 153–157. [CrossRef]

10. Iqbal, S.; Nguyen, T.H.; Nguyen, H.A.; Nguyen, T.T.; Maischberger, T.; Kittl, R.; Haltrich, D. Characterization of a heterodimeric GH2 β-galactosidase from *Lactobacillus sakei* Lb790 and formation of prebiotic galacto-oligosaccharides. *J. Agric. Food Chem.* **2011**, *59*, 3803–3811. [CrossRef] [PubMed]

11. Maischberger, T.; Leitner, E.; Nitisinprasert, S.; Juajun, O.; Yamabhai, M.; Nguyen, T.H.; Haltrich, D. β-galactosidase from *Lactobacillus pentosus*: Purification, characterization and formation of galacto-oligosaccharides. *Biotechnol. J.* **2010**, *5*, 838–847. [CrossRef]

12. Nguyen, T.T.; Nguyen, H.A.; Arreola, S.L.; Mlynek, G.; Djinović-Carugo, K.; Mathiesen, G.; Nguyen, T.H.; Haltrich, D. Homodimeric β-galactosidase from *Lactobacillus delbrueckii* subsp. *bulgaricus* DSM 20081: Expression in *Lactobacillus plantarum* and biochemical characterization. *J. Agric. Food Chem.* **2012**, *60*, 1713–1721. [CrossRef]

13. Black, B.A.; Lee, V.S.Y.; Zhao, Y.Y.; Hu, Y.; Curtis, J.M.; Ganzle, M.G. Structural identification of novel oligosaccharides produced by *Lactobacillus bulgaricus* and *Lactobacillus plantarum*. *J. Agric. Food Chem.* **2012**, *60*, 4886–4894. [CrossRef]

14. Liu, G.X.; Kong, J.; Lu, W.W.; Kong, W.T.; Tian, H.; Tian, X.Y.; Huo, G.C. Beta-galactosidase with transgalactosylation activity from *Lactobacillus fermentum* K4. *J. Dairy Sci.* **2011**, *94*, 5811–5820. [CrossRef]

15. Nie, C.; Liu, B.; Zhang, Y.; Zhao, G.; Fan, X.; Ning, X.; Zhang, W. Production and secretion of *Lactobacillus crispatus* β-galactosidase in *Pichia pastoris*. *Protein Expr. Purif.* **2013**, *92*, 88–93. [CrossRef] [PubMed]

16. Nguyen, T.H.; Splechtna, B.; Yamabhai, M.; Haltrich, D.; Peterbauer, C. Cloning and expression of the β-galactosidase genes from *Lactobacillus reuteri* in *Escherichia coli*. *J. Biotechnol.* **2007**, *129*, 581–591. [CrossRef]

17. Pham, M.L.; Leister, T.; Nguyen, H.A.; Do, B.C.; Pham, A.T.; Haltrich, D.; Yamabhai, M.; Nguyen, T.H.; Nguyen, T.T. Immobilization of β-galactosidases from *Lactobacillus* on chitin using a chitin-binding domain. *J. Agric. Food Chem.* **2017**, *65*, 2965–2976. [CrossRef]

18. Michon, C.; Langella, P.; Eijsink, V.G.; Mathiesen, G.; Chatel, J.M. Display of recombinant proteins at the surface of lactic acid bacteria: Strategies and applications. *Microb. Cell Fact.* **2016**, *15*, 70. [CrossRef]

19. Nguyen, H.M.; Mathiesen, G.; Stelzer, E.M.; Pham, M.L.; Kuczkowska, K.; Mackenzie, A.; Agger, J.W.; Eijsink, V.G.; Yamabhai, M.; Peterbauer, C.K.; et al. Display of a β-mannanase and a chitosanase on the cell surface of *Lactobacillus plantarum* towards the development of whole-cell biocatalysts. *Microb. Cell Fact.* **2016**, *15*, 169. [CrossRef] [PubMed]

20. Schneewind, O.; Missiakas, D.M. Protein secretion and surface display in Gram-positive bacteria. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 1123–1139. [CrossRef] [PubMed]

21. Leenhouts, K.; Buist, G.; Kok, J. Anchoring of proteins to lactic acid bacteria. *Antonie van Leeuwenhoek* **1999**, *76*, 367–376. [PubMed]

22. Proft, T. Sortase-mediated protein ligation: An emerging biotechnology tool for protein modification and immobilisation. *Biotechnol. Lett.* **2009**, *32*, 1–10.

23. Diep, D.B.; Mathiesen, G.; Eijsink, V.G.H.; Nes, I.F. Use of lactobacilli and their pheromone-based regulatory mechanism in gene expression and drug delivery. *Curr. Pharm. Biotechnol.* **2009**, *10*, 62–73. [CrossRef]

24. Boekhorst, J.; De Been, M.W.H.J.; Kleerebezem, M.; Siezen, R.J. Genome-wide detection and analysis of cell wall-bound proteins with LPxTG-like sorting motifs. *J. Bacteriol.* **2005**, *187*, 4928–4934. [CrossRef] [PubMed]

25. Marraffini, L.A.; Dedent, A.C.; Schneewind, O. Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 192–221. [CrossRef] [PubMed]

26. Visweswaran, G.R.; Leenhouts, K.; van Roosmalen, M.; Kok, J.; Buist, G. Exploiting the peptidoglycan-binding motif, LysM, for medical and industrial applications. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 4331–4345. [CrossRef]

27. Turner, M.S.; Hafner, L.M.; Walsh, T.; Giffard, P.M. Identification and characterization of the novel LysM domain-containing surface protein Sep from *Lactobacillus fermentum* BR11 and its use as a peptide fusion partner in *Lactobacillus* and *Lactococcus*. *Appl. Environ. Microbiol.* **2004**, *70*, 3673–3680. [CrossRef] [PubMed]

28. Raha, A.R.; Varma, N.R.S.; Yusoff, K.; Ross, E.; Foo, H.L. Cell surface display system for *Lactococcus lactis*: A novel development for oral vaccine. *Appl. Microbiol. Biotechnol.* **2005**, *68*, 75–81. [CrossRef]

29. Steen, A.; Buist, G.; Leenhouts, K.J.; El Khattabi, M.; Grijpstra, F.; Zomer, A.L.; Venema, G.; Kuipers, O.P.; Kok, J. Cell wall attachment of a widely distributed peptidoglycan binding domain is hindered by cell wall constituents. *J. Biol. Chem.* **2003**, *278*, 23874–23881. [CrossRef]

30. Okano, K.; Zhang, Q.; Kimura, S.; Narita, J.; Tanaka, T.; Fukuda, H.; Kondo, A. System using tandem repeats of the cA peptidoglycan-binding domain from *Lactococcus lactis* for display of both N-and C-terminal fusions on cell surfaces of lactic acid bacteria. *Appl. Environ. Microbiol.* **2008**, *74*, 1117–1123. [CrossRef] [PubMed]

31. Boekhorst, J.; Wels, M.; Kleeberezem, M.; Siezen, R.J. The predicted secretome of *Lactobacillus plantarum* WCFS1 sheds light on interactions with its environment. *Microbiology* **2006**, *152*, 3175–3183. [CrossRef] [PubMed]

32. Kleerebezem, M.; Boekhorst, J.; van Kranenburg, R.; Molenaar, D.; Kuipers, O.P.; Leer, R.; Tarchini, R.; Peters, S.A.; Sandbrink, H.M.; Fiers, M.W.; et al. Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1990–1995. [CrossRef] [PubMed]

33. Mesnage, S.; Tosi-Couture, E.; Fouet, A. Production and cell surface anchoring of functional fusions between the SLH motifs of the *Bacillus anthrasis* S-layer proteins and the *Bacillus subtilis* levansucrase. *Mol. Microbiol.* **1999**, *31*, 927–936. [CrossRef]

34. Mesnage, S.; Weber-Levy, M.; Haustant, M.; Mock, M.; Fouet, A. Cell surface-exposed tetanus toxin fragment C produced by recombinant *Bacillus anthracis* protects against tetanus toxin. *Infect. Immun.* **1999**, *67*, 4847–4850.

35. Bosma, T.; Kanninga, R.; Neef, J.; Audouy, S.A.; van Roosmalen, M.L.; Steen, A.; Buist, G.; Kok, J.; Kuipers, O.P.; Robillard, G.; et al. Novel surface display system for proteins on non-genetically modified gram-positive bacteria. *Appl. Environ. Microbiol.* **2006**, *72*, 880–889. [CrossRef] [PubMed]

36. Fredriksen, L.; Kleiveland, C.R.; Olsen Hult, L.T.; Lea, T.; Nygaard, C.S.; Eijsink, V.G.H.; Mathiesen, G. Surface display of N-terminally anchored invasin by *Lactobacillus plantarum* activates NF-κB in monocytes. *Appl. Environ. Microbiol.* **2012**, *78*, 5864–5871. [CrossRef] [PubMed]

37. Kuczkowska, K.; Mathiesen, G.; Eijsink, V.G.H.; Øynebråten, I. *Lactobacillus plantarum* displaying CCL3 chemokine in fusion with HIV-1 Gag derived antigen causes increased recruitment of T cells. *Microb. Cell Fact.* **2015**, *14*, 1. [CrossRef] [PubMed]

38. Desvaux, M.; Dumas, E.; Chafsey, I.; Hebraud, M. Protein cell surface display in Gram-positive bacteria: From single protein to macromolecular protein structure. *FEMS Microbiol. Lett.* **2006**, *256*, 1–15. [CrossRef]

39. Joris, B.; Englebert, S.; Chu, C.-P.; Kariyama, R.; Daneo-Moore, L.; Shockman, G.D.; Ghuysen, J.-M. Modular design of the *Enterococcus hirae* muramidase-2 and *Streptococcus faecalis* autolysin. *FEMS Microbiol. Lett.* **1992**, *91*, 257–264. [CrossRef]

40. Buist, G.; Steen, A.; Kok, J.; Kuipers, O.P. LysM, a widely distributed protein motif for binding to (peptido) glycans. *Mol. Microbiol.* **2008**, *68*, 838–847. [CrossRef] [PubMed]

41. Mesnage, S.; Dellarole, M.; Baxter, N.J.; Rouget, J.B.; Dimitrov, J.D.; Wang, N.; Fujimoto, Y.; Hounslow, A.M.; Lacroix-Desmazes, S.; Fukase, K.; et al. Molecular basis for bacterial peptidoglycan recognition by LysM domains. *Nat. Commun.* **2014**, *5*, 4269. [CrossRef] [PubMed]

42. Xu, W.; Huang, M.; Zhang, Y.; Yi, X.; Dong, W.; Gao, X.; Jia, C. Novel surface display system for heterogonous proteins on *Lactobacillus plantarum*. *Lett. Appl. Microbiol.* **2011**, *53*, 641–648. [CrossRef] [PubMed]

43. De Man, J.C.; Rogosa, M.; Sharpe, M.E. A medium for the cultivation of lactobacilli. *J. Appl. Bacteriol.* **1960**, *23*, 130–135. [CrossRef]

44. Målbakken, N. Development of a Non-GMO Tuberculosis Vaccine, Using *Lactobacillus* as a Delivery Vehicle. Master's Thesis, Norwegian University of Life Sciences, Ås, Norway, 2014.

45. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]

Bacillus subtilis Lipase A—Lipase or Esterase?

Paula Bracco [1], Nelleke van Midden [1], Epifanía Arango [1], Guzman Torrelo [1], Valerio Ferrario [2], Lucia Gardossi [2] and Ulf Hanefeld [1,*]

[1] Biokatalyse, Afdeling Biotechnologie, Technische Universiteit Delft, Van der Maasweg 9, 2629 HZ Delft, The Netherlands; paulabracco@gmail.com (P.B.); nellekevanmidden@gmail.com (N.v.M.); epifaniarango@gmail.com (E.A.); guzman.torrelo@hotmail.com (G.T.)

[2] Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, Via Licio Giorgieri 1, 34127 Trieste, Italy; valerio.ferrario@gmail.com (V.F.); gardossi@units.it (L.G.)

* Correspondence: u.hanefeld@tudelft.nl; Tel.: +31-15-278-9304

Abstract: The question of how to distinguish between lipases and esterases is about as old as the definition of the subclassification is. Many different criteria have been proposed to this end, all indicative but not decisive. Here, the activity of lipases in dry organic solvents as a criterion is probed on a minimal α/β hydrolase fold enzyme, the *Bacillus subtilis* lipase A (BSLA), and compared to *Candida antarctica* lipase B (CALB), a proven lipase. Both hydrolases show activity in dry solvents and this proves BSLA to be a lipase. Overall, this demonstrates the value of this additional parameter to distinguish between lipases and esterases. Lipases tend to be active in dry organic solvents, while esterases are not active under these circumstances.

Keywords: hydrolase; lipase; esterase; *Bacillus subtilis* lipase A; transesterification; organic solvent; water activity

1. Introduction

Lipases and esterases both catalyze the hydrolysis of esters. This has led to the longstanding question: how can we distinguish between a lipase and an esterase? As the simple hydrolysis of an ester does not suffice, a range of different criteria has been suggested [1–5]. (1) The oldest distinction is the kinetic and structural criterion of interfacial activation, which was already described in 1936 [6]. However, several lipases do not fulfill this; in particular, the much-used *Candida antarctica* lipase B (CALB) does not [7]. (2) Directly linked to the interfacial activation is the lid that covers the active site of many lipases and, via a conformational change, makes the active site more accessible once an interface is present. Again, this is not the case for all lipases [1–3,7,8]. (3) Primary sequence data were shown not to be distinctive enough [2]. (4) Substrates and inhibitors, such as Orlistat, can be utilized to distinguish between esterases and lipases but, again, they are not precise. However, the different substrate ranges are indicative. Esterases tend to be capable of the hydrolysis of water-soluble esters and, in general, short and/or branched side chain esters, while lipases hydrolyze triglycerides, apolar esters, substituted with linear side chains, as well as waxes. This is seen as a reliable but not decisive criterion [2,9]. (5) The activity of the enzyme in the presence of (water-miscible) organic solvents has been proposed as a property of lipases, but other enzymes fulfill this criterion, too [2,10–14]. (6) A parameter already investigated some time ago is the activity of lipases in the absence of water, i.e., in modestly polar, water-non-miscible solvents at very low water activities (a_w). Out of all enzymes tested, only lipases and the closely related cutinases are active at low a_w [1–5,10,15–19]. While not all lipases display this property, it is highly distinctive [20,21].

To probe whether a_w is indeed a suitable parameter to distinguish between lipases and esterases and between lipases and other hydrolases in general, we studied the behavior of *Bacillus subtilis*

lipase A (BSLA) [9]. BSLA is a small (181 amino acids, 19 kDa) serine hydrolase (Figure 1). It is neither interfacially activated nor does it have a lid (criteria one and two) [9,22–24] and sequence data are not conclusive, but it is a minimal α/β hydrolase fold enzyme [9,23,25]. The substrate range clearly qualifies BSLA as a lipase, as does the stability in the presence of solvents [9,26–29]. This stability has even been significantly improved in recent mutational studies and BSLA mutants can be very stable in the presence of water-miscible solvents, such as dimethyl sulfoxide (DMSO), dioxane and trifluoroethanol [30,31]. Studies on BSLA in dry organic solvents are, however, missing. As an experimental parameter, we demonstrate the activity of BSLA in dry toluene. Toluene is not water-miscible and has a logP of 2.5 [32]. It is commonly used in organic synthesis and is highly suitable for lipases and also other enzymes with an α/β hydrolase fold. To date, only lipases were shown to be active in toluene with a very low a_w [1,3].
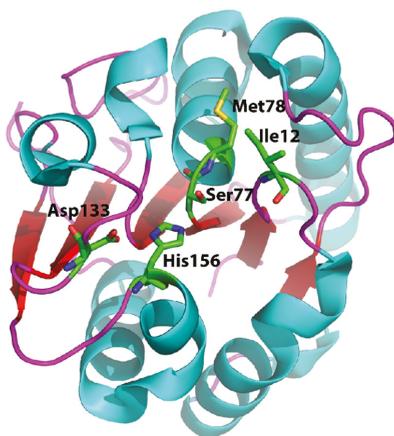

Figure 1. *Bacillus subtilis* lipase A (BSLA) is the smallest serine hydrolase with an α/β hydrolase fold. With only 181 amino acids, it has a molecular weight of 19 kDa. The depicted BSLA structure is pdb 1R50 and the catalytic triad His156, Ser77 and Asp133 and the oxyanion hole Ile12 and Met78 are highlighted. The figure was created with PyMOL.

Additionally, we extend the structural assignment of the hydrolase character with the GRID-based (Fortran program [33]) Global Positioning System in Biological Space (Bio GPS) investigation [33]. BioGPS utilizes surface shape and polarity as criteria. It is neither based on direct sequence comparison, nor on structure superimposition [33,34]. Earlier studies with this method had placed CALB in both the esterase and lipase group. CALB works extremely well in dry solvents and is, therefore, often applied in reactions that require these conditions, such as dynamic kinetic resolutions [1,3,7,35]. On the other hand, it misses interfacial activation and major conformational changes do not take place when CALB comes into contact with an apolar second phase (see above). As such, BioGPS recognized the ambivalence in the assignment of CALB as a lipase well.

Here, we describe the investigation of BSLA by BioGPS and a comparison to other lipases, in particular CALB. We also probe the lipase character of both BSLA and CALB at different a_w. In this manner, new experimental and computational criteria for the esterases and lipases are introduced and investigated.

2. Results

2.1. BioGPS

BioGPS descriptors can be utilized to explore enzyme active site properties and to group them according to their similarities and differences. As such, they can help to explore promiscuous activities.

In an earlier study, the character of CALB was investigated in a set of 42 serine hydrolases. The set contained 11 amidases, nine proteases, 11 esterases and 11 lipases, one of them being CALB [33]. Here, we expand this set with BSLA, utilizing the pdb 1R50 with a resolution of 1.4 Å for the structural information (Table S1). Three probes were used to map specific electrostatic and geometrical active site properties. The O-probe evaluates the H-bond donor properties of the enzymes; the N1 probe, on the contrary, evaluates the H-bond acceptor properties and the DRY probe evaluates the hydrophobic interactions [33]. The DRY probe is clearly of special importance for enzymes that accept hydrophobic substrates, as is the case for lipases.

Considering each property separately, the O-probe located BSLA (pdb 1R50) not among the lipases, but in the amidases cluster, together with a number of esterases (Figure S1a). Equally, the N1-probe (Figure S1b) placed BSLA among the amidases. The DRY probe (Figure S1c), again, placed BSLA amongst the amidases and esterases. This is, in all cases, in contrast to CALB (pdb 1TCA) but it should also be noted that *Candida rugosa* lipase (CRL), a classic lipase with a prominent movement of the lid (criteria one and two) is also always outside the lipase cluster in the different analyses. The previous study ascribed this behavior to the lower hydrophobic nature of the active site of CRL (pdb 1CRL) when compared to the other lipases [8,33].

In the global score, which considers all the mapped properties of the BioGPS together, BSLA can be found firmly among the amidases and esterases (Figure 2), while CALB is in the lipase cluster in the area overlapping with the esterase cluster. CRL, again, is outside the lipase cluster and indeed seems to take a separate position.
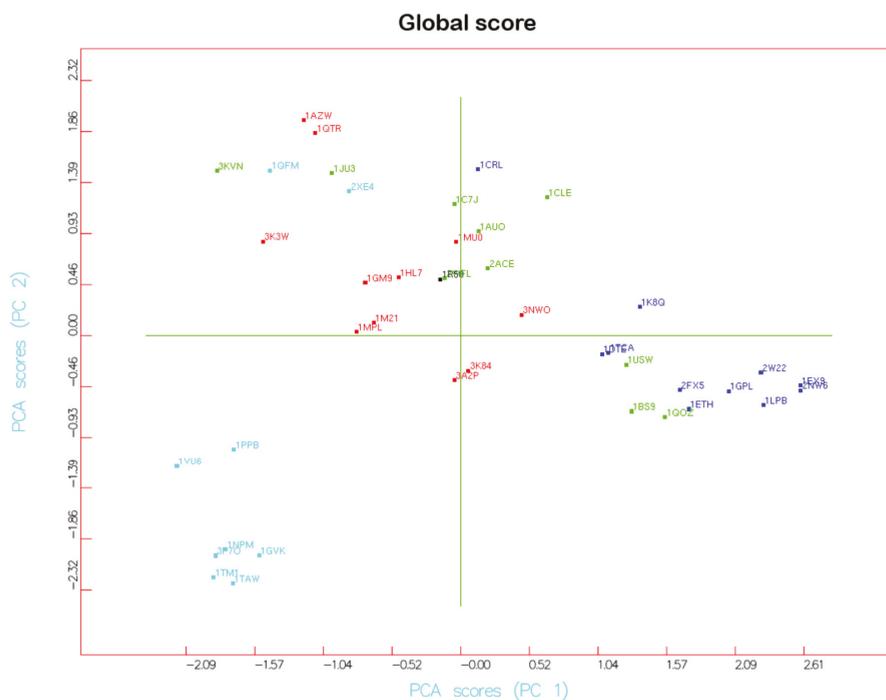
Figure 2. BioGPS of 43 serine hydrolases, for BSLA the data of pdb 1R50 were utilized (global score). Each analyzed enzyme structure is placed within a multidimensional space. Relative distances between each enzyme and all the other enzymes are determined by a statistical principal component analysis. The pdb codes of the processed enzyme structures are indicated in different colors according to their class: lipases in blue, amidases in red, proteases in cyan and esterases in green; the BSLA structure is in black.

BSLA is, according to its substrate range, very clearly a lipase and not an esterase. Amidase activity has to date not been reported for BSLA. While initially surprising, these results also indicate that the study should be extended with an activity assay for amidases.

2.2. Amidase Activity

To probe for amidase activity in BSLA, an amidase activity assay is utilized (Scheme 1). This assay employs benzyl chloroacetamide as standard amide. The released amine reacts with 4-nitro-7-chloro-benzo-2-oxa-1,3-diazole, yielding an adduct that can directly be quantified spectrophotometrically at 475 nm [36].
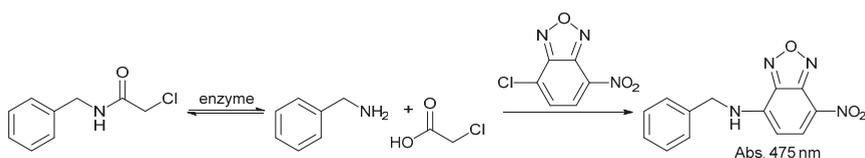
Scheme 1. Amidase activity assay [36]. The assay can be quantified spectrophotometrically. BSLA showed no activity in this assay, ruling out amidase activity.

BSLA showed no activity in this assay. A control experiment with another serine hydrolase, the acyltransferase from *Mycobacterium smegmatis* (*Ms*Act), was performed. This enzyme is an acyltransferase [37,38] and displays promiscuous amidase activity [39,40]. *Ms*Act exhibited activity in this 24 h assay (> 20% conversion of the 5 mM substrate), showing that even minor, promiscuous activities are detectable. This rules out amidase activity for BSLA and supports the earlier assignment of the enzyme as a lipase.

2.3. BSLA Activity in Dry Organic Solvents

To probe the activity of BSLA at low a_w, toluene was used as the solvent and the transesterification of 1-octanol with vinyl acetate was performed as a test reaction (Scheme 2). The use of 1-octanol as a long chain aliphatic compound is a good substrate for lipases [1–5,9] and vinyl acetate is a readily available and widely utilized acyl donor in lipase catalyzed acylation reactions [1,3,41,42]. All reactions were performed with lyophilized BSLA. In parallel, CALB was also tested to ensure direct comparability with one of the most-used lipases. CALB was utilized both as lyophilized enzyme and immobilized as Novozym 435. The latter preparation is most commonly employed, both in the laboratory and on industrial scale [43].
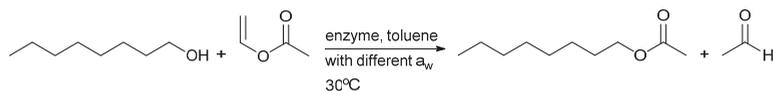
Scheme 2. Test reaction for the activity of BSLA at low a_w. The reaction was performed in toluene at 30 °C, with a ratio of 1-octanol to vinyl acetate of 1:5 and a_w < 0.1, 0.23 and 0.75.

BSLA and CALB were produced by expressing the codon-optimized genes in *E. coli* BL21 (DE3) within pET22b. Subsequent purification gave both enzymes a good purity (Figure 3). With this expression system, both enzymes are not glycosylated. The CALB Novozym 435 produced and immobilized by Novozymes, however, is expressed in *Aspergillus oryzae* and it is, therefore, glycosylated [44].
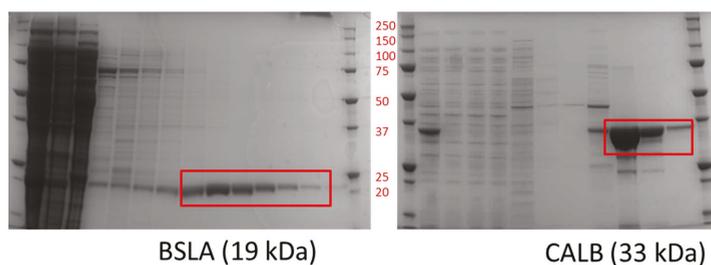
Figure 3. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gels of purified BSLA (19 kDa) and Candida antarctica lipase B (CALB) (33 kDa).

Three different a_w were tested < 0.1 to establish whether BSLA shows the activity in dry solvent only observed for lipases, with a_w = 0.23 as a low value at which most enzymes lose all their activity and a_w = 0.75, an activity at which most enzymes are active [10,21,45,46]. To rigorously ascertain these values of the solvent and reagents, including the internal standard, decane and the enzyme preparations were equilibrated via the vapor phase with dried molecular sieves (activated at elevated temperatures, 5 Å) for a_w < 0.1 [47]. For the other a_w, the enzyme preparations and the other components were equilibrated via the gas phase with an oversaturated solution of potassium acetate (a_w = 0.23) and sodium chloride (a_w = 0.75) [48–52]. For all components, the water content was determined by Karl Fischer titration and equilibrations were considered complete when no changes were observed any more (24–48 h, Table 1). As vinyl acetate was found to negatively affect the Karl Fischer titration, it was freshly distilled and dried with activated molecular sieves for 16 h before use. The activity of the different enzyme preparations was also followed with the tributyrin and *p*-nitrophenol acetate activity assays [2,5,53–56] during equilibration, to establish optimal equilibration times. For BSLA, a small loss of activity over time was observed, while both CALB preparations were stable.

Table 1. Equilibration to different a_w via vapor phase over a saturated solution of salt [47,51] and via the salt pair method [50]. All reaction components, except the acyl donor, were mixed and equilibrated overnight at 30 °C. Finally, dried and freshly distilled vinyl acetate was added in order to start the reaction. The water content was determined by Karl Fischer Titration after 48 h.

a_w	Agent (Vapor Phase or Salt Pair)	Moles of H_2O/mol of Salt	Water Content (ppm)
<0.1	Mol. sieves	0	~20
0.25	NaAc anhydr. (salt pair)	1.5	~180
0.57	Na_2HPO_4 anhydr. (salt pair)	5.0	~360
0.23	KAc (vapor phase)	NA [a]	~120
0.75	NaCl (vapor phase)	NA [a]	~400

[a] Not applicable (NA).

Once reagents and enzymes were equilibrated, the reactions were performed with 100 mM 1-octanol and 500 mM vinyl acetate in previously equilibrated toluene at 30 °C and 1000 rpm (Figure 4). Equal activity of the enzymes (Units) was utilized as determined with the tributyrin activity assay. CALB and, in particular, the well-established commercial preparation of CALB, Novozym 435, performed very well. In both cases, full conversion to 1-octyl acetate was observed. In comparison, BSLA displayed lower conversions (Figure 4). However, the key indicator for a lipase is its activity at low a_w. Here, BSLA and Novozym 435 performed best. For the synthesis of 1-octyl acetate, the trend is a reduction in specific rate at higher a_w (Figure 5). BSLA is very active in dry solvent, as is Novozym 435. Both display lower activities at higher a_w. CALB does not follow this trend.
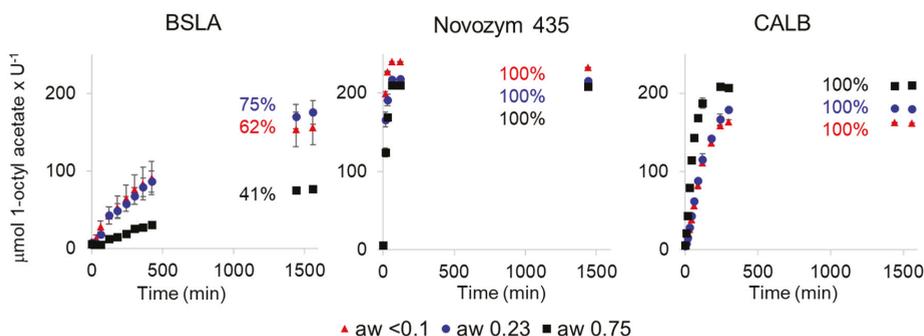
Figure 4. Activity of BSLA, CALB and Novozym 435 in toluene with different a_w. U = μmol butyric acid × min^{-1} in tributyrin activity assay, 0.5–1.2 U of catalyst, 1-octanol (100 mM), vinyl acetate (5 eq.), ISTD: Decane (500 mM), 1 mL reaction volume, 24 h, 30 °C and 1000 rpm. Blanks were performed in the absence of enzyme and showed no conversion. Final conversions are given as inset; the color corresponds to the a_w.
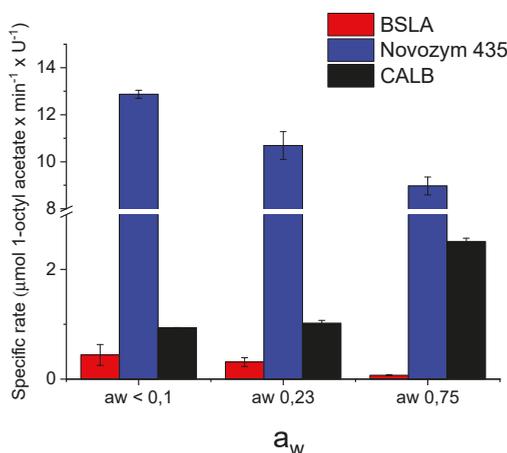
Figure 5. Activity of BSLA, CALB and Novozym 435 in toluene with different a_w. Reaction conditions: 0.5–1.2 U of catalyst, 100 mM 1-octanol, 500 mM vinyl acetate, 500 mM decane (ISTD), in dry toluene (1 mL reaction) at 30 °C and 1000 rpm. U: μmol butyric acid × min^{-1}. Blanks were performed in the absence of enzyme and showed no conversion.

In an earlier study, it had been demonstrated, for different CALB preparations, that this change in activity in the synthesis reaction to 1-octyl acetate can be due to the hydrolysis of the acyl donor vinyl acetate [47]. Therefore, the synthesis reaction at a_w < 0.1 was repeated for BSLA with a 1-octanol to vinyl acetate ratio of 1:1 (Figure 6). Almost the same rate and conversion was observed as with the 1:5 ratio, indicating that, at this low a_w, essentially no hydrolysis occurred, as was the case for Novozym 435, as reported earlier. Overall, these differences in performance at altered a_w can be ascribed to several influences [47,57,58]. Novozyme 435 is an immobilized enzyme and its high activity can be linked to the dispersion of the enzyme on a large surface, promoting its mass transfer and preventing particle aggregation. In contrast, the lyophilized enzymes have a reduced accessibility of the individual enzymes in the preparation. Furthermore, it is well established that immobilized enzymes are better protected against the acetaldehyde that is a side product of the acylation reaction [59]. A

difference in susceptibility to acetaldehyde induced deactivation might also cause the alterations in rate between the two pure enzymes. However, similarly, ionization and water clustering can influence the activity [60,61], leading to these alterations. To demonstrate that the observed effect is general, the experiments were repeated, but this time with BSLA that was dried by co-lyophilization with a salt to establish the desired a_w [62,63]. The enzyme is now in a different environment and two different a_w were established, < 0.1 and 0.57. At < 0.1, very similar results were obtained. Equally, at higher a_w, the ester formation slowed down as before, but could be restarted by adding additional vinyl acetate (Figure 7).

Figure 6. Activity of BSLA, toluene at $a_w < 0.1$. Reaction conditions: 0.5–1.2 U of catalyst, 100 mM 1-octanol, 100 mM or 500 mM vinyl acetate, 500 mM decane (ISTD), in dry toluene (1 mL reaction) at 30 °C and 1000 rpm. U: μmol butyric acid × min^{-1}. Blanks were performed in the absence of enzyme and showed no conversion.
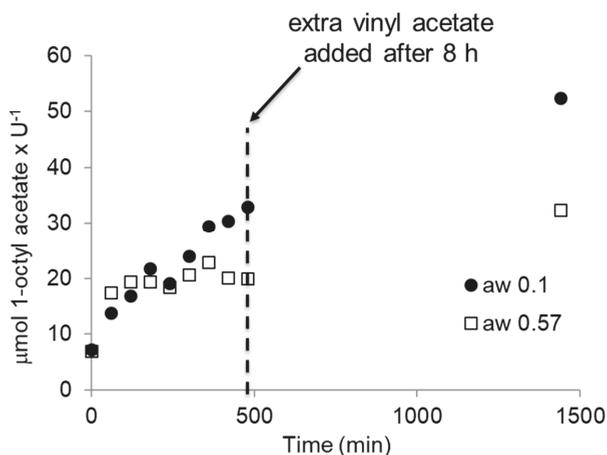
Figure 7. Activity of BSLA co-lyophilized with the appropriate salt, toluene at $a_w < 0.1$ or 0.57. Reaction conditions: 0.5–1.2 U of catalyst, 100 mM 1-octanol, 100 mM vinyl acetate, 500 mM decane (ISTD), in dry toluene (1 mL reaction) at 30 °C and 1000 rpm. U: μmol butyric acid × min^{-1}. Blanks were performed in the absence of enzyme, i.e., in the presence of salt, and showed no conversion. After 8 h (480 min) an additional equivalent of vinyl acetate was added.

To confirm this activity of BSLA (equilibrated via the gas phase) in dry toluene as a general property, the reaction was repeated in dry methyl-t-butyl ether (MTBE) at the same low a_w < 0.1. Enzymes display the same activity in organic solvents when these have the same a_w [64]. Indeed, the BSLA-catalyzed esterification displayed a very similar reaction progress in MTBE and toluene (Figure 8). This confirms the activity of BSLA at low a_w, in line with the earlier observed catalytic activity of CALB at low a_w [47] and of *Rhizomucor miehei* lipase at very low a_w [65].
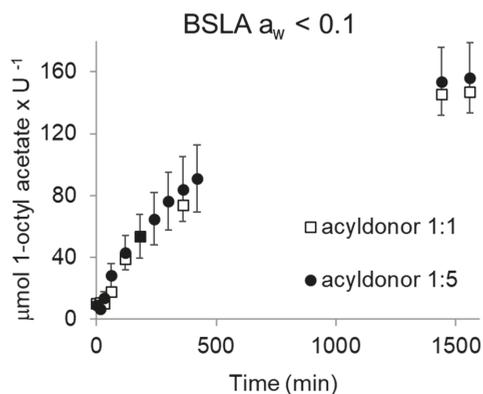

Figure 8. Activity of BSLA, in MTBE and toluene a_w < 0.1. Reaction conditions: 0.5–1.2 U of catalyst, 100 mM 1-octanol, 500 mM vinyl acetate, 500 mM decane (ISTD), in dry solvent (1 mL reaction) at 30 °C and 1000 rpm. U: μmol butyric acid × min^{-1}. Blanks were performed in the absence of enzyme and showed no conversion.
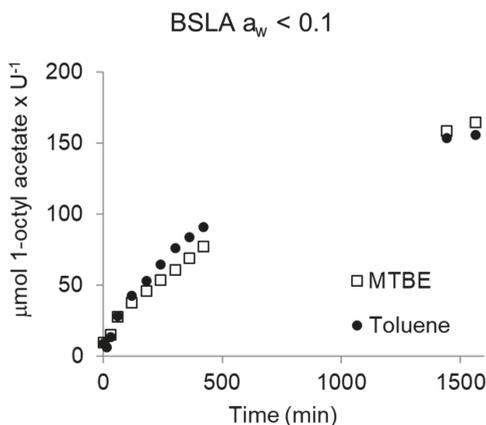
3. Discussion

Interestingly, the BioGPS analysis seems to identify features of the BSLA active site which are shared by other amidase enzymes. In particular, BSLA seems to share similar H-bond capabilities with amidases, as evidenced by the single-probe clustering. The possible promiscuous amidase activity of BSLA was probed with an amidase activity assay (Scheme 1) [36]. This revealed a complete absence of amidase activity. While indicative, this is not conclusive, as this might also be due to substrate specificity. Amidases are characterized by a developed network of H-bond acceptors and donors as described in previous work [33]. The aromatic moiety of the substrate molecule might thus prevent a good interaction with such H-bond/hydrophilic network. More generally, for amidases, the necessity of a hydrogen bond network that stabilizes the NH hydrogen to suppress its deprotonation was reported earlier [66]. Given the very open active site of BSLA, the minimal serine hydrolase with an α/β hydrolase fold, it is not entirely surprising that this type of hydrogen bond network has never been described for this enzyme.

The test of a_w as parameter for the assignment of a serine hydrolase as lipase gave conclusive results. BSLA and CALB displayed good activity at low a_w. In line with earlier results, the synthetic catalytic activity of CALB varies depending on the preparation. Earlier studies had shown that the observed synthetic catalytic activity competes with the hydrolytic activity—that is to say, the parallel hydrolysis of vinyl acetate [47]. This can lead to an apparent decrease in synthetic activity, as is observed for Novozym 435 and BSLA at higher a_w (Figure 5). This trend has already been reported for Novozym 435 [47]. Just like Novozym 435 [47], BSLA displays essentially no hydrolysis of vinyl acetate at low a_w < 0.1. This is confirmed in experiments with a ratio of 1:1 of alcohol to vinyl acetate; a similar rate of synthesis was observed (Figures 6 and 7). The fact that free CALB displays higher synthetic rates at higher a_w is also in line with the literature [47]. It had earlier been demonstrated for

CALB that the ratio of synthesis to hydrolysis depends on the preparation of the enzyme used and that it increases with a_w for purified, free CALB [47]. The activity of BSLA at low a_w was proven also with a different solvent, MTBE (Figure 8).

BioGPS is a complimentary computational tool to investigate the character of an enzyme and delivers a useful input to help us explore the scope of an enzyme more thoroughly. The parameter a_w is an indicative tool to determine whether an enzyme is a lipase or and esterase. Just like the substrate scope, it is not absolute, but is highly indicative. Essentially, a serine hydrolase that is active at low a_w is a lipase and not an esterase, while the reverse statement is not valid. Or, as it was recently summarized: "This long-standing and biased question could be compared to the search for differences between humans and mammals, which implicitly means that one does not consider humans as mammals! Obviously, lipases are a special kind of esterases like humans are a special kind of mammals." [2].

4. Materials and Methods

4.1. Materials

Chemicals and Enzymes

1-propanol, 1-octanol, toluene extra dry, decane, *p*-nitrophenylbutyrate, *p*-nitrophenol, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), tributyrin and 2-methyl-2-propanol, butyric acid and caprylic acid were purchased from Sigma-Aldrich (Schnelldorf, Germany) and Acros (Geel, Belgium), and used without previous purification. Vinyl acetate was purchased from Sigma-Aldrich and distilled before use. Novozym 435 (immobilized lipase B from *Candida antarctica*) was made available by Novozymes (Bagsværd, Denmark). Bovine serum albumin protein and lysozyme from chicken egg whites were purchased from Sigma Aldrich. Bradford reagent was purchased from Biorad (Hercules, C.A., USA). Medium and buffer components were purchased from BD, Merck (Darmstadt, Germany) or J.T. Baker (Geel, Belgium).

Strains and Plasmids

Strains *Escherichia coli* (*E. coli*) HB2151 and *E. coli* HB2151 pCANTAB 5E bsla were kindly provided by Prof. Bauke Dijkstra and Prof. Wim Quax, University of Groningen, the Netherlands. Strains *Escherichia coli* (*E. coli*) BL21 (DE3), *E. coli* TOP10 and plasmid pET22b(+) were utilized for all further work.

4.2. Methods

Cloning pET22bbsla and pET22bcalb

The gene of BSLA (as confirmed by sequencing ID: CP011115.1, range from 292296 to 292841, protein AKCA5803.1) was amplified by PCR from vector pCANTAB 5E BSLA using primers BSLA F: 5'-CCTTTCTATGCGGCCCAGC–3' and BSLA + XhoI R: 5'-CCGCTCGAGCGCCTTCGTATTCTGG-3'. Thereby, restriction site XhoI was introduced for subsequent cloning of BSLA (NcoI, XhoI) into vector pET22b+ (in frame with pelB and His-tag signals). The resulting vector was named pET22bBSLA. The wild type CALB was synthesized by BaseClear (Leiden, The Netherlands). The codon-optimized genes were cloned into pET22b(+) using the NcoI and NotI restriction sites as previously described [67], in order to be in frame with the pelB sequence and a C-terminal His-tag of the plasmid.

Expression and Purification of BSLA

This protocol was adapted from [24]. A freshly grown colony of *E. coli* HB 2151 pCANTAB 5E BSLA was used to inoculate a 1 L shake flask containing 100 mL of 2xTY medium (1.6% *w/v* bactotryptone, 1% *w/v* bacto yeast extract and 0.5% *w/v* sodium chloride), ampicillin (100 µg/mL final concentration) and isopropyl-β-d-galactopyranoside (IPTG, 1 mM final concentration). After 16 h at 28 °C and 150 rpm (Innova Incubator, Hamburg, Germany) the cells were harvested and washed with

10 mM Tris Buffer pH 7.4 and stored at −20 °C. The periplasm isolation protocol was adapted from [67] and consisted of the resuspension of the overexpressed cells in 1 mL of 10 mM Tris buffer pH 8.0 containing sucrose (25 % *w/v*), EDTA (2 mM) and lysozyme (0.5 mg/mL). After incubation on ice for 20 min, 250 μL of 10 mM Tris buffer pH 8.0 containing sucrose (20% *w/v*) and MgCl$_2$ (125 mM) was added. The suspension was centrifuged and the supernatant containing the periplasmic fraction was desalted using a PD10 column (GE, Healthcare, New York, N.Y., USA) to 100 mM potassium phosphate buffer pH 7.4. Afterwards, the solution was shock-frozen with liquid nitrogen and stored at −20 °C for future biocatalysis applications. BSLA was purified mainly from the media. First, proteins were precipitated by adding 50% *v/v* saturated ammonium sulphate (2.8 M final concentration) for 5 h at 4 °C. After centrifugation, the solid fraction was dissolved in 100 mL of 100 mM potassium phosphate buffer pH 7.4, filtered through a 0.45 μm filter and loaded into a 5 mL His-Trap previously equilibrated column (GE Healthcare) using a NGC chromatography system (BIORAD, Hercules, C.A., USA). The loaded proteins were washed with equilibration buffer potassium phosphate (100 mM, pH 7.4) containing 500 mM NaCl and 20 mM imidazole. The His-tagged BSLA was eluted with a linear gradient from 0–100% potassium phosphate (100 mM, pH 7.4) containing 500 mM NaCl and 500 mM imidazole. The progress of the purification was monitored at 280 nm. Fractions containing the target protein (as confirmed by SDS-PAGE and activity assay, 50–60% of the gradient) were combined, concentrated and desalted with a PD-10 column (GE Healthcare) to potassium phosphate buffer (100 mM, pH 7.4). The purified enzyme (68–148 μg/L medium) was aliquoted (2.5 U/vial, units determined by tributyrin assay) and freeze dried for 16 h, −80 °C and stored at −20 °C under nitrogen atmosphere.

Protein Sequence of BSLA-His (AKCA5803.1):

MAAEHNPVVMVHGIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNNGPVLSR FVQKVLDETGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGGANRLTTGKALPGTDP NQKILYTSIYSSADMIVMNYLSRLDGARNVQIHGVGHIGLLYSSQVNSLIKEGLNGGGQNTKALEH HHHHH

Expression and Purification of CALB

An LB-Amp plate (100 μg/mL) was used to freshly grow *E. coli* BL21 (DE3) pET22bCALB from a −80 °C DMSO stock. After incubation at 37 °C for 16 h, a single colony was used to inoculate a 5 mL LB-Amp (100 μg/mL) preculture and grown for 8 h at the same temperature. Large-scale expressions were carried out in 0.5 L of ZYM-5052 media (placed in 2 L shake flask), 2% *v/v* of the preculture was used for inoculation. After 17 h expression at 22 °C and 170 rpm, an optical density (600 nm) of approximately 3 was obtained in all cases. Afterwards, cells were spun down, washed with 10 mM potassium phosphate buffer pH 7.4 and stored at −20 °C. ZYM-5052 medium [68]: The main cultures were grown in ZYM-5052 medium containing 50 mL 50xM (Na$_2$HPO$_4$·12H$_2$O 448 g/L, KH$_2$PO$_4$ 170 g/L, NH$_4$Cl 134 g/L, Na$_2$SO$_4$ 35.5 g/L), 20 mL 50x5052 (100 g/L α-D-lactose, 250 g/L glycerol and 25 g/L glucose dissolved in ddH$_2$O) and 2 mL of MgSO$_4$ solution (1 M in ddH$_2$O) and filled to 1 L with ZY medium (casamino acids 10 g/L—tryptone in this case—yeast extract 5 g/L). Additionally, 0.2 mL of trace element solution was added to the media. The purification of His-tagged CALB was performed from the periplasmic fraction, as described for BSLA in the previous sections.

Protein Sequence of CALB-His (Sequence ID: 4K6G_A):

MALPSGSDPAFSQPKSVLDAGLTCQGASPSSVSKPILLVPGTGTTGPQSFDSNWIPLSTQLGYTPC WISPPPFMLNDTQVNTEYMVNAITALYAGSGNNKLPVLTWSQGGLVAQWGLTFFPSIRSKVDRLMA FAPDYKGTVLAGPLDALAVSAPSVWQQTTGSALTTALRNAGGLTQIVPTTNLYSATDEIVQPQVSNS PLDSSYLFNGKNVQAQAVCGPLFVIDHAGSLTSQFSYVVGRSALRSTTGQARSADYGITDCNPLPAN DLTPEQKVAAAALLAPAAAAIVAGPKQNCEPDLMPYARPFAVGKRTCSGIVTPAAALEHHHHHH

Bradford Assay

Total protein concentration was determined using Bradford reagent in a microtiter plate (MTP) reader format (96 well plates) [69]. Properly diluted samples were mixed with Bradford reagent (5x), incubated at room temperature (RT) for 5 min and the absorbance measured at 595 nm (in triplicate). The calibration curve was carried out using bovine serum albumin protein, as is standard.

Lipase Activity: Tributyrin Assay

A tributyrin assay for determining lipase activity was performed according to the literature [54]. The assay is based on pH change by acid formation when tributyrin is hydrolyzed by the enzyme. *p*-Nitrophenol was used as pH indicator (colorless at pH 5.5 and yellow at pH 7.5) and the acid concentration was determined by a calibration curve with known amounts of butyric acid (from 0 mM to 40 mM). A negative control was performed by adding buffer instead of an enzyme sample. The substrate consumption (0.8 mM initial concentration) was monitored at 410 nm, 30 °C for 15 min, every 38 s by a microtiter plate reader (in 96-well plates, Synergy 2, BioTek, Winooski, V.T., USA). Plates were shaken for 5 s before every read. The different buffers needed for this assay contained 2.5 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS) (pH 7.2), CHAPS (to dissolve acids) and β-cyclodextrin (to dissolve acids into the solution, to increase the linearity). The activity was determined in U, which is equivalent to μmol acid formed per minute. The assays were done in triplicate. For performing this assay with immobilized enzymes, a larger scale (3 mL) in glass vials with a magnetic stirrer was applied. These were placed on a stirring platform and, for Novozym 435, samples (120 μL) were taken over time and placed in a 96-well plate. If desired, the assay can also be performed with Trioctanoin.

Esterase/Lipase Activity: *p*-Nitrophenol Assay

This protocol was adapted from [53] to an MTP reader equipped for 96 well plates. The enzymatic hydrolysis of *p*-nitrophenyl butyrate with the concomitant formation of *p*-nitrophenol was monitored at 405 nm, 37 °C and recorded for 30 min. For this, a calibration curve of *p*-nitrophenol in potassium phosphate buffer (100 mM, pH 7.4) was prepared (levels from 0-500 μM, 200 μL total volume, in triplicate) and control reactions without enzyme extracts were performed. Lyophilized cell-free extract or pure enzymes were re-dissolved in potassium phosphate buffer (100 mM, pH 7.4) (approximately 20–30 mg/mL) and proper dilutions were added into a preheated potassium phosphate buffer (100 mM, pH 7.4) solution containing 3 mM *p*-nitrophenyl butyrate. The esterase activity measured was corrected by subtracting the activity observed in the controls (no enzyme). By definition, one unit of enzyme (U) is equivalent to 1 μmol of *p*-nitrophenol formed per minute.

Karl Fischer Titration

A Metrohm KF Coulometer Karl Fischer titration setup was used, according to the manufacturer's instructions, to determine the water content in ppm. Samples (100 μL) were taken from the solvent and injected into the system in duplicate. In general, master mixes of toluene after equilibration with $a_w < 0.1$, a_w 0.23 and a_w 0.75 contained 20, 120 and 360 ppm respectively. A deviation of 5–10 ppm per sample was observed.

Equilibration of Solvents/Enzymes to the Desired a_w

All materials, reagents, enzymes and solvents used for the biocatalytic reactions were carefully dried and kept under nitrogen atmosphere with molecular sieves (5 Å) at all times. In all cases, the water content was monitored by Karl Fischer titration and as standard parameter compounds with a water content below 100 ppm were considered dry and suitable for the reaction.

Vapor Phase Method

Oversaturated salt solutions and activated molecular sieves were used to equilibrate the solvents and enzymes needed in the transesterification reaction with a desired water activity [47–50]. In the case of working with dry systems, a master mix was prepared including solvent, substrates (without vinyl acetate) and internal standard, all components previously dried with activated molecular sieves

achieving a_w < 0.1. Lyophilized BSLA, Novozym 435 and CALB were dried over silica in desiccators under a vacuum at room temperature (20–25 °C) for 24, 48 and 24 h, respectively. In order to achieve higher water activities, the master mix and the enzymes were equilibrated over saturated salt solutions of potassium acetate (KAc) and sodium chloride (NaCl) at 30 °C for 48 h, resulting in a_w 0.23 and a_w 0.75 at 30 °C, respectively [51]. As exceptions, BSLA and free CALB were equilibrated for a shorter period of only 24 h.

Salt Pairs Method

The protocol was adapted from [62]. The enzymes were lyophilized with anhydrous salts (Na_2HPO_4 or NaAc) in a ratio 1:99 (3 mg pure BSLA or CALB enzyme and 297 mg of the respective salt). For the background reaction (no-enzyme), only lyophilized salts were added. An amount of 10 mg of the co-lyophilized enzyme was added under a nitrogen atmosphere to the previously dried reaction components (except vinyl acetate) and a specific amount of water was introduced under the nitrogen atmosphere. The moles of water added to the reaction mixture were calculated in order to generate the couple of hepta- and dihydrated phosphates in the case of Na_2HPO_4 (5 moles of water per mol of salt, a_w ~ 0.57) and the couple of tri- and anhydrous acetate in the case of NaAc (1.5 moles of water per mol of salt, a_w ~ 0.25) [50]. After overnight equilibration, the last substrate was added (freshly distilled and dry vinyl acetate) to begin the reaction. In the case of the dry system, no water was added (a_w < 0.1).

Transesterification Catalyzed by Lipases in Organic Solvents under Fixed Water Activities

The protocol was adapted from [47]. The reaction conditions included substrates 1-propanol, 1-octanol, 2-octanol or benzylacohol (100 mM), vinyl acetate freshly distilled (1 or 5 equiv. in respect to the initial substrate concentration) and decane (ISTD, 500 mM final concentration). A range of 0.5–1 U of purified enzymes, 1 mg of immobilized Novozym 435 were tested as catalysts. Toluene or methyl-*t*-butyl ether were used as media (1 mL total volume in GC airtight vials). Reactions were carried out for 25 h, 30 °C and 1000 rpm (thermoblock Eppendorf, Hamburg, Germany). Negative controls were run for both substrates in absence of enzyme. All reactions were performed in duplicate and monitored over time by gas chromatography.

Analytics: GC and GC-MS

Gas chromatograph (GC) and gas chromatograph-mass spectrometry (GC-MS) methods were adapted from [47]. Samples were injected in a gas chromatograph (GC-2014, Shimadzu, Kyoto, Japan) equipped with a CP Sil 5 column (50 m × 0.53 mm × 1.0 um). Injector and detector temperatures were set to 340 and 360 °C, respectively. The initial column temperature was set to 35 °C for 5 min, followed by an increase of 15 °C/min up to 60 °C for 0.5 min and 15 °C/min up to 160 °C and hold for 2 min. Finally, a burnout was introduced, 30 °C/min up to 325 °C. The retention times for 1-propanol, vinyl acetate, toluene, decane, 1-octanol and 1-octylacetate were 1.69, 1.87, 6.64, 10.52, 11.19 and 12.79 min, respectively. To confirm the product's structure, samples were also injected in a gas chromatograph-mass spectrometer (GC-MS QP2010s, Kyoto, Japan) equipped with a CP Sil 5 (25 m × 0.25 mm × 0.4 μm). The injector, interface and ion source temperatures were set to 315, 250 and 200 °C, respectively. The retention times for vinyl acetate, 1-octanol and 1-octylacetate were 1.78, 11.57 and 12.93 min, respectively.

Amidase Activity Assay: Hydrolysis of Benzyl Chloroacetamide

This protocol was adapted from the literature [36]. The biocatalysis conditions included a total volume 500 μL in a 2 mL Eppendorf tube containing 5 mM benzyl chloroacetamide (stock solution of 500 mM in THF), 100 μg/mL enzyme (as quantified by Bradford assay), in 25 mM potassium phosphate buffer pH 7.0 with 10 % *v/v* THF. The conversion was carried out for 24 h at 37 °C and 500 rpm. Afterwards, the derivatization of 200 μL of reaction mixture was carried out with 50 μL of NBDCl (20 mM in DMSO) for 1 h, at 37 °C and 500 rpm. UV detection at 475 nm was performed.

BSLA Structure

The structure of BSLA was taken from the Protein Data Bank (PDB). The structure downloaded (1R50) was treated by removing all molecules but the protein chain with the software PyMOL. The thus-generated structure was used for visual inspection with PyMOL as well as input for the BioGPS analysis.

4.3. BioGPS Computational Analysis

The BioGPS analysis and projection was taken from the previous published work. The BSLA BioGPS analysis was performed using the BioGPS software provided by Molecular Discovery Ltd. (Borehamwood, Hertfordshire, UK) by projecting the enzyme according to its active site properties in the previously performed analysis. The identification of the BSLA active site and the calculation of its properties has been performed as previously described. Specifically, FLAPsite was used for automatic active site identification. The active site was mapped using a GRID approach and the resulting computed properties were considered as electrondensity-like fields centered on each atom, which correspond to the so-called pseudo-molecular interaction fields (pseudo-MIFs). Four different properties were mapped: the active site shape (H probe), H-bond donor properties (O probe), H-bond acceptor capabilities (N1 probe), and hydrophobicity (DRY probe). The magnitude of the interaction of the N1 and O probes also includes, implicitly, information about the charge contribution, as these probes already have a partially positive and negative charge, respectively. The pseudo-MIF points were filtered, by means of a weighted energy-based and space-coverage function, and then used for the generation of quadruplets obtained from all possible combinations of the four pseudo-MIF points. Thus, the BSLA active site was described by a series of quadruplets. Finally, BSLA was projected according to its series of quadruples and scored by the previously performed BioGPS analysis.

5. Conclusions

The longstanding question "what differentiates lipases from esterases?" has led to a list of six parameters that are indicative, but not decisive. Here, we have probed BioGPS and a_w as parameters to distinguish between lipases and esterases, utilizing the minimal serine hydrolase with an α/β fold, BSLA, as a test enzyme. While BioGPS has been used successfully to address similar questions earlier, it was not indicative in this case. The clear assignment of BSLA as either esterase or lipase was a challenging task. The high catalytic activity of BSLA at low a_w clearly demonstrated this serine hydrolase to be a lipase. In future studies, activity at low a_w should, therefore, be utilized to support the differentiation of lipases and esterases.

References

1. Bornscheuer, U.T.; Kazlauskas, R. *Hydrolases in Organic Synthesis*; Wiley: Hoboken, NJ, USA, 2005; pp. 61–184.
2. Ben Ali, Y.; Verger, R.; Abousalham, A. Lipases or Esterases: Does It Really Matter? Toward a New Bio-Physico-Chemical Classification. In *Lipases and Phospholipases: Methods and Protocols*; Book Series: Methods in Molecular Biology; Sandoval, G., Ed.; Springer Science+Business Media: New York, NY, USA, 2012; Volume 861, pp. 31–51.
3. Paravidino, M.; Bohm, P.; Gröger, H.; Hanefeld, U. Hydrolysis and Formation of Carboxylic Acid Esters. In *Enzyme Catalysis in Organic Synthesis*; Wiley: Hoboken, NJ, USA, 2012; pp. 249–362.
4. Arpigny, J.L.; Jaeger, K.E. Bacterial lipolytic enzymes: Classification and properties. *Biochem. J.* **1999**, *343*, 177–183. [CrossRef]
5. Jaeger, K.-E.; Kovacic, F. Determination of Lipolytic Enzyme Activities. In *Pseudomonas Methods and Protocols*; Book Series: Methods in Molecular Biology; Filloux, A., Ramos, J.L., Eds.; Springer Science+Business Media: New York, NY, USA, 2014; Volume 1149, pp. 111–134.
6. Holwerda, K.; Verkade, P.E.; De Willigen, A.H.A. Vergleichende Untersuchungen über die Verseifungsgeschwindigkeit einiger Einsäuriger Triglyceride unter Einfluss von Pankreasextrakt. I. Der Einfluss des Verteilungszustandes der Triglyceride auf die Verseifungsgeschwindigkeit. *Rec. Trav. Chim. Pays Bas* **1936**, *55*, 43–57. [CrossRef]
7. Kirk, O.; Christensen, M.W. Lipases from *Candida antarctica*: Unique Biocatalysts from a Unique Origin. *Org. Proc. Res. Develop.* **2002**, *6*, 446–451. [CrossRef]
8. Grochulski, P.; Li, Y.; Schrag, J.D.; Cygler, M. Two conformational states of *Candida rugosa* lipase. *Protein Sci.* **1994**, *3*, 82–91. [CrossRef]
9. Eggert, T.; van Pouderoyen, G.; Pencreac'h, G.; Douchet, I.; Verger, R.; Dijkstra, B.W.; Jaeger, K.-E. Biochemical properties and three-dimensional structures of two extracellular lipolytic enzymes from *Bacillus subtilis*. *Colloids Surf. B Biointerfaces* **2002**, *26*, 37–46. [CrossRef]
10. Adlercreutz, P. Comparison of lipases and glycoside hydrolases as catalysts in synthesis reactions. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 513–519. [CrossRef]
11. Lopes, D.B.; Fraga, L.P.; Fleuri, L.F.; Macedo, G.A. Lipase and esterase—To what extent can this classification be applied accurately? *Ciênc. Tecnol. Aliment. Camp.* **2011**, *31*, 608–613. [CrossRef]
12. Berlemont, R.; Spee, O.; Delsaute, M.; Lara, Y.; Schuldes, J.; Simon, C.; Power, P.; Daniel, R.; Galleni, M. Novel organic solvent-tolerant esterase isolated by metagenomics: Insights into the lipase/esterase classification. *Rev. Argent. Microbiol.* **2013**, *45*, 3–12.
13. Adlercreutz, P. Fundamentals of Biocatalysis in Neat Organic Solvents. In *Organic Synthesis with Enzymes in Non-Aqueous Media*; Carrea, G., Riva, S., Eds.; WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008; pp. 3–24.
14. Serdakowski, A.L.; Dordick, J.S. Activating Enzymes for Use in Organic Solvents. In *Organic Synthesis with Enzymes in Non-Aqueous Media*; Carrea, G., Riva, S., Eds.; WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008; pp. 47–74.
15. Svensson, I.; Wehtje, E.; Adlercreutz, P.; Mattiasson, B. Effects of water activity on reaction rates and equilibrium positions in enzymatic esterifications. *Biotechnol. Bioeng.* **1994**, *44*, 549–556. [CrossRef]
16. Wehtje, E.; Svensson, I.; Adlercreutz, P.; Mattiasson, B. Continuous control of water activity during biocatalysis in organic media. *Biotechnol. Tech.* **1993**, *7*, 873–878. [CrossRef]
17. Ferrario, V.; Pellis, A.; Cespugli, M.; Guebitz, G.; Gardossi, L. Nature Inspired Solutions for Polymers: Will Cutinase Enzymes Make Polyesters and Polyamides Greener? *Catalysts* **2016**, *6*, 205. [CrossRef]
18. Pellis, A.; Ferrario, V.; Zartl, B.; Brandauer, M.; Gamerith, C.; Acero, E.H.; Ebert, C.; Gardossi, L.; Guebitz, G.M. Enlarging the tools for efficient enzymatic polycondensation: Structural and catalytic features of cutinase 1 from *Thermobifida cellulosilytica*. *Catal. Sci. Technol.* **2016**, *6*, 3430–3442. [CrossRef]
19. Pellis, A.; Ferrario, V.; Cespugli, M.; Corici, L.; Guarneri, A.; Zartl, B.; Acero, E.H.; Ebert, C.; Guebitz, G.; Gardossi, L. Fully renewable polyesters via polycondensation catalyzed by *Thermobifida cellulosilytica* cutinase 1: An integrated approach. *Green Chem.* **2017**, *19*, 490–502. [CrossRef]
20. Valivety, R.H.; Halling, P.J.; Peilow, A.D.; Macrae, A.R. Relationship between water activity and catalytic activity of lipases in organic media. Effects of supports, loading and enzyme preparation. *Eur. J. Biochem.* **1994**, *222*, 461–466. [CrossRef]

21. Valivety, R.H.; Halling, P.J.; Peilow, A.D.; Macrae, A.R. Lipases from different sources vary widely in dependence of catalytic activity on water activity. *Biochim. Biophys. Acta Protein Struct. Mol. Enzym.* **1992**, *1122*, 143–146. [CrossRef]

22. Gupta, R.; Gupta, N.; Rathi, P. Bacterial lipases: An overview of production, purification and biochemical properties. *Appl. Microbiol. Biotechnol.* **2004**, *64*, 763–781. [CrossRef]

23. Eggert, T.; van Pouderoyen, G.; Dijkstra, B.W.; Jaeger, K.-E. Lipolytic enzymes LipA and LipB from *Bacillus subtilis* differ in regulation of gene expression, biochemical properties, and three-dimensional structure. *FEBS Lett.* **2001**, *502*, 89–92. [CrossRef]

24. Boersma, Y.L.; Pijning, T.; Bosma, M.S.; van der Sloot, A.; Godinho, L.F.; Dröge, M.J.; Winter, R.T.; van Pouderoyen, G.; Dijkstra, B.W.; Quax, W.J. Loop Grafting of *Bacillus subtilis* Lipase A: Inversion of Enantioselectivity. *Chem. Biol.* **2008**, *15*, 782–789. [CrossRef]

25. van Pouderoyen, G.; Eggert, T.; Jaeger, K.-E.; Dijkstra, B.W. The crystal structure of *Bacillus subtilis* lipase: A minimal α/β hydrolase fold enzyme. *J. Mol. Biol.* **2001**, *309*, 215–226. [CrossRef]

26. Augustyniak, W.; Brzezinska, A.A.; Pijning, T.; Wienk, H.; Boelens, R.; Dijkstra, B.W.; Reetz, M.T. Biophysical characterization of mutants of *Bacillus subtilis* lipase evolved for thermostability: Factors contributing to increased activity retention. *Protein Sci.* **2012**, *21*, 487–497. [CrossRef]

27. Kamal, Z.; Ahmad, S.; Molugu, T.R.; Vijayalakshmi, A.; Deshmukh, M.V.; Sankaranarayanan, R.; Rao, N.M. In Vitro Evolved Non-Aggregating and Thermostable Lipase: Structural and Thermodynamic Investigation. *J. Mol. Biol.* **2011**, *413*, 726–741. [CrossRef]

28. Ahmad, S.; Kamal, Z.; Sankaranarayanan, R.; Rao, N.M. Thermostable *Bacillus subtilis* Lipases: In Vitro Evolution and Structural Insight. *J. Mol. Biol.* **2008**, *381*, 324–340. [CrossRef]

29. Rajakumara, E.; Acharya, P.; Ahmad, S.; Sankaranaryanan, R.; Rao, N.M. Structural basis for the remarkable stability of *Bacillus subtilis* lipase (Lip A) at low pH. *Biochim. Biophys. Acta Proteins Proteom.* **2008**, *1784*, 302–311. [CrossRef]

30. Frauenkron-Machedjou, V.J.; Fulton, A.; Zhao, J.; Weber, L.; Jaeger, K.E.; Schwaneberg, U.; Zhu, L. Exploring the full natural diversity of single amino acid exchange reveals that 40–60% of BSLA positions improve organic solvents resistance. *Bioresour. Bioprocess.* **2018**, *5*, 2. [CrossRef]

31. Markel, U.; Zhu, L.; Frauenkron-Machedjou, V.J.; Zhao, J.; Bocola, M.; Davari, M.D.; Jaeger, K.-E.; Schwaneberg, U. Are Directed Evolution Approaches Efficient in Exploring Nature's Potential to Stabilize a Lipase in Organic Cosolvents? *Catalysts* **2017**, *7*, 142. [CrossRef]

32. Laane, C.; Boeren, S.; Vos, K.; Veeger, C. Rules for optimization of biocatalysis in organic solvents. *Biotechnol. Bioeng.* **1987**, *30*, 81–87. [CrossRef]

33. Ferrario, V.; Siragusa, L.; Ebert, C.; Baroni, M.; Foscato, M.; Cruciani, G.; Gardossi, L. BioGPS Descriptors for Rational Engineering of Enzyme Promiscuity and Structure Based Bioinformatic Analysis. *PLoS ONE* **2014**, *9*, 109354. [CrossRef]

34. Cross, S.; Baroni, M.; Goracci, L.; Cruciani, G. GRID-Based Three-Dimensional Pharmacophores I: FLAPpharm, a Novel Approach for Pharmacophore Elucidation. *J. Chem. Inf. Model.* **2012**, *52*, 2587–2598. [CrossRef]

35. Veum, L.; Kanerva, L.T.; Halling, P.J.; Maschmeyer, T.; Hanefeld, U. Optimisation of the Enantioselective Synthesis of Cyanohydrin Esters. *Adv. Synth. Catal.* **2005**, *347*, 1015–1021. [CrossRef]

36. Henke, E.; Bornscheuer, U.T. Fluorophoric Assay for the High-Throughput Determination of Amidase Activity. *Anal. Chem.* **2003**, *75*, 255–260. [CrossRef]

37. Mathews, I.; Soltis, M.; Saldajeno, M.; Ganshaw, G.; Sala, R.; Weyler, W.; Cervin, M.A.; Whited, G.; Bott, R. Structure of a Novel Enzyme That Catalyzes Acyl Transfer to Alcohols in Aqueous Conditions. *Biochemistry* **2007**, *46*, 8969–8979. [CrossRef] [PubMed]

38. Mestrom, L.; Claessen, J.G.R.; Hanefeld, U. Enzyme-Catalyzed Synthesis of Esters in Water. *ChemCatChem* **2019**, *11*, 2004–2010. [CrossRef]

39. Land, H.; Hendil-Forssell, P.; Martinelle, M.; Berglund, P. One-pot biocatalytic amine transaminase/acyl transferase cascade for aqueous formation of amides from aldehydes or ketones. *Catal. Sci. Technol.* **2016**, *6*, 2897–2900. [CrossRef]

40. Contente, M.L.; Farris, S.; Tamborini, L.; Molinari, F.; Paradisi, F. Flow-based enzymatic synthesis of melatonin and other high value tryptamine derivatives: A five-minute intensified process. *Green Chem.* **2019**, *21*, 3263–3266. [CrossRef]

41. Hanefeld, U. Reagents for (ir)reversible enzymatic acylations. *Org. Biomol. Chem.* **2003**, *1*, 2405–2415. [CrossRef]

42. Paravidino, M.; Hanefeld, U. Enzymatic acylation: Assessing the greenness of different acyl donors. *Green Chem.* **2011**, *13*, 2651–2657. [CrossRef]

43. Ortiz, C.; Ferreira, M.L.; Barbosa, O.; dos Santos, J.C.S.; Rodrigues, R.C.; Berenguer-Murcia, Á.; Briand, E.L.; Fernandez-Lafuente, R. Novozym 435: The "perfect" lipase immobilized biocatalyst? *Catal. Sci. Technol.* **2019**, *9*, 2380–2420. [CrossRef]

44. Basso, A.; Braiuca, P.; Cantone, S.; Ebert, C.; Linda, P.; Spizzo, P.; Caimi, P.; Hanefeld, U.; Degrassi, G.; Gardossi, L. In Silico Analysis of Enzyme Surface and Glycosylation Effect as a Tool for Efficient Covalent Immobilisation of CalB and PGA on Sepabeads®. *Adv. Synth. Catal.* **2007**, *349*, 877–886. [CrossRef]

45. Ma, L.; Persson, M.; Adlercreutz, P. Water activity dependence of lipase catalysis in organic media explains successful transesterification reactions. *Enzym. Microb. Technol.* **2002**, *31*, 1024–1029. [CrossRef]

46. Paravidino, M.; Sorgedrager, M.J.; Orru, R.V.A.; Hanefeld, U. Activity and Enantioselectivity of the Hydroxynitrile Lyase *Me*HNL in Dry Organic Solvents. *Chem. Eur. J.* **2010**, *16*, 7596–7604. [CrossRef]

47. Secundo, F.; Carrea, G.; Soregaroli, C.; Varinelli, D.; Morrone, R. Activity of different *Candida antarctica* lipase B formulations in organic solvents. *Biotechnol. Bioeng.* **2001**, *73*, 157–163. [CrossRef]

48. Valivety, R.H.; Halling, P.J.; Macrae, A.R. Reaction rate with suspended lipase catalyst shows similar dependence on water activity in different organic solvents. *Biochim. Biophys. Acta Protein Struct. Mol. Enzym.* **1992**, *1118*, 218–222. [CrossRef]

49. Fontes, N.; Partridge, J.; Halling, P.J.; Barreiros, S. Zeolite molecular sieves have dramatic acid-base effects on enzymes in nonaqueous media. *Biotechnol. Bioeng.* **2002**, *77*, 296–305. [CrossRef] [PubMed]

50. Halling, P.J. Salt hydrates for water activity control with biocatalysts in organic media. *Biotechnol. Tech.* **1992**, *6*, 271–276. [CrossRef]

51. Greenspan, L. Humidity fixed points of binary saturated aqueous solutions. *J. Res. Natl. Bur. Stand. Sect. A Phys. Chem.* **1977**, *81*, 89–96. [CrossRef]

52. Johnson, J.R.; Affsprung, H.E.; Christian, S.D. The molecular complexity of water in organic solvents. Part II. *J. Chem. Soc. A* **1966**, 77–78. [CrossRef]

53. Gupta, R.; Rathi, P.; Gupta, N.; Bradoo, S. Lipase assays for conventional and molecular screening: An overview. *Biotechnol. Appl. Biochem.* **2003**, *37*, 63–71. [CrossRef]

54. Mateos-Díaz, E.; Rodríguez, J.A.; de los Ángeles Camacho-Ruiz, M.; Mateos-Díaz, J.C. High-Throughput Screening Method for Lipases/Esterases. In *Lipases and Phospholipases: Methods and Protocols*; Book Series: Methods in Molecular Biology; Sandoval, G., Ed.; Springer Science + Business Media: New York, NY, USA, 2012; Volume 861, pp. 89–100.

55. Híreš, M.; Rapavá, N.; Šimkovič, M.; Varečka, Ľ.; Berkeš, D.; Kryštofová, S. Development and Optimization of a High-Throughput Screening Assay for Rapid Evaluation of Lipstatin Production by Streptomyces Strains. *Curr. Microbiol.* **2018**, *75*, 580–587. [CrossRef]

56. Nalder, T.D.; Ashton, T.D.; Pfeffer, F.M.; Marshall, S.N.; Barrow, C.J. 4-Hydroxy-N-propyl-1,8-naphthalimide esters: New fluorescence-based assay for analysing lipase and esterase activity. *Biochimie* **2016**, *128–129*, 127–132. [CrossRef]

57. Corici, L.; Ferrario, V.; Pellis, A.; Ebert, C.; Lotteria, S.; Cantone, S.; Voinovich, D.; Gardossi, L. Large scale applications of immobilized enzymes call for sustainable and inexpensive solutions: Rice husk as renewable alternative to fossil-based organic resins. *RSC Adv.* **2016**, *6*, 63256–63270. [CrossRef]

58. Secundo, F.; Carrea, G. Lipase activity and conformation in neat organic solvents. *J. Mol. Catal. B Enzym.* **2002**, *19–20*, 93–102. [CrossRef]

59. Hanefeld, U.; Gardossi, L.; Magner, E. Understanding enzyme immobilisation. *Chem. Soc. Rev.* **2009**, *38*, 453–468. [CrossRef] [PubMed]

60. Stauch, B.; Fisher, S.J.; Cianci, M. Open and closed states of *Candida antarctica* lipase B: Protonation and the mechanism of interfacial activation. *J. Lipid Res.* **2015**, *56*, 2348–2358. [CrossRef] [PubMed]

61. Banik, S.D.; Nordblad, M.; Woodley, J.M.; Peters, G.H. Effect of Water Clustering on the Activity of *Candida antarctica* Lipase B in Organic Medium. *Catalysts* **2017**, *7*, 227. [CrossRef]

62. Ebert, C.; Gardossi, L.; Linda, P. Control of enzyme hydration in penicillin amidase catalysed synthesis of amide bond. *Tetrahedron Lett.* **1996**, *37*, 9377–9380. [CrossRef]

63. Ru, M.T.; Dordick, J.S.; Reimer, J.A.; Clark, D.S. Optimizing the salt-induced activation of enzymes in organic solvents: Effects of lyophilization time and water content. *Biotechnol. Bioeng.* **1999**, *63*, 233–241. [CrossRef]

64. Partridge, J.; Dennison, P.R.; Moore, B.D.; Halling, P.J. Activity and mobility of subtilisin in low water organic media: Hydration is more important than solvent dielectric. *Biochim. Biophys. Acta Protein Struct. Mol. Enzym.* **1998**, *1386*, 79–89. [CrossRef]

65. Valivety, R.H.; Halling, P.J.; Macrae, A.R. *Rhizomucor miehei* lipase remains highly active at water activity below 0.0001. *FEBS Lett.* **1992**, *301*, 258–260. [CrossRef]

66. Syrén, P.-O.; Hult, K. Amidases Have a Hydrogen Bond that Facilitates Nitrogen Inversion, but Esterases Have Not. *ChemCatChem* **2011**, *3*, 853–860. [CrossRef]

67. Larsen, M.W.; Zielinska, D.F.; Martinelle, M.; Hidalgo, A.; Jensen, L.J.; Bornscheuer, U.T.; Hult, K. Suppression of Water as a Nucleophile in *Candida antarctica* Lipase B Catalysis. *ChemBioChem* **2010**, *11*, 796–801. [CrossRef]

68. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **2005**, *41*, 207–234. [CrossRef] [PubMed]

69. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]

Hydrolysis of Glycosyl Thioimidates by Glycoside Hydrolase Requires Remote Activation for Efficient Activity

Laure Guillotin [1], Zeinab Assaf [1], Salvatore G. Pistorio [2], Pierre Lafite [1], Alexei V. Demchenko [2] and Richard Daniellou [1,*]

[1] Institut de Chimie Organique et Analytique, ICOA, Université d'Orléans, CNRS, UMR 7311, BP6759 Rue de Chartres, F-45067 Orléans CEDEX 2, France; guillotin.laure@gmail.com (L.G.); assaf_zeinab@hotmail.fr (Z.A.); pierre.lafite@univ-orleans.fr (P.L.)

[2] Department of Chemistry and Biochemistry, University of Missouri-St. Louis, One University Boulevard, St. Louis, MO 63121, USA; salvopistorio@hotmail.com (S.G.P.); demchenkoa@umsl.edu (A.V.D.)

[*] Correspondence: richard.daniellou@univ-orleans.fr; Tel.: +33-2-38-49-49-78

Abstract: Chemoenzymatic synthesis of glycosides relies on efficient glycosyl donor substrates able to react rapidly and efficiently, yet with increased stability towards chemical or enzymatic hydrolysis. In this context, glycosyl thioimidates have previously been used as efficient donors, in the case of hydrolysis or thioglycoligation. In both cases, the release of the thioimidoyl aglycone was remotely activated through a protonation driven by a carboxylic residue in the active site of the corresponding enzymes. A recombinant glucosidase (*Dt*Gly) from *Dictyoglomus themophilum*, previously used in biocatalysis, was also able to use such glycosyl thioimidates as substrates. Yet, enzymatic kinetic values analysis, coupled to mutagenesis and in silico modelling of *Dt*Gly/substrate complexes demonstrated that the release of the thioimidoyl moiety during catalysis is only driven by its leaving group ability, without the activation of a remote protonation. In the search of efficient glycosyl donors, glycosyl thioimidates are attractive and efficient. Their utility, however, is limited to enzymes able to promote leaving group release by remote activation.

Keywords: glycoside hydrolase; thioglycosides; biocatalysis

1. Introduction

Enzymes proved to be efficient synthetic tools for the eco-compatible synthesis of many classes of compounds. Non-organic solvents, mild experimental conditions and high regio- or stereo- specificity inherent to biocatalyzed reactions have increased the added value of enzymes in transformation processes, from the laboratory bench to the industrial scale [1]. Moreover, genetic modifications of recombinant enzymes are now powerful tools to easily alter versatility and properties of the engineered proteins. Rational mutagenesis, directed evolution, or even de novo design have dramatically broadened the applicability of enzymes in biocatalysis [2].

In the glycochemistry field, a vast array of carbohydrate-metabolizing enzymes (CAZYmes), including glycoside hydrolases (GH) or glycosyltransferases (GT), has been engineered and used for the chemo-enzymatic synthesis of glycosides [3]. The corresponding methodologies have proved useful in numerous applications ranging from glycosylated natural products to pharmaceuticals [4,5]. However, only few examples in the literature have been describing the use of CAZYmes for the preparation of synthetic thioglycosides that exhibit a sulphur atom linking the glycone and aglycone counterparts instead of more conventional oxygen or nitrogen atoms [6]. Interestingly, when compared to the corresponding *O*-glycosides, *S*-glycosides are highly stable towards enzymatic and acidic hydrolyses.

As a result, thioglycosides have been used as substrate analogues or inhibitors of *O*-GH involved in many diseases including cancer, lysosomal storage disorder, viral and bacterial infections [7,8].

Activated glycosyl donors have been used for a long time, especially in chemoenzymatic synthesis of oligosaccharides [9–11]. In retaining GH, where the stereochemistry of the anomeric carbon is conserved, these activated donors are of high interest because they enable the formation of the glycosyl-enzyme intermediate through the release of the leaving group (Figure 1). This first step is common to all enzymatic activities (hydrolase [12], transglycosidase [13], halogenase [12] and thioligase [14]) because the final outcome of the reaction only depends on the nature of the nucleophile that will attack the glycosyl-enzyme intermediate in the second step. Depending on the reaction and the substrate employed, this step can be rate-determining.
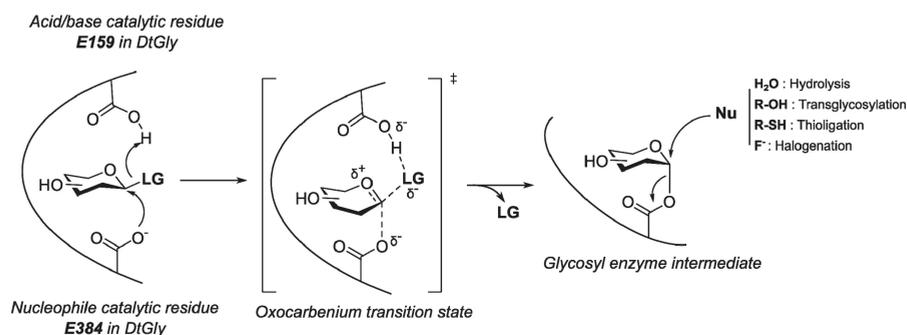
Figure 1. Schematic mechanism of the first step involving the glycosyl-enzyme intermediate formation in retaining GH. The leaving group (LG) release can also be catalysed through another catalytic residue according to its nature. Depending on the nucleophile (Nu) attacking the intermediate, three reactions can take place—hydrolysis, transglycosylation or thioligation.

In addition to the well characterized *O*-glycosides bearing a potent leaving group, some activated *S*-glycosides have been reported as efficient substrates for thioligases [15] or glycoside hydrolases [16]. This latter hydrolytic activity is peculiar as very few examples of *S*-glycosides hydrolysis by glycoside hydrolases have been reported in literature [16–26]. Among those examples, putting aside GlcNAcase, GH4 and myrosinase that do not operate through the canonical GH mechanism, only almond β-glucosidase GH1, *Aspergillus niger* GH3 [16,22,27], *Micromonospora viridifaciens* sialidase [21], *Caldocellum saccharolyticum* glucosidase [24] and *Oryza sativa* Os4BGlu12 [23] have been isolated and identified as thioglycoside hydrolases (Table 1).

Table 1. Comparison of *S*- and *O*-glycoside hydrolysis by GH.

Enzyme	Organism	Substrates Tested	Relative Activity *S*-vs. *O*-(%)	Ref
β-D-Glucosidase	Sweet almond	pNPSGlc	0.13 [a]	[22]
		pNPSGal	0.07 [a]	
		pNPSFuc	0.06 [a]	
		GlcSBiz	80 [a]	[16]
		Glc*S*(*N*-Me)Biz	10 [a]	
		GlcSBox	5 [a]	
β-D-Glucosidase	*A. niger*	GlcSBiz	5 [a]	[16]
		Glc*S*(*N*-Me)Biz	1 [a]	
Sialidase	*M. viridifaciens*	Substituted pNPSNeuAc	0.01–60	[21]
Os4BGlu12	*O. sativa*	pNPSGlc	0.5 [a]	[23]
		OctylSGlc	0.1 [b]	

[a] ratio of k_{cat}/K_M for thioglycoside substrate vs. corresponding *para*-nitrophenyl glycoside. [b] ratio of k_{cat}/K_M for octyl *S*-glucoside vs. octyl *O*-glucoside.

In most cases, *S*-substrate hydrolysis is much less efficient than the rate observed for the corresponding *O*-substrate. Indeed, thioglycosides are less efficient substrates because no general acid/base catalysis is available [28]. Yet, a new class of reactive thioglucosides (Figure 2, Table 1) bearing a thioimidoyl moiety was reported, which were efficiently hydrolysed by almond GH1, as well as *A. niger* GH3 [16]. In both cases, the authors demonstrated that benzoxazolyl 1-thio-β-D-glucopyranoside (GlcSBox) and benzimidazolyl 1-thio-β-D-glucopyranoside (GlcSBiz) hydrolyses were catalysed by remote activation of the C-S bond through protonation of the ring nitrogen in the aglycone. Such remote activation was also described in the case of Araf51 [15], which was able to use similar arabinofuranosyl thioimidates as glycosyl donors in thioglycoligation reaction [29–35]. In the context of chemoenzymatic synthesis of glycosides, these substrates are attractive because of their high stability towards chemical hydrolysis in aqueous solutions, as well as efficient leaving group ability [15].

Figure 2. Substrates used in this study.

In this work, we demonstrated that *Dt*Gly, a GH previously used in chemoenzymatic synthesis of *O*-glycosides, was able to hydrolyse these glycosyl thioimidates. Combined in vitro enzymatic analysis with in silico modelling of the enzyme-substrate interaction have helped us to decipher the molecular mechanism of this rare hydrolysis.

2. Results

2.1. DtGly Can Hydrolyze Thioglycosides

*Dt*Gly (uniprot B5YCI2_DICT6) is encoded by *dicth_0359* gene in the thermophile *Dictyoglomus thermophilum* genome. We have recently reported the cloning, expression and purification of this protein [36]. As many other *D. thermophilum* proteins [37–41], *Dt*Gly was found to be thermostable and also exhibited a wide substrate specificity, as it is able to hydrolyse *p*NP β-D-glucoside, *p*NP β-D-galactoside and *p*NP β-D-fucoside. Moreover, our previous study demonstrated that *Dt*Gly could be used in chemoenzymatic synthesis of glycosides, thereby serving as an attractive biocatalyst that needed to be assayed for other substrates [36].

In this context, we have focused on thioglycoside hydrolysis, as few examples of *S*-GH are available in literature. Three *S*-containing substrates were tested, namely GlcSBiz, GlcSBox and GlcSTaz that bear benzimidazolyl, benzoxazolyl and thiazolinyl aglycones, respectively (Figure 2).

Unlike *p*NP-Glc, wherein the hydrolysis can be easily monitored by quantification of the released *p*NP group, hydrolysis rates of the *S*-containing substrates were determined by quantification of the released glucose. This was achieved by monitoring *o*-dianisidine oxidation enzymatically coupled to glucose production [42]. This methodology applied to *p*NP-Glc hydrolysis gave similar kinetic values to those previously reported using *p*NP quantification (*data not shown*).

All three *S*-containing substrates were hydrolysed by *Dt*Gly (Table 2), with K_M values higher but in the same order of magnitude, as those observed for *p*NP-Glc (2- to 5-fold increase). However, the catalytic rate k_{cat} was decreased by one order of magnitude, indicating that the reaction is dramatically slowed in the case of *S*-containing substrates. Therefore, the catalytic efficiencies of

GlcSBiz, GlcSBox and GlcSTaz were found to be 20 to 40 times lower than the value determined for pNP-Glc.

Table 2. Kinetic parameters of WT and acid/base E159Q mutant of *Dt*Gly. pNP-Glc hydrolysis activity was measured by pNP release quantification. Other substrate hydrolysis activities were determined by quantification of the released glucose. All experiments were done in three independent replicates and are expressed as Mean ± SD.

Enzyme	Substrate	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1}.mM^{-1})
WT	pNP-Glc [a]	460 ± 40	31 ± 0.7	67
WT	GlcSBiz	1533 ± 114	0.23 ± 0.01	0.15
WT	GlcSBox	2246 ± 289	0.38 ± 0.03	0.17
WT	GlcSTaz	880 ± 52	0.31 ± 0.01	0.35
E159Q	pNP-Glc	200 ± 20	0.20 ± 0.01	1.0
E159Q	GlcSBox	445 ± 40	0.06 ± 0.01	0.13

[a] Previously reported data [36].

GlcSBiz, GlcSBox and GlcSTaz have previously been used as substrates for sweet almond and *A. niger* β-glucosidases [16], yet with a much different behaviour. GlcSBiz was hydrolysed by this enzyme as efficiently as pNP-Glc. Kinetics analysis proved that GlcSBiz was efficiently hydrolysed by those glucosidases thanks to the remote protonation of the imidazole ring nitrogen. A much lower activity was observed for GlcSBox and no activity could be observed for GlcSTaz.

To better understand the chemistry underlying the thioglucoside hydrolysis by *Dt*Gly, we first investigated whether these substrates were efficiently binding in the active site, because low GH activities can arise from a second binding mode of substrates, as already reported [43]. Inhibition of pNP-Glc hydrolysis by GlcSBiz demonstrated that the latter is a competitor in the active site to pNP-Glc (Figure 3). Moreover, it efficiently binds into the active site, as an inhibitory constant K_i of 177 ± 11 µM was calculated from the inhibition curves.
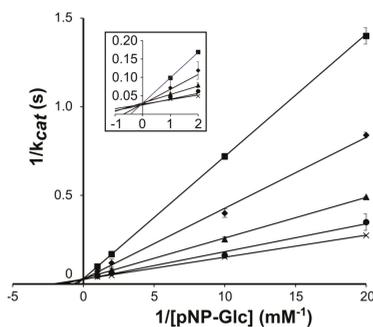

Figure 3. Lineweaver plot of wild-type *Dt*Gly inhibition with increasing concentrations of GlcSBiz. Data are expressed as mean ± SD from three independent experiments. Inhibitor concentrations are respectively depicted as crosses (0 µM), circles (100 µM), triangles (250 µM), diamonds (500 µM) and squares (1000 µM). Inset: 2X zoom on axes origin highlighting the intersection of fitted lines on *y*-axis.

2.2. Identification of Residues Surrounding the Thioglycoside Substrates in DtGly Active Site

Structural analysis of *Dt*Gly was carried out to identify potential residues that might be involved in *S*-containing substrate hydrolysis mechanism. Despite our efforts to crystallize *Dt*Gly, no diffracting crystal could be obtained, thus we decided to build a homology model of the enzyme. To do so, a 3D structure of β-glycosidase from *Pyrococcus horikoshii* was chosen because of its high sequence identity (resp. homology) with *Dt*Gly of 45% (resp. 63%) [44].

An initial model of residues 2-416 was built using a ModWeb server (ModPipe Protein Quality Score of 1.6, considered as reliable); missing residues were then added and the overall model was equilibrated by several cycles of energy minimization and molecular dynamics (Figure 4A).
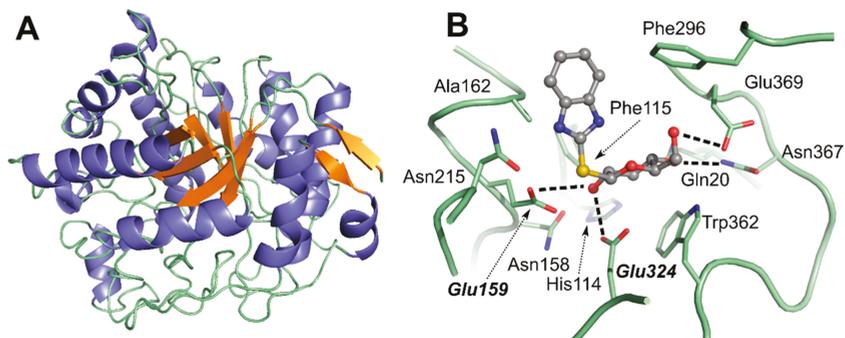

Figure 4. (**A**) Overall representation of *Dt*Gly model. Helices and sheets are respectively coloured in blue and orange. (**B**) Model of docked GlcSBiz in *Dt*Gly active site. Residues surrounding the ligand binding pocket are depicted as sticks. For clarity purposes, hydrogens are not represented. Catalytic residues Glu159 (acid/base) and Glu324 (nucleophile) are highlighted in bold. H-bonds are indicated as dashed lines.

In order to evaluate potential roles of active site residues in sulphur-containing substrate hydrolysis, modelling of substrate-bound *Dt*Gly were done by molecular docking. Using the conformation of glucosides in other closely related GH1 x-ray structures (β-glucosidase from *Thermotoga maritima* PDB 1OIM and 1OIF [45]), GlcSBiz, GlcSBox and GlcSTaz were independently docked into the *Dt*Gly active site. Figure 4B depicts the residues surrounding GlcSBiz, as well as the network of H-bonds between the sugar moiety and several polar residues (Gln20, Glu159, Glu324, Asn367, Glu369). An additional H-π interaction between glucose and Trp362 is also visible, as already seen for other GH [40]. The same interactions were found for other substrates or conformations (see the Supplementary Materials).

In the context of identification of potential residues involved in the *S*-glycoside activation during hydrolysis this model confirms that no acidic residue except Glu159 was close enough to remotely protonate aglycone moieties of substrates, as expected considering in vitro assays.

2.3. DtGly Hydrolysis of S-Glycosides Does Not Involve General Acid/Base Catalysis

In our model, the catalytic glutamate Glu159, which acts as the acid/base residue in retaining GH mechanism [46], is the only one close enough to activate thioglycoside hydrolysis. Although direct protonation on sulphur cannot occur in the case of thiogycosides [28], examples of distant protonation of the aglycone by a catalytic residue have been reported [15]. We have thus generated two mutants, namely Glu159Ala and Glu159Gln to assess the potential role of this residue in the thioglycoside hydrolysis. Unlike Glu159Gln that could be purified to homogeneity, Glu159Ala mutant was not soluble after cell lysis and thus could not be purified. This mutant was left aside for further experiments.

Glu159Gln mutation led to a dramatic decrease of catalytic efficiency for *p*NPGlc, as shown in Table 2. K_M values for this substrate are lower but in the same order of magnitude (200 μM vs. 460 μM), which can be explained by conservation of the active site structure in the mutant and decreased k_{cat} value by a factor of 150 because K_M is related to k_{cat}. This loss of hydrolytic activity upon acid/base mutation is usual, as reported in many other studies, especially those concerning thioligase generation [14,40].

When using GlcSBox as a substrate, *Dt*Gly Glu159Gln exhibits a reduced k_{cat} value (0.38 to 0.06 s^{-1}), as expected because nucleophile water attack is not activated by deprotonation. However, second order rate constant remains unchanged, indicating that the first step of the reaction is not compromised

by the removal of the glutamate residue. Thus, the release of the thiol leaving group is not activated by Glu159 and is only dictated by its leaving capability (i.e., pKa).

3. Discussion

We have previously used *D. thermophilum* *Dt*Gly as a versatile tool for synthesis of glycosides and looked for alternate substrates for this enzyme. Thioimidate glycosyl donors have been used for a long time in organic synthesis to generate a wide range of glycosides and glycans [34,47–49]. In this context, we tested previously reported glycosyl thioimidates as substrates for almond GH1 and *A. niger* GH3 [16]. Examples of cloning and characterization of thioglycoside hydrolases are scarcely available in literature and even fewer studies on mechanism underlying the thioglycoside hydrolysis by GH have been published.

*Dt*Gly is able to hydrolyse *S*-glycosides, with lower activities than those observed for *O*-glycosides. This hydrolytic activity is rate-limited by release of the thiol-containing leaving group and not water nucleophilic attack, unlike generally accepted mechanism for the *O*-glycoside hydrolysis [19,28]. The modelling of substrate-*Dt*Gly complexes as well as mutagenesis of the acid/base residue also demonstrated that no residue was able to remotely protonate the benzimidazole group nor the sulphur atom. If *Dt*Gly is able to hydrolyse *S*-containing substrates without general acid catalysis, the hydrolysis rate is limited by the leaving group capability, as no remote activation is possible. The pKa of leaving groups 2-mercaptobenzimidazole (for GlcSBiz) and 2-mercaptobenzoxazole (for GlcSBox) have been experimentally determined at 5.8 [50] and 6.58 [51]. To our knowledge, no value is available for 2-mercaptothiazoline (for GlcSTaz).

In the case of almond and *A. niger* glycosidases, a remote protonation occurring on a nitrogen atom of the benzimidazole moiety of GlcSBiz was shown to accelerate the leaving group release, thus increasing the catalytic rate to a value close to those observed for *O*-glycosides. Another GH exhibiting thioglycosidase activity on 2'-thio-benzimidazolyl arabinosides activated by remote deprotonation, namely Ara*f*51 [15], was also reported. The modelling of Ara*f*51/substrate complex demonstrated that the nucleophile catalytic residue was responsible for the remote protonation on imidazole nitrogen, mostly because a furanosyl ring is much more flexible than a pyranosyl ring and allows the nucleophile residue to interact with the aglycone.

4. Materials and Methods

4.1. Materials

para-Nitrophenyl β-D-glucopyranoside (*p*NP-Glc) was purchased from Carbosynth (Oxford, UK). 2-benzoxazolyl 1-thio-β-D-glucopyranoside (GlcSBox) [52], 2-benzimidazolyl 1-thio-β-D-glucopyranoside (GlcSBiz) [16] and 2-thiazolinyl 1-thio-β-D-glucopyranoside (GlcSTaz) [49] were prepared as previously described. Otherwise specified, all other chemicals were purchased from Thermo Fisher Scientific (Waltham, MA, USA) and were of purest quality available. Mutagenic primers were purchased from Eurofins Genomics (Ebersberg, Germany) and WT *Dt*Gly coding expression plasmid (pET28a-*dtgly*) was prepared as previously described [36].

4.2. Production of WT and E159Q DtGly

pET28a-*dtgly* was used as a template for mutagenic PCR in the Quikchange Site-directed mutagenesis kit (Agilent, Les Ulis, France). Primers containing desired mutation on acid/base residue position (E159) were constituted of a pair of complement oligonucleotides designed using Agilent tools website (www.genomics.agilent.com, mutated codons are highlighted in bold): *Dt*Gly E159A: 5'-gaattactggatgactataaat**gcg**cccaatgcttatgcctttt-3' and *Dt*Gly E159Q: 5'-atcttgtgaattactggatgactataaat**cag**cccaatgcttatg-3'. Mutagenesis procedure was performed according to the kit procedure. Sequences of pET28a-*dtgly*E159A and pET28a-*dtgly*E159Q were verified by Sanger sequencing at Eurofins Genomics (Ebersberg, Germany).

Production and purification of *Dt*Gly variants was done as previously reported [36]. Briefly, *Escherishia coli* Rosetta(DE3) transformed with expression plasmids were grown in LB medium supplemented with chloramphenicol (34 μg/mL) and kanamycin (30 μg/mL) at 37 °C until OD_{600} reached 0.6. Induction was then done by addition of 1 mM IPTG and incubated overnight a 25 °C. Cells were harvested, lyzed by freeze-thaw cycles and sonication and supernatant was clarified by heat treatment for 15 min at 70 °C before centrifugation. Finally, supernatant was loaded on a Nickel column (HisPure, Thermo Scientific) and purified by elution with lysis buffer containing 500 mM imidazole.

4.3. pNP Release Quantification Assay

*Dt*Gly variants (WT or mutants) activity towards *p*NP-Glc hydrolysis was determined at 37 °C in a 200 μL incubation containing 5 ng of the enzyme, 0.01–10 mM *p*NP-Glc, Citrate-Phosphate buffer (20 mM, pH 6) and 0.1–1 mM GlcBox for inhibition studies. After 20 min of incubation, 100 μL of Na_2CO_3 1 M was added and released *p*NP was quantified at 405 nm (ε_{405} = 19,500 $cm^{-1}.M^{-1}$). Prism 4 (GraphPad) was used to fit data according to Michaelis-Menten model, or competitive inhibition model and retrieve kinetic parameters.

4.4. Glucose Release Assay

To determine Glc*S*Box, Glc*S*Taz and Glc*S*Biz hydrolysis rate by DtGly variants, produced glucose was quantified using a continuous coupled enzyme assay [42]. Incubations were similar as those for *p*NP-Glc hydrolysis with the addition of glucose oxidase from *Aspergillus niger* (Sigma-Aldrich, Saint Louis, MO, USA, 0.4 u), horseradish peroxidase (Sigma-Aldrich, 0.4 u) and *o*-dianisidine (Sigma-Aldrich, 100 μM). Dianisidine oxidation coupled to glucose production was monitored at 442 nm during 30 min. Prism 4 (GraphPad) was used to fit data according to Michaelis-Menten model and retrieve kinetic parameters.

4.5. Computational Studies

The structure of β-glycosidase from Pyrococcus horikoshii [44] (PDB 1VFF, 45%/63% sequence identity/homology) was used as a template for homology model building using a ModWeb server from the A. Sali Laboratory (https://modbase.compbio.ucsf.edu/modweb/). The resulting model was prepared with AmberTools [53] and equilibrated using NAMD software [54] and Amber fb15 force field [55] (3 cycles of 10,000 minimization steps and 0.5 ns dynamics at 100 K).

Docking of Glc*S*Box, Glc*S*Taz and Glc*S*Biz substrates into *Dt*Gly active site model was done by firstly applying AM1-BCC charges on ligands [56]. Then each substrate was placed 10 Å away, facing the active site (according to PDB 1VFF). *Dt*Gly-substrate complexes were formed using steered molecular dynamics [57] at 100 K using the structural alignment of glucose moiety in its binding pocket as the final orientation according to closest structures bearing a ligand in their active site (β-glucosidase from *Thermotoga maritima* PDB 1OIM, 1OIF and 1OIF) [45]. *Dt*Gly backbone was kept constrained during the whole procedure. Finally, protein-ligand complexes models were equilibrated by releasing substrate constraints and applying several cycles of energy minimization (10,000 steps, steepest descent) followed by molecular dynamics (100 K, 1 ns). Final complex models were obtained by a final energy minimization. For each substrate, several initial conformations were tested, mostly by rotation of the glycosidic bond. All structural figures were drawn using the PyMOL Molecular Graphics system 1.8 (www.pymol.org).

5. Conclusions

This study demonstrates that glycosyl thioimidates are not universal glycosyl donors for chemoenzymatic syntheses. While the above examples of efficient enzymatic activities using such substrates were reported in literature, they rely on an activation by protonation of the aglycone moiety, either with a distant carboxylic acid residue (almond GH1) or the catalytic nucleophile (Ara*f*51, Figure 5). *Dt*Gly seems to be the paradigm of the general case of an enzyme that can use those substrates

without acid catalysis, yet with a much lower activity. This study paves the way for broadening *Dt*Gly applications in biocatalysis. Identification of efficient substrates and mutation into a thioligase are currently under further investigation.
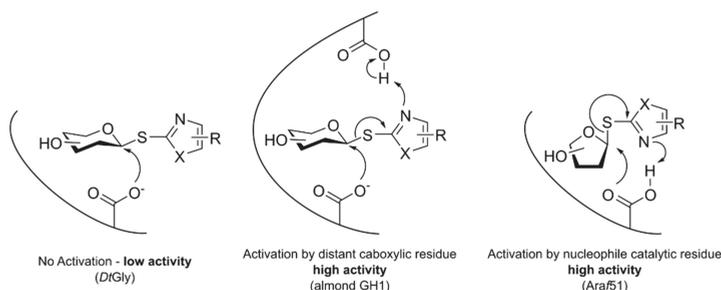

Figure 5. Glycosyl thioimidates require remote activation to promote the leaving group release.

Author Contributions: Conceptualization, R.D. and A.V.D.; Enzymatic studies, L.G., P.L. and Z.A.; data curation and modelling, P.L.; chemical synthesis of substrates, S.G.P. writing—original draft preparation, P.L.; writing—review and editing, P.L., R.D. and A.V.D.; project administration, R.D.; funding acquisition, R.D. and A.V.D.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Sheldon, R.A.; Woodley, J.M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **2018**, *118*, 801–838. [CrossRef] [PubMed]
2. Quin, M.B.; Schmidt-Dannert, C. Engineering of Biocatalysts: From Evolution to Creation. *ACS Catal.* **2011**, *1*, 1017–1021. [CrossRef] [PubMed]
3. Hancock, S.M.; Vaughan, M.D.; Withers, S.G. Engineering of glycosidases and glycosyltransferases. *Curr. Opin. Chem. Biol.* **2006**, *10*, 509–519. [CrossRef] [PubMed]
4. Li, T.-L.; Liu, Y.-C.; Lyu, S.-Y. Combining biocatalysis and chemoselective chemistries for glycopeptide antibiotics modification. *Curr. Opin. Chem. Biol.* **2012**, *16*, 170–178. [CrossRef] [PubMed]
5. Ati, J.; Lafite, P.; Daniellou, R. Enzymatic synthesis of glycosides: From natural O- and N-glycosides to rare C- and S-glycosides. *Beilstein J. Org. Chem.* **2017**, *13*, 1857–1865. [CrossRef]
6. Guillotin, L.; Lafite, P.; Daniellou, R. Chapter 10 Enzymatic thioglycosylation: Current knowledge and challenges. In *Carbohydrate Chemistry*; The Royal Society of Chemistry: London, UK, 2014; Volume 40, pp. 178–194. ISBN 978-1-84973-965-8.
7. Driguez, H. Thiooligosaccharides as Tools for Structural Biology. *Chembiochem* **2001**, *2*, 311–318. [CrossRef]
8. Wardrop, D.J.; Waidyarachchi, S.L. Synthesis and biological activity of naturally occurring α-glucosidase inhibitors. *Nat. Prod. Rep.* **2010**, *27*, 1431–1468. [CrossRef]
9. O'Neill, E.C.; Field, R.A. Enzymatic synthesis using glycoside phosphorylases. *Carbohydr. Res.* **2015**, *403*, 23–37. [CrossRef]
10. Lairson, L.L.; Henrissat, B.; Davies, G.J.; Withers, S.G. Glycosyltransferases: Structures, Functions and Mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521–555. [CrossRef]
11. De Bruyn, F.; Maertens, J.; Beauprez, J.; Soetaert, W.; De Mey, M. Biotechnological advances in UDP-sugar based glycosylation of small molecules. *Biotechnol. Adv.* **2015**, *33*, 288–302. [CrossRef]
12. Rye, C.S.; Withers, S.G. Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **2000**, *4*, 573–580. [CrossRef]

13. Wang, L.-X.; Huang, W. Enzymatic transglycosylation for glycoconjugate synthesis. *Curr. Opin. Chem. Biol.* **2009**, *13*, 592–600. [CrossRef] [PubMed]

14. Jahn, M.; Marles, J.; Warren, R.A.J.; Withers, S.G. Thioglycoligases: Mutant Glycosidases for Thioglycoside Synthesis. *Angew. Chem. Int. Ed.* **2003**, *42*, 352–354. [CrossRef] [PubMed]

15. Almendros, M.; Danalev, D.; François-Heude, M.; Loyer, P.; Legentil, L.; Nugier-Chauvin, C.; Daniellou, R.; Ferrières, V. Exploring the synthetic potency of the first furanothioglycoligase through original remote activation. *Org. Biomol. Chem.* **2011**, *9*, 8371–8378. [CrossRef]

16. Avegno, E.A.-B.; Hasty, S.J.; Parameswar, A.R.; Howarth, G.S.; Demchenko, A.V.; Byers, L.D. Reactive thioglucoside substrates for β-glucosidase. *Arch. Biochem. Biophys.* **2013**, *537*, 1–4. [CrossRef] [PubMed]

17. Goodman, I.; Fouts, J.R.; Bresnick, E.; Menegas, R.; Hitchings, G.H. A Mammalian Thioglycosidase. *Science* **1959**, *130*, 450–451. [CrossRef]

18. Meulenbeld, G.H.; Hartmans, S. Thioglucosidase activity from *Sphingobacterium sp.* strain OTG1. *Appl. Microbiol. Biotechnol.* **2001**, *56*, 700–706. [CrossRef] [PubMed]

19. Macauley, M.S.; Stubbs, K.A.; Vocadlo, D.J. O-GlcNAcase catalyzes cleavage of thioglycosides without general acid catalysis. *J. Am. Chem. Soc.* **2005**, *127*, 17202–17203. [CrossRef] [PubMed]

20. Cetinbaş, N.; Macauley, M.S.; Stubbs, K.A.; Drapala, R.; Vocadlo, D.J. Identification of Asp174 and Asp175 as the key catalytic residues of human O-GlcNAcase by functional analysis of site-directed mutants. *Biochemistry* **2006**, *45*, 3835–3844. [CrossRef]

21. Narine, A.A.; Watson, J.N.; Bennet, A.J. Mechanistic requirements for the efficient enzyme-catalyzed hydrolysis of thiosialosides. *Biochemistry* **2006**, *45*, 9319–9326. [CrossRef]

22. Shen, H.; Byers, L.D. Thioglycoside hydrolysis catalyzed by beta-glucosidase. *Biochem. Biophys. Res. Commun.* **2007**, *362*, 717–720. [CrossRef] [PubMed]

23. Sansenya, S.; Opassiri, R.; Kuaprasert, B.; Chen, C.J.; Cairns, J.R.K. The crystal structure of rice (*Oryza sativa L.*) Os4BGlu12, an oligosaccharide and tuberonic acid glucoside-hydrolyzing beta-glucosidase with significant thioglucohydrolase activity. *Arch. Biochem. Biophys.* **2011**, *510*, 62–72. [CrossRef] [PubMed]

24. Wathelet, J.-P.; Iori, R.; Leoni, O.; Rollin, P.; Mabon, N.; Marlier, M.; Palmieri, S. A recombinant β-O-glucosidase from *Caldocellum saccharolyticum* to hydrolyse desulfo-glucosinolates. *Biotechnol. Lett.* **2001**, *23*, 443–446. [CrossRef]

25. Reese, E.T.; Clapp, R.C.; Mandels, M. A thioglucosidase in fungi. *Arch. Biochem. Biophys.* **1958**, *75*, 228–242. [CrossRef]

26. Yip, V.L.Y.; Withers, S.G. Family 4 glycosidases carry out efficient hydrolysis of thioglycosides by an alpha,beta-elimination mechanism. *Angew. Chem. Int. Ed.* **2006**, *45*, 6179–6182. [CrossRef] [PubMed]

27. Niemec-Cyganek, A.; Szeja, W. Heteroaryl thioglycosides, a new class of substrated for glycosidases. *Pol. J. Chem.* **2003**, *77*, 969–973.

28. Jensen, J.L.; Jencks, W.P. Hydrolysis of benzaldehyde O,S-acetals. *J. Am. Chem. Soc.* **1979**, *101*, 1476–1488. [CrossRef]

29. Hanessian, S.; Bacquet, C.; Lehong, N. Chemistry of the glycosidic linkage. Exceptionally fast and efficient formation of glycosides by remote activation. *Carbohydr. Res.* **1980**, *80*, C17–C22. [CrossRef]

30. Ferrières, V.; Blanchard, S.; Fischer, D.; Plusquellec, D. A novel synthesis of D-galactofuranosyl, D-glucofuranosyl and D-mannofuranosyl 1-phosphates based on remote activation of new and free hexofuranosyl donors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3515–3518. [CrossRef]

31. Demchenko, A.V.; Malysheva, N.N.; De Meo, C. S-Benzoxazolyl (SBox) Glycosides as Novel, Versatile Glycosyl Donors for Stereoselective 1,2-Cis Glycosylation. *Org. Lett.* **2003**, *5*, 455–458. [CrossRef]

32. Demchenko, A.V.; Pornsuriyasak, P.; De Meo, C.; Malysheva, N.N. Potent, Versatile and Stable: Thiazolyl Thioglycosides as Glycosyl Donors. *Angew. Chem. Int. Ed.* **2004**, *43*, 3069–3072. [CrossRef] [PubMed]

33. Euzen, R.; Guégan, J.-P.; Ferrières, V.; Plusquellec, D. First *O*-Glycosylation from Unprotected 1-Thioimidoyl Hexofuranosides Assisted by Divalent Cations. *J. Org. Chem.* **2007**, *72*, 5743–5747. [CrossRef] [PubMed]

34. Kamat, M.N.; De Meo, C.; Demchenko, A.V. S-Benzoxazolyl as a Stable Protecting Moiety and a Potent Anomeric Leaving Group in Oligosaccharide Synthesis. *J. Org. Chem.* **2007**, *72*, 6947–6955. [CrossRef] [PubMed]

35. Hasty, S.J.; Demchenko, A.V. Glycosyl Thioimidates as Versatile Building Blocks for Organic Synthesis. *Chem. Heterocycl. Compd.* **2012**, *48*, 220–240. [CrossRef] [PubMed]

36. Guillotin, L.; Cancellieri, P.; Lafite, P.; Landemarre, L.; Daniellou, R. Chemo-enzymatic synthesis of 3-O-(β-d-glycopyranosyl)-sn-glycerols and their evaluation as preservative in cosmetics. *Pure Appl. Chem.* **2017**, *89*, 11302. [CrossRef]

37. Fukusumi, S.; Kamizono, A.; Horinouchi, S.; Beppu, T. Cloning and nucleotide sequence of a heat-stable amylase gene from an anaerobic thermophile, *Dictyoglomus thermophilum*. *Eur. J. Biochem.* **1988**, *174*, 15–21. [CrossRef] [PubMed]

38. Shi, R.; Li, Z.; Ye, Q.; Xu, J.; Liu, Y. Heterologous expression and characterization of a novel thermo-halotolerant endoglucanase Cel5H from *Dictyoglomus thermophilum*. *Bioresour. Technol.* **2013**, *142*, 338–344. [CrossRef] [PubMed]

39. Nakajima, M.; Imamura, H.; Shoun, H.; Horinouchi, S.; Wakagi, T. Transglycosylation Activity of *Dictyoglomus thermophilum* Amylase A. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 2369–2373. [CrossRef]

40. Guillotin, L.; Richet, N.; Lafite, P.; Daniellou, R. Is the acid/base catalytic residue mutation in β-d-mannosidase Dt Man from *Dictyoglomus thermophilum* sufficient enough to provide thioglycoligase activity? *Biochimie* **2017**, *137*, 190–196. [CrossRef] [PubMed]

41. Guillotin, L.; Kim, H.; Traore, Y.; Moreau, P.; Lafite, P.; Coquoin, V.; Nuccio, S.; De Vaumas, R.; Daniellou, R. Biochemical Characterization of the α-L-Rhamnosidase Dt Rha from *Dictyoglomus thermophilum*: Application to the Selective Derhamnosylation of Natural Flavonoids. *ACS Omega* **2019**, *4*, 1916–1922. [CrossRef]

42. Huggett, A.S.G.; Nixon, D.A. Use of glucose oxidase, peroxidase and o-dianisidine in determination of blood and urinary glucose. *Lancet* **1957**, *270*, 368–370. [CrossRef]

43. Guillotin, L.; Lafite, P.; Daniellou, R. Unraveling the Substrate Recognition Mechanism and Specificity of the Unusual Glycosyl Hydrolase Family 29 BT2192 from *Bacteroides thetaiotaomicron*. *Biochemistry* **2014**, *53*, 1447–1455. [CrossRef] [PubMed]

44. Akiba, T.; Nishio, M.; Matsui, I.; Harata, K. X-ray structure of a membrane-bound β-glycosidase from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Proteins Struct. Funct. Bioinform.* **2004**, *57*, 422–431. [CrossRef] [PubMed]

45. Zechel, D.L.; Boraston, A.B.; Gloster, T.; Boraston, C.M.; Macdonald, J.M.; Tilbrook, D.M.G.; Stick, R.V.; Davies, G.J. Iminosugar Glycosidase Inhibitors: Structural and Thermodynamic Dissection of the Binding of Isofagomine and 1-Deoxynojirimycin to β-Glucosidases. *J. Am. Chem. Soc.* **2003**, *125*, 14313–14323. [CrossRef] [PubMed]

46. Zechel, D.L.; Withers, S.G. Dissection of nucleophilic and acid-base catalysis in glycosidases. *Curr. Opin. Chem. Biol.* **2001**, *5*, 643–649. [CrossRef]

47. Zhu, X.; Schmidt, R.R. New Principles for Glycoside-Bond Formation. *Angew. Chem. Int. Ed.* **2009**, *48*, 1900–1934. [CrossRef] [PubMed]

48. Codée, J.D.C.; Litjens, R.E.J.N.; Van den Bos, L.J.; Overkleeft, H.S.; Van der Marel, G.A. Thioglycosides in sequential glycosylation strategies. *Chem. Soc. Rev.* **2005**, *34*, 769–782. [CrossRef] [PubMed]

49. Pornsuriyasak, P.; Demchenko, A.V. S-Thiazolinyl (STaz) Glycosides as Versatile Building Blocks for Convergent Selective, Chemoselective and Orthogonal Oligosaccharide Synthesis. *Chem.-Eur J.* **2006**, *12*, 6630–6646. [CrossRef] [PubMed]

50. Jerez, G.; Kaufman, G.; Prystai, M.; Schenkeveld, S.; Donkor, K.K. Determination of thermodynamic pK_a values of benzimidazole and benzimidazole derivatives by capillary electrophoresis. *J. Sep. Sci.* **2009**, *32*, 1087–1095. [CrossRef]

51. Akhond, M.; Ghaedi, M.; Tashkhourian, J. Development of a New Copper(II) Ion-selective Poly(vinyl chloride) Membrane Electrode Based on 2-Mercaptobenzoxazole. *Bull. Korean Chem. Soc.* **2005**, *26*, 882–886.

52. Kamat, M.N.; Rath, N.P.; Demchenko, A. V Versatile Synthesis and Mechanism of Activation of S-Benzoxazolyl Glycosides. *J. Org. Chem.* **2007**, *72*, 6938–6946. [CrossRef] [PubMed]

53. Case, D.A.; Darden, T.A.; III, T.E.C.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Walker, R.C.; Zhang, W.; Merz, K.M.; et al. *AMBER 12 2012*; University of California: San Francisco, CA, USA, 2012.

54. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]

55. Wang, L.-P.; McKiernan, K.A.; Gomes, J.; Beauchamp, K.A.; Head-Gordon, T.; Rice, J.E.; Swope, W.C.; Martínez, T.J.; Pande, V.S. Building a More Predictive Protein Force Field: A Systematic and Reproducible Route to AMBER-FB15. *J. Phys. Chem. B* **2017**, *121*, 4023–4039. [CrossRef] [PubMed]

56. Jakalian, A.; Jack, D.B.; Bayly, C.I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641. [CrossRef] [PubMed]

57. Lafite, P.; André, F.; Zeldin, D.C.; Dansette, P.M.; Mansuy, D. Unusual Regioselectivity and Active Site Topology of Human Cytochrome P450 2J2. *Biochemistry* **2007**, *46*, 10237–10247. [CrossRef] [PubMed]

Article

Deciphering the Role of V88L Substitution in NDM-24 Metallo-β-Lactamase

Zhihai Liu [1,2], Alessandra Piccirilli [3], Dejun Liu [2], Wan Li [2], Yang Wang [2] and Jianzhong Shen [2,*]

[1] Agricultural Bio-pharmaceutical Laboratory, College of Chemistry and Pharmaceutical Sciences, Qingdao Agricultural University, Qingdao 266109, China

[2] Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Veterinary Medicine, China Agricultural University, Beijing 100193, China

[3] Dipartimento di Scienze Cliniche Applicate e Biotecnologiche, Università degli Studi dell'Aquila, 67100 L'Aquila, Italy

[*] Correspondence: sjz@cau.edu.cn; Tel.: +86-10-62732803; Fax: +86-10-62731032

Abstract: The New Delhi metallo-β-lactamase-1 (NDM-1) is a typical carbapenemase and plays a crucial role in antibiotic-resistance bacterial infection. Phylogenetic analysis, performed on known NDM-variants, classified NDM enzymes in seven clusters. Three of them include a major number of NDM-variants. In this study, we evaluated the role of the V88L substitution in NDM-24 by kinetical and structural analysis. Functional results showed that V88L did not significantly increase the resistance level in the NDM-24 transformant toward penicillins, cephalosporins, meropenem, and imipenem. Concerning ertapenem, *E. coli* DH5α/NDM-24 showed a MIC value 4-fold higher than that of *E. coli* DH5α/NDM-1. The determination of the k_{cat}, K_m, and k_{cat}/K_m values for NDM-24, compared with NDM-1 and NDM-5, demonstrated an increase of the substrate hydrolysis compared to all the β-lactams tested, except penicillins. The thermostability testing revealed that V88L generated a destabilized effect on NDM-24. The V88L substitution occurred in the β-strand and low β-sheet content in the secondary structure, as evidenced by the CD analysis data. In conclusion, the V88L substitution increases the enzyme activity and decreases the protein stability. This study characterizes the role of the V88L substitution in NDM-24 and provides insight about the NDM variants evolution.

Keywords: New Delhi metallo-β-lactamase; NDM-24; kinetic profile; secondary structure

1. Introduction

Metallo-β-lactamases (MBLs) are a group of enzymes that confer high resistance to most β-lactams. The active site of these enzymes contains one or two zinc ions, that are crucial for catalytic mechanism [1]. Based on their amino acid sequences, MBLs have been divided into subclasses B1, B2, and B3 [2]. Among subclass B1, the New Delhi metallo-β-lactamase (NDM-1) is one of the most widespread carbapenemase. NDM-1 was first identified in 2008 in a clinical strain of *Klebsiella pneumoniae* [3]. NDM-1 producing bacteria can hydrolyse all β-lactams (except monobactams), including carbapenems, the "last resort" antibiotics used in clinical therapy. NDM-1 genes are located on plasmids that mediate their dissemination across different bacterial strains [4,5]. However, the clinical success of NDM is also due to the fact that it is a lipoprotein anchored to the outer membrane, resulting in an unusual stability of NDM-1 and enabling secretion, in Gram-negative bacteria [6–8].

To date, more than 26 variants differing by a limited number of substitutions have been identified [9]. Previous studies revealed that these substitutions have contributed to NDM-1 to increase the hydrolytic activity toward several β-lactams resulting in an increment of resistance in the host bacteria [10]. Crystal structures showed that NDM-1 presents the typical αβ/βα fold of MBLs [11,12]. In this enzyme, the zinc ions are coordinated by six conserved residues: H120, H122, and H189 for Zn1

(BBL numbering) and D124, C208, and H250 for Zn2 (BBL numbering). The active site is surrounded by Loop 3 (residues 67-73) and Loop 10 (residues 210-230), involved in the substrate accommodation [12]. The most frequent substitution in NDM-1 is M154L, found in 11 NDM variants (NDM-4, -5, -7, -8, -12, -13, -15, -17, -19, -20 and -21) [9,13–16]. Indeed, V88L has been reported in five NDM variants (NDM-5, -17, -20, -21 and -24). Other frequent substitutions are A233V (NDM-6, -15, -19 and -27), D130G (NDM-8 and -14), D130N (NDM-7 and -19), and D95N (NDM-3 and -27) [10]. The single substitutions of M154L and D130G seem to increase the carbapenemase activity in NDM-4 and NDM-14, respectively [17,18]. Moreover, the combination D130G/M154L (NDM-8), reduces the hydrolysis toward carbapenems [19]. The main goal of the study was to evaluate the role of the V88L substitution in the NDM-24 enzyme. The NDM-24 was generated in the laboratory by a site-directed mutagenesis of NDM-1 and the enzyme properties, protein structure, and thermal stability were studied compared with NDM-1 and NDM-5.

2. Results and Discussion

2.1. Phylogenetic Analysis

A phylogenetic analysis of NDM-1 variants was performed in order to classify these enzymes based on their amino acid similarities. Overall, the NDM variants were classified into three major clusters (NDM-1, NDM-4, and NDM-24), two minor clusters (NDM-3 and NDM-6), and two divergent proteins (NDM-14 and NDM-10). As shown in Figure 1, the NDM variants are well categorized. The NDM-1 cluster includes eight variants that showed only one amino acid replacement, except for NDM-18 where an insertion of five amino acids have been found (position 48-52). In the NDM-4 group, all variants possess the replacement at position 154. In particular, except for NDM-11 containing the M154V substitution, all variants shared M154L. In the NDM-24 group, Valine at position 88 has been replaced by a Leucine (V88L). Concerning the two minor groups, similar characteristics were observed with the D95N and A233V substitution for the NDM-3 and NDM-6 clusters, respectively.
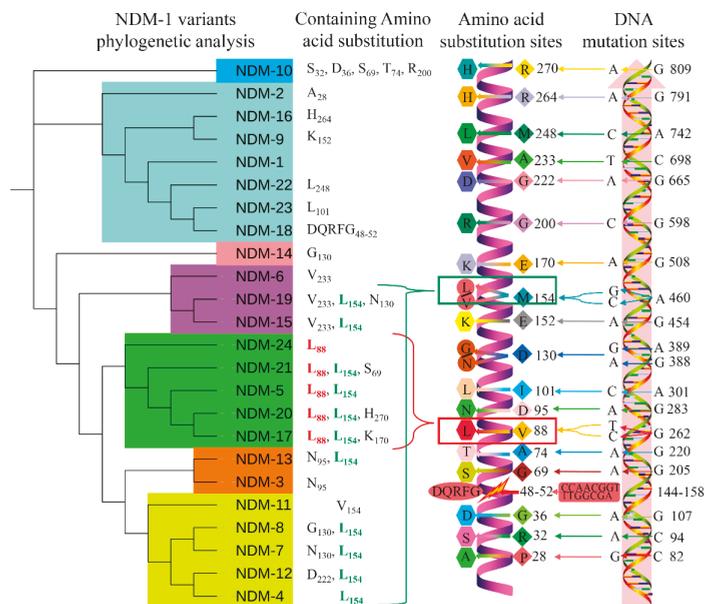

Figure 1. New Delhi metallo-β-lactamase-1 (NDM-1) variants phylogenetic analysis. Phylogenetic groups were differently coloured: For example, the NDM-24 cluster is coloured in green.

2.2. Functional Study

The NDM-24 variant was obtained by a site-directed mutagenesis by using the NDM-1 as template. All genes (*bla*_{NDM-1}, *bla*_{NDM-5}, and *bla*_{NDM-24}) were cloned into pHSG398, which were controlled by the same promoter and thus the same expression. The *E. coli* DH5α recombinant strains were used to test susceptibility to a wide panel of β-lactams. As shown in Table 1, the results of the susceptibility test revealed that NDM-1, NDM-5, and NDM-24 confer resistance to most β-lactams with similar MIC values, suggesting that the NDM enzymes were successfully expressed in the host cells. A different behaviour was observed for carbapenems, for which the MIC values were markedly lower than those of penicillins and cephalosporins with the exception of cefepime, as previously reported [15–20]. Concerning ertapenem, NDM-24 and NDM-5 showed an increase of the MIC values of the 4- and 8-fold with respect to NDM-1. Based on the data obtained, the V88L substitution enhances resistance toward ertapenem.

Table 1. Antimicrobial susceptibility of *E. coli* BL21(DE3)-DH5α carrying *bla*_{NDM-1}, *bla*_{NDM-5}, and *bla*_{NDM-24}.

Antibiotic	MIC (μg/mL)			
	E. coli DH5α/pHSG398	*E. coli* DH5α/pHSG398 -NDM-24	*E. coli* DH5α/pHSG398 -NDM-1	*E. coli* DH5α/pHSG398 -NDM-5
Ampicillin	2	>256	>256	>256
Penicillin G	16	>256	>256	>256
Aztreonam	0.031	0.031	0.031	0.031
Cefepime	0.031	2	1	2
Cefotaxime	0.062	32	64	64
Cefoxitin	2	128	128	128
Ceftazidime	0.125	256	256	256
Cefazolin	2	128	128	256
Ertapenem	0.015	1	0.25	2
Imipenem	0.062	2	2	2
Meropenem	0.031	1	1	2

2.3. Characteristics of Enzyme Activity

In order to obtain soluble and active enzymes, the recombinant plasmids were successfully expressed in the *E. coli* BL21 (DE3) cells as described in the methods section. After purification, the enzymes were checked on SDS-PAGE to confirm the solubility and purity (>90%) (Figure S1). The MALDI-TOF mass spectrometry was used to confirm the molecular mass of the three enzymes, which corresponds to 24884,024 Da (Figure S2). To investigate the enzyme activity, the kinetic parameters for NDM-1, NDM-5, and NDM-24 were determined (Table 2).

All the NDM variants of this study were able to hydrolyse all the β-lactams tested. Compared with NDM-1, NDM-24 showed lower K_m values for penicillins and ceftazidime whereas for carbapenems they are quite similar. Comparing the k_{cat} values, NDM-24 hydrolyses all β-lactams, except penicillins, better than NDM-1 and NDM-5. In particular, the k_{cat} values of NDM-24 are 2.26-, 1.61-, 2.73-, 2.02-, 2.17-, and 1.75-fold higher than NDM-1 towards penicillin G, ceftazidime, cefepime, imipenem, meropenem, and ertapenem, respectively. This was also confirmed by a slight increase of catalytic efficiency. This result stated the important role of V88L in the substrate hydrolysis. The contribution of V88L is likely that of M154L as demonstrated by the calculation of the k_{cat}/K_m rates (Table 2). This was possibly due to differences in the intrinsic properties, such as the enzyme stability, protein expression, and adaptability [21–24], and nutritional conditions of bacteria in vivo/vitro. Comparing the k_{cat}/K_m values of carbapenems, the carbapenemase activity of NDM-5 was similar to NDM-24, but higher than NDM-1. A recent study showed that an increase of the catalytic efficiency (k_{cat}/K_m) for meropenem has been ascertained in NDM-5 (V88L and M154L). In NDM-4, which contains only M154L, no significant change has been observed, suggesting that V88L might play a role in enhancing the NDM enzymes activity rather than M154L. Moreover, an increase of the carbapenemase activity was also observed in the evolutionary NDM variants, such as NDM-17 (V88L, M154L, and E170K) and NDM-20 (V88L, M154L, and R270H) [10,15,16].

Table 2. Kinetic parameters of NDM-1, NDM-5, and NDM-24 toward β-lactams [a].

Kinetic Parameters	Enzyme	β-Lactams [b]						
		AMP	PEN	TAG	FEP	MEM	IPM	ETP
K_m (μM)	NDM-24	638.79 ± 23.86	331.30 ± 29.43	173.85 ± 9.73	318.93 ± 10.86	266.24 ± 27.03	338.20 ± 24.23	125.23 ± 19.08
	NDM-1	1249.98 ± 210.94	224.57 ± 13.57	213.90 ± 11.01	173.55 ± 19.46	284.24 ± 7.87	234.83 ± 7.44	105.54 ± 3.09
	NDM-5	825.00 ± 0.29	315.21 ± 46.68	76.45 ± 4.76	179.64 ± 12.19	275.16 ± 36.87	292.97 ± 13.76	82.18 ± 3.86
k_{cat} (s^{-1})	NDM-24	259.94 ± 23.52	179.10 ± 8.17	43.13 ± 1.06	22.98 ± 0.34	151.75 ± 6.69	173.16 ± 8.83	110.31 ± 7.62
	NDM-1	254.34 ± 28.96	79.28 ± 1.96	26.73 ± 0.71	8.42 ± 0.63	75.18 ± 3.44	79.81 ± 5.15	62.89 ± 1.15
	NDM-5	346.13 ± 31.30	214.13 ± 12.11	26.96 ± 0.75	13.05 ± 0.24	142.48 ± 17.91	149.63 ± 2.02	83.18 ± 1.67
k_{cat}/K_m (μM^{-1} s^{-1})	NDM-24	0.41	0.54	0.25	0.072	0.57	0.51	0.88
	NDM-1	0.20	0.35	0.13	0.046	0.26	0.34	0.60
	NDM-5	0.40	0.68	0.35	0.073	0.52	0.51	1.01
k_{cat}/K_m (μM^{-1} s^{-1}) ratio for:	NDM-24/NDM-1	2.00	1.53	1.98	1.49	2.16	1.51	1.46
	NDM-5/NDM-24	1.03	1.26	1.42	1.01	0.91	1.00	1.15
	NDM-5/NDM-1	2.07	1.92	2.82	1.50	1.96	1.50	1.68

[a] The proteins were initially purified with a His-tag, which was removed after purification. Each kinetic value is the mean of three different measurements; the error was below 5%. [b] β-lactams: AMP, ampicillin; TAG, ceftazidime; PEN, penicillin G; FEP, cefepime; IPM, imipenem; MEM, meropenem; ETP, ertapenem.

2.4. Thermal Stability

As previously reported, mutations in the NDM variants can affect the enzymes stability, resulting in changing the persistence lifetime in the bacterial host, and consequent antibiotics resistance [25]. For determining whether the V88L substitution influences the NDM-24 stability property, circular dichroism CD was used to assay the protein stability by recording signal changes. NDM-5 was used as reference to analyze the effect of M154L. Compared with NDM-1 and NDM-5, NDM-24 (59.41 ± 0.06 °C) possessed the lowest melting temperature (Figure 2). Notably, the V88L destabilized effect was compensated by M154L in NDM-5 with a remarkable higher thermal temperature than NDM-24 (69.13 ± 3.6 °C compared to 59.41 ± 0.06 °C). Moreover, NDM-5 showed a higher stability than NDM-1 suggesting the destabilized role of M154L. This was in agreement with a previous document that the M154L mutation would be a turning point for the NDM variants, in which combing M154L with additional substitutions benefit for the NDM enzymes exhibiting increased thermostability [10]. In the NDM-24 group there are four variants (NDM-5, -17, -20, and -21) in which the combination of the V88L and M154L substitutions takes favorable results in terms of the stability and environmental selection.
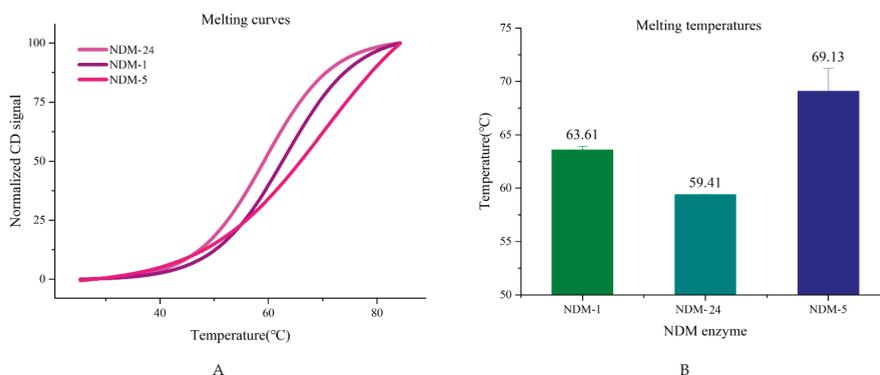
Figure 2. (**A**) Thermal stability melting curves. (**B**) Melting temperatures of the NDM enzymes as determined by the circular dichroism analysis: NDM-1, 63.61 ± 0.57 °C; NDM-V88L, 59.41 ± 0.06 °C; and NDM-5, 69.13 ± 3.6 °C. Data are the means of triplicate experiments, with error bars showing the standard deviation (±SD).

2.5. Structure Analysis

Previous reports indicated that mutations in the NDM influence the α-helical, β-sheets content, and loop flexibility [26]. For example, the Q123A substitution in NDM-1 leads to a decrease of the α-helical content [27]. To know if the V88L substitution could modify the NDM-24 structure, a secondary structure was determined by the Far-UV CD spectrum. All NDM variants CD spectrum data were fitted and shown in Figure 3. The spectrum signals were superimposable at most wavelengths, and showed characteristics of αβ/βα fold, a typical and conservative protein structure in MBLs [28]. The presence of positive bands at 192 nm and two negative peaks at 208 nm, a minimum peak, and 220 nm, suggesting the dominance of the β-sheets and α-helical content. The major differences were observed in the nearby 192 nm, symbolizing α-helical peak, and 208–220 nm, a α-helical and β-sheets bonds. Overall, the α-helical content was found ranging between 13%–20% in NDM-1, NDM-5, and NDM-24 (Table 3), in agreement with previous reports and the content of the β-sheet was high around 30% [27]. Compared to NDM-1, NDM-24 possesses a higher α-helical content and lower β-sheet content, suggesting that V88L was responsible for the secondary structure content changes of NDM-24. Furthermore, the secondary predicted result (Figure S3) confirmed that the V88L substitution occurred in the β-strand terminal, which may be prematurely terminated, leading to a decrease in the β-sheet content. Kumar et al. claimed that 152A, located in the β-strand, drastically influenced the NDM-5

activity and protein thermolability, by reducing the β-sheet content [26]. Our analysis demonstrates that the emergence of M154L (in NDM-5) caused the α-helical to decrease and the β-sheet to increase relative to NDM-24, while the α-helical and β-sheet content of NDM-5 were between NDM-1 and NDM-24. In addition, the 3D model of NDM-24 (Figure 4) was generated by using NDM-1 (PDB accession: 5N0H, 4RBS, 4HKY, and 4EYL) and NDM-5 (PDB accession: 6MGY, and 4TZE) as a template. Although the residue 88L is away from the active site groove and far to the active loops (Loop 3 and Loop 10), differences in the structure content, stability, and enzyme activity were ascertained. Several studies confirmed that non-activity sites substitution can influence the NDM catalytic efficiency [29], and our results about the V88L substitution support this theory.
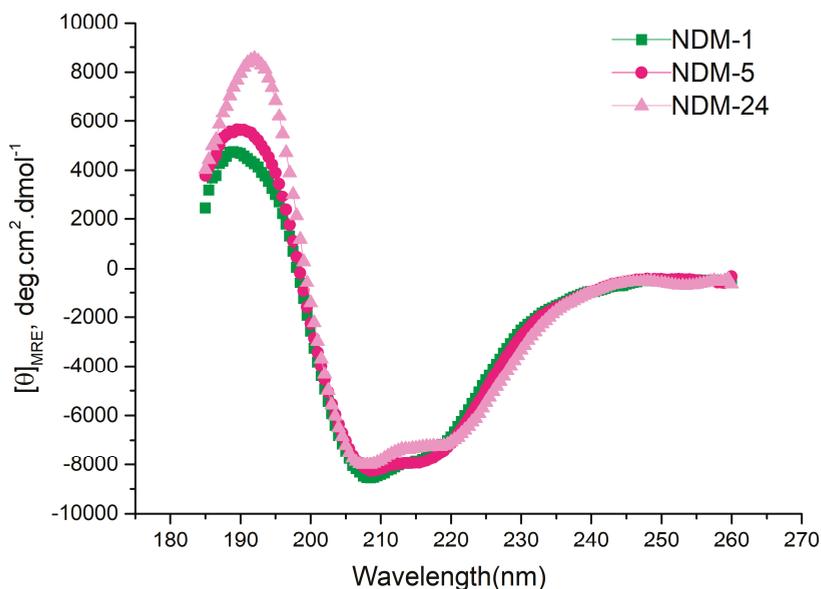
Figure 3. Normalized circular dichroism (CD) spectra of the NDM enzymes tested. MRE: Mean residue ellipticity.

Table 3. Proportions of various secondary structural elements in the NDM-1, NDM-5, and NDM-24 enzymes.

Program Algorithms [a]	Structural Elements [b]	SMP50(9) [c]			SP37(3) [c]			SP29(1) [c]		
		NDM-1	NDM-5	NDM-24	NDM-1	NDM-5	NDM-24	NDM-1	NDM-5	NDM-24
SELCON3	H(r)	0.070	0.078	0.092	0.062	0.074	0.092	0.059	0.079	0.087
	H(d)	0.085	0.088	0.089	0.081	0.088	0.089	0.078	0.087	0.086
	S(r)	0.215	0.199	0.195	0.228	0.214	0.195	0.231	0.191	0.196
	S(d)	0.115	0.109	0.108	0.117	0.113	0.108	0.118	0.107	0.108
	Trn	0.214	0.211	0.194	0.218	0.214	0.194	0.226	0.214	0.215
	Unrd	0.284	0.287	0.261	0.282	0.279	0.261	0.287	0.292	0.285
	H(r)+H(d)	0.155	0.166	0.181	0.143	0.162	0.181	0.137	0.166	0.173
	S(r)+S(d)	0.33	0.308	0.303	0.345	0.327	0.303	0.349	0.298	0.304
CONTINLL	H(r)	0.054	0.075	0.091	0.046	0.079	0.097	0.071	0.078	0.093
	H(d)	0.079	0.092	0.101	0.089	0.095	0.103	0.092	0.096	0.100
	S(r)	0.217	0.208	0.187	0.202	0.205	0.182	0.197	0.197	0.183
	S(d)	0.114	0.113	0.108	0.112	0.111	0.107	0.113	0.111	0.107
	Trn	0.233	0.220	0.220	0.248	0.216	0.216	0.231	0.222	0.225
	Unrd	0.303	0.292	0.293	0.304	0.293	0.294	0.297	0.297	0.292
	H(r)+H(d)	0.133	0.167	0.192	0.135	0.174	0.200	0.163	0.174	0.193
	S(r)+S(d)	0.331	0.321	0.295	0.314	0.316	0.289	0.310	0.308	0.290

[a] The CDPro program package was used to analyse the data using two algorithms: CONTINLL and SELCON3.
[b] H(r), regular α-helix; H(d), distorted α-helix; S(r), regular β-strand; S(d), distorted β-strand; Trn, turns; Unrd, unordered. [c] The reference protein sets (IBasis sets) were adopted.

Note: 5N0H, 4RBS, 4HKY, 4EYL, 6MGY and 4TZE are shown in green, yellow, wheat, blue, cyan and orange, respectively. The hydrolyzed meropenem (ligand) is shown in the binding pocket of 4EYL with blue surface. The VAL in 5N0H, 4RBS, 4HKY and 4EYL marked in red, while the LEU in 6MGY and 4TZE marked in dark.
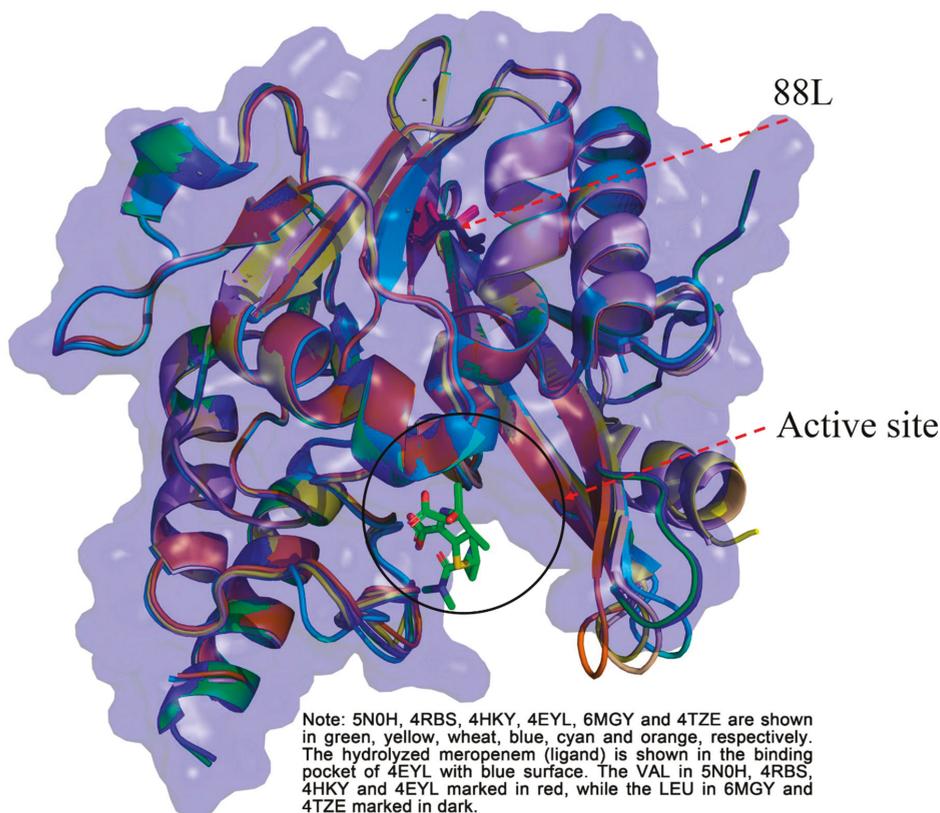
Figure 4. Cartoon model of NDM-24. To acquire a credible model, the 6 NDM crystal structure (NDM-1(5N0H, 4RBS, 4HKY, and 4EYL) and NDM-5(6MGY, and 4TZE)) were adopted. The residue 88L and active pocket were labelled.

3. Material and Methods

3.1. Site-Directed Mutagenesis, Cloning and Expression of NDM Variants

The bla_{NDM-1} and bla_{NDM-5} encoding genes were obtained from clinical *Escherichia coli* strains as previously described [15,16]. Site-directed mutagenesis was performed to generate bla_{NDM-24} using the pHSG398/NDM-1 plasmid as template and primers listed in Table S1, as previously described [30].

First, the bla_{NDM}-genes were cloned into a pHSG398 vector (TaKaRa Bio, Dalian, China) using the *BamHI* and *XhoI* restriction sites. In a second PCR experiment, the bla_{NDM} variants were amplified without a signal peptide introducing the Tobacco Etch Virus (TEV) at the N-terminal sequence. In order to obtain enzymes overexpression, the amplicons were inserted into a pET-28a(+) plasmid. The *E. coli* DH5α competent cells were used as a non-expression host. *E. coli* BL21(DE3) was used for enzymes overexpression. The authenticity of recombinant plasmids was verified by PCR and sequencing was with Sanger.

3.2. Antimicrobial Susceptibility Tests

The phenotypic profile has been characterized by a microdilution method using a bacterial inoculum of 5×10^5 CFU/ml according to the Clinical Laboratory and Standards Institute [31,32]. *E. coli* ATCC25922 was used as a negative control.

3.3. Production and Purification of NDM-1, NDM-5, and NDM-24

NDM-1, NDM-5, and NDM-24 were extracted from 0.5 L of culture of *E. coli* BL21(DE3)/pETNDM-1, *E. coli* BL21 (DE3) /pETNDM-5, and *E. coli* BL21 (DE3)/pETNDM-24, respectively. The cultures were grown at 37 °C to achieve an A_{600} of 0.5 L, and 0.4 mM of IPTG was added. After addition of the IPTG, the cultures were incubated at 20 °C for 16 h, under aerobic conditions. Thereafter, a cell supernatant containing the soluble NDM protein was obtained from the lytic bacteria by centrifuging at 10^4 rpm. Proteins purification followed the manufacturer's instructions (Qiagen, Hilden, Germany) by using the Ni-nitrilotriacetic acid (NTA) agarose. The turbo tobacco etch virus (TEV) protease (Accelagen, San Diego, CA, USA) was used to gain the untagged protein without the His tags. The SDS-PAGE was performed to estimate the NDM purity enzymes. The protein concentration was determined using a BCA Protein Quantification Kit (Vazyme Biotech Co., Ltd, Nanjing, China). The β-lactamase activity was monitored at each purification step using the colour change of nitrocefin 1 mg/mL, a chromogenic cephalosporin, according to the previous report [20].

3.4. Determination of Kinetic Parameters

Steady-state kinetic experiments were performed following the hydrolysis of each substrate at 25 °C in a 50 mM phosphate buffer, pH 7.0 in the presence of 20 μM Zn SO_4. The data were collected with a SpectraMax M5 multi-detection microplate reader (Molecular Devices, Sunnyvale, CA, USA) as previously described [33]. Kinetic parameters were determined under initial-rate conditions using the GraphPad Prism® version 5.01 (San Diego, CA, USA) to generate the Michaelis-Menten curves, or by analysing the complete hydrolysis time courses [34,35]. Each kinetic value is the mean of the results of three different measurements. The error was below 5%. NDM-5 was used as a reference to normalize.

3.5. Circular Dichroism and Structure Analysis

The circular dichroism (CD) spectra (185 to 260 nm) were determined with a Chirascan Plus CD spectrophotometer (Applied Photophysics Ltd, Leatherhead, UK) equipped with a Peltier temperature-controlled cell holder, at 25 °C using a 1-mm pathlength cuvette and the Savitzky-Golay filter was explored to the baseline-correct spectra data. Protein concentrations were diluted to 0.05-0.2 mg/mL with a 5 mM sodium phosphate buffer pH 7.0 [36]. 207 nm spectrum data was used as the baseline to normalize and calibrate data for eliminating minor errors due to different concentrations [37]. The analysis was performed using the CONTINLL and SELCON3 algorithms with reference data sets three and nine, respectively [38]. The super-secondary (tertiary) structures of the proteins were analysed by the CDPro software package, which is available at the CDPro website: https://sites.bmb.colostate.edu/sreeram/CDPro/ [38,39]. To assay the location of the V88L substitution and analyse its effect on the structure, the pharmacophore modeling and screening software program Discovery Studio (version 2018) was employed to generate a three-dimensional (3D) interconnected model of NDM-24 using NDM-1 and NDM-5 as a template, in which reliable data of the crystal structure were collected from the PDB database.

3.6. Thermal Stability Testing

The melting temperature (T_m) was used to show the protein stability. The determination of T_m was performed by recording the CD signal change at 222 nm. Data were collected at a ramp of 1 °C /min with a temperature range from 20-90 °C. The two-state model using nonlinear regression (Boltzmann) in the OriginPro 9.1.64 (OriginLab, Northampton, MA, USA) was applied to analyse the data. When 50% of the protein melts, the temperature is defined as the *Tm*, representing thermal stability.

4. Conclusions

Our study explored the NDM-24 biological function and probed the V88L substitution role on the structure, enzyme activity, and stability. In brief, this non-active site change enhances the

enzyme activity by increasing the turnover rate, related with an indirect effect on conformation. However, the loss cost caused by V88L significantly decreased the protein stability, and would shorten the persistence lifetime in the cell, so that the resistance to antibiotics hardly exhibits an outstanding elevation for the NDM-24-producing transformants. Meanwhile, alterations in the secondary content, such as lowering the β-sheet, have an interesting role in the NDM instability, being relevant to the V88L substitution occurring in the β-strand. According to previous data, the V88L/M154L combination appears to be favorable in the NDM evolution under an environmental pressure selection [14]. Further analysis about the significance of non-active-site residues will help in the comprehension of the resistance mechanism and broaden insight in the development of inhibitors, such as potential antibiotics candidate by reducing the protein stability lifetime.

References

1. Palzkill, T. Metallo-beta-lactamase structure and function. *Ann. N. Y. Acad. Sci.* **2013**, *1277*, 91–104. [CrossRef] [PubMed]

2. Garau, G.; García-Sáez, I.; Bebrone, C.; Anne, C.; Mercuri, P.; Galleni, M.; Frère, J.M.; Dideberg, O. Update of the standard numbering scheme for class B beta-lactamases. *Antimicrob. Agents Chemother.* **2004**, *48*, 2347–2349. [CrossRef] [PubMed]

3. Yong, D.; Toleman, M.A.; Giske, C.G.; Cho, H.S.; Sundman, K.; Lee, K.; Walsh, T.R. Characterization of a New Metallo-β-Lactamase Gene, blaNDM-1, and a Novel Erythromycin Esterase Gene Carried on a Unique Genetic Structure in Klebsiella pneumoniae Sequence Type 14 from India. *Antimicrob. Agents Chemother.* **2009**, *53*, 5046–5054. [CrossRef] [PubMed]

4. Bonnin, R.A.; Poirel, L.; Carattoli, A.; Nordmann, P. Characterization of an IncFII Plasmid Encoding NDM-1 from Escherichia coli ST131. *PLoS ONE* **2012**, *7*, 34752. [CrossRef] [PubMed]

5. Dolejska, M.; Villa, L.; Poirel, L.; Nordmann, P.; Carattoli, A. Complete sequencing of an IncHI1 plasmid encoding the carbapenemase NDM-1, the ArmA 16S RNA methylase and a resistance-nodulation-cell division/multidrug efflux pump. *J. Antimicrob. Chemother.* **2013**, *68*, 34–39. [CrossRef] [PubMed]

6. King, D.; Strynadka, N. Crystal structure of New Delhi metallo-betalactamase reveals molecular basis for antibiotic resistance. *Protein Sci.* **2011**, *20*, 1484–1491. [CrossRef] [PubMed]

7. Gonzalez, L.J.; Bahr, G.; Nakashige, T.G.; Nolan, E.M.; Bonomo, R.A.; Vila, A.J. Membrane anchoring stabilizes and favors secretion of New Delhi metallo-beta-lactamase. *Nat. Chem. Biol.* **2016**, *12*, 516–522. [CrossRef] [PubMed]

8. Gonzalez, L.J.; Bahr, G.; Vila, A.J. Lipidated beta-lactamases: From bench to bedside. *Future Microbiol.* **2016**, *11*, 1495–1498. [CrossRef]

9. Khan, A.U.; Maryam, L.; Zarrilli, R. Structure, genetics and worldwide spread of New Delhi metallo-beta-lactamase (NDM): A threat to public health. *BMC Microbiol.* **2017**, *17*, 101. [CrossRef]

10. Cheng, Z.; Thomas, P.W.; Ju, L.; Bergstrom, A.; Mason, K.; Clayton, D.; Miller, C.; Bethel, C.R.; Vanpelt, J.; Tierney, D.L.; et al. Evolution of New Delhi metallo-β-lactamase (NDM) in the clinic: Effects of NDM mutations on stability, zinc affinity, and mono-zinc activity. *J. Biol. Chem.* **2018**, *293*, 12606–12618. [CrossRef]

11. Zhang, H.O.; Hau, Q. Crystal structure of NDM-1 reveals a common β-lactam hydrolysis mechanism. *FASEB J.* **2011**, *25*, 2574–2582. [CrossRef] [PubMed]

12. Kim, Y.; Cunningham, M.A.; Mire, J.; Tesar, C.; Sacchettini, J.; Joachimiak, A. NDM-1, the ultimate promiscuous enzyme: Substrate recognition and catalytic mechanism. *FASEB J.* **2013**, *27*, 1917–1927. [CrossRef] [PubMed]

13. Liu, L.; Feng, Y.; McNally, A.; Zong, Z. bla NDM-21, a new variant of blaNDM in an Escherichia coli clinical isolate carrying blaCTX-M-55 and rmtB. *J. Antimicrob. Chemother.* **2018**, *73*, 2336–2339. [CrossRef] [PubMed]

14. Bahr, G.; Vitor-Horen, L.; Bethel, C.R.; Bonomo, R.A.; González, L.J.; Vila, A.J. Clinical evolution of New Delhi Metallo-β-lactamase (NDM) optimizes resistance under Zn(II) deprivation. *Antimicrob. Agents Chemother.* **2018**, *62*, 1817–1849. [CrossRef] [PubMed]

15. Liu, Z.; Li, J.; Wang, X.; Liu, D.; Ke, Y.; Wang, Y.; Shen, J. Novel Variant of New Delhi Metallo-β-lactamase, NDM-20, in Escherichia coli. *Front. Microbiol.* **2018**, *9*, 248. [CrossRef] [PubMed]

16. Liu, Z.; Wang, Y.; Walsh, T.R.; Liu, D.; Shen, Z.; Zhang, R.; Yin, W.; Yao, H.; Li, J.; Shen, J. Plasmid-mediated novel *bla*NDM-17 gene encoding a carbapenemase with enhanced activity in a ST48 *Escherichia coli* strain. *Antimicrob. Agents Chemother.* **2017**, *61*, 2216–2233. [CrossRef] [PubMed]

17. Nordmann, P.; Boulanger, A.E.; Poirel, L. NDM-4 Metallo-β-Lactamase with Increased Carbapenemase Activity from Escherichia coli. *Antimicrob. Agents Chemother.* **2012**, *56*, 2184–2186. [CrossRef]

18. Zou, D.; Huang, Y.; Zhao, X.; Liu, W.; Dong, D.; Li, H.; Wang, X.; Huang, S.; Wei, X.; Yan, X.; et al. A Novel New Delhi Metallo-β-Lactamase Variant, NDM-14, Isolated in a Chinese Hospital Possesses Increased Enzymatic Activity against Carbapenems. *Antimicrob. Agents Chemother.* **2015**, *59*, 2450–2453. [CrossRef]

19. Tada, T.; Miyoshi-Akiyama, T.; Dahal, R.K.; Sah, M.K.; Ohara, H.; Kirikae, T.; Pokhrel, B.M. NDM-8 Metallo-β-Lactamase in a Multidrug-Resistant Escherichia coli Strain Isolated in Nepal. *Antimicrob. Agents Chemother.* **2013**, *57*, 2394–2396. [CrossRef]

20. Tada, T.; Shrestha, B.; Miyoshi-Akiyama, T.; Shimada, K.; Ohara, H.; Kirikae, T.; Pokhrel, B.M. NDM-12, a Novel New Delhi Metallo-β-Lactamase Variant from a Carbapenem-Resistant *Escherichia coli* Clinical Isolate in Nepal. *Antimicrob. Agents Chemother.* **2014**, *58*, 6302–6305. [CrossRef]

21. Ines, S.; Emma, K.; Rudolf, R.; Rumyana, M.; Anne Marie, Q.; Adolf, B. VIM-15 and VIM-16, two new VIM-2-like metallo-beta-lactamases in *Pseudomonas aeruginosa* isolates from Bulgaria and Germany. *Antimicrob. Agents Chemother.* **2008**, *52*, 2977.

22. Patricia, M.; Tomatis, P.E.; Mussi, M.A.; Fernando, P.; Viale, A.M.; Limansky, A.S.; Vila, A.J. Biochemical characterization of metallo-beta-lactamase VIM-11 from a *Pseudomonas aeruginosa* clinical strain. *Antimicrob. Agents Chemother.* **2008**, *52*, 2250.

23. Jose-Manuel, R.M.; Patrice, N.; Nicolas, F.; Laurent, P. VIM-19, a metallo-beta-lactamase with increased carbapenemase activity from *Escherichia coli* and *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* **2010**, *54*, 471–476.

24. Pierre, B.; Carine, B.; Te-Din, H.; Warda, B.; Yves, D.; Ariane, D.; Kurt, H.; Youri, G. Detection and characterization of VIM-31, a new variant of VIM-2 with Tyr224His and His252Arg mutations, in a clinical isolate of *Enterobacter cloacae*. *Antimicrob. Agents Chemother.* **2012**, *56*, 3283.

25. Corbin, B.D.; Seeley, E.H.; Raab, A.; Feldmann, J.; Miller, M.R.; Torres, V.J.; Anderson, K.L.; Dattilo, B.M.; Dunman, P.M.; Gerads, R.; et al. Metal Chelation and Inhibition of Bacterial Growth in Tissue Abscesses. *Science* **2008**, *319*, 962–965. [CrossRef] [PubMed]

26. Kumar, G.; Issa, B.; Kar, D.; Biswal, S.; Ghosh, A.S. E152A substitution drastically affects NDM-5 activity. *FEMS Microbiol. Lett.* **2017**, *364*. [CrossRef] [PubMed]

27. Ali, A.; Azam, M.W.; Khan, A.U. Non-active site mutation (Q123A) in New Delhi metallo-β-lactamase (NDM-1) enhanced its enzyme activity. *Int. J. Biol. Macromol.* **2018**, *112*, 1272–1277. [CrossRef] [PubMed]

28. Carfi, A.; Pares, S.; Duée, E.; Galleni, M.; Duez, C.; Frère, J.M.; Dideberg, O. The 3-D structure of a zinc metallo-beta-lactamase from Bacillus cereus reveals a new type of protein fold. *EMBO J.* **1995**, *14*, 4914–4921. [CrossRef] [PubMed]

29. Piccirilli, A.; Brisdelli, F.; Aschi, M.; Celenza, G.; Amicosante, G.; Perilli, M. Kinetic Profile and Molecular Dynamic Studies Show that Y229W Substitution in an NDM-1/L209F Variant Restores the Hydrolytic Activity of the Enzyme toward Penicillins, Cephalosporins, and Carbapenems. *Antimicrob. Agents Chemother.* **2019**, *63*, e02270-18. [CrossRef]

30. Meini, M.-R.; Tomatis, P.E.; Weinreich, D.M.; Vila, A.J. Quantitative Description of a Protein Fitness Landscape Based on Molecular Features. *Mol. Boil. Evol.* **2015**, *32*, 1774–1787. [CrossRef]

31. Clinical and Laboratory Standards Institute. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria that Grow Aerobically: Approved Standard*, 11th ed.; CLSI document M07; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2018.

32. Clinical and Laboratory Standards Institute. *Performance Standards for Antimicrobial Susceptibility Testing*, 28th ed.; CLSI document M100; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2018.

33. Crowder, M.W.; Walsh, T.R.; Banovic, L.; Pettit, M.; Spencer, J. Overexpression, purification, and characterization of the cloned metallo-beta-lactamase L1 from Stenotrophomonas maltophilia. *Antimicrob. Agents Chemother.* **1998**, *42*, 921. [CrossRef] [PubMed]

34. De Meester, F.; Joris, B.; Reckinger, G. Automated analysis of enzyme inactivation phenomena. Application to β-lactamases and DD-peptidases. *Biochem. Pharmacol.* **1987**, *36*, 2393–2403. [CrossRef]

35. Segel, I.H. *Biochemical Calculations*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 1976; pp. 236–241.

36. Liu, Z.; Zhang, R.; Li, W.; Yang, L.; Liu, D.; Wang, S.; Shen, J.; Wang, Y. Amino acid changes at the VIM-48 C-terminus result in increased carbapenem resistance, enzyme activity and protein stability. *J. Antimicrob. Chemother.* **2018**, *74*, 885–893. [CrossRef] [PubMed]

37. Raussens, V.; Ruysschaert, J.-M.; Goormaghtigh, E. Protein concentration is not an absolute prerequisite for the determination of secondary structure from circular dichroism spectra: A new scaling method. *Anal. Biochem.* **2003**, *319*, 114–121. [CrossRef]

38. Sreerama, N.; Venyaminov, S.Y.; Woody, R.W. Estimation of the number of α-helical and β-strand segments in proteins using circular dichroism spectroscopy. *Protein Sci.* **1999**, *8*, 370–380. [CrossRef] [PubMed]

39. Sreerama, N.; Venyaminov, S.Y.; Woody, R.W. Analysis of Protein Circular Dichroism Spectra Based on the Tertiary Structure Classification. *Anal. Biochem.* **2001**, *299*, 271–274. [CrossRef] [PubMed]

Decoding Essential Amino Acid Residues in the Substrate Groove of a Non-Specific Nuclease from *Pseudomonas syringae*

Lynn Sophie Schwardmann, Sarah Schmitz, Volker Nölle and Skander Elleuche *

Miltenyi Biotec B.V. & Co. KG, Friedrich-Ebert-Straße 68, 51429 Bergisch Gladbach, Germany;
l.schwardmann@web.de (L.S.S.); sarahs@miltenyibiotec.de (S.S.); VolkerN@miltenyibiotec.de (V.N.)
* Correspondence: skander.elleuche@miltenyibiotec.de; Tel.: +49-2204-8306-4452

Abstract: Non-specific nucleases (NSN) are of interest for biotechnological applications, including industrial downstream processing of crude protein extracts or cell-sorting approaches in microfabricated channels. Bacterial nucleases belonging to the superfamily of phospholipase D (PLD) are featured for their ability to catalyze the hydrolysis of nucleic acids in a metal-ion-independent manner. In order to gain a deeper insight into the composition of the substrate groove of a NSN from *Pseudomonas syringae*, semi-rational mutagenesis based on a structure homology model was applied to identify amino acid residues on the protein's surface adjacent to the catalytic region. A collection of 12 mutant enzymes each with a substitution to a positively charged amino acid (arginine or lysine) was produced in recombinant form and biochemically characterized. Mutations in close proximity to the catalytic region (inner ring) either dramatically impaired or completely abolished the enzymatic performance, while amino acid residues located at the border of the substrate groove (outer ring) only had limited or no effects. A K119R substitution mutant displayed a relative turnover rate of 112% compared to the original nuclease. In conclusion, the well-defined outer ring of the substrate groove is a potential target for modulation of the enzymatic performance of NSNs belonging to the PLD superfamily.

1. Introduction

Non-specific nucleases (NSN) are a group of enzymes that hydrolyze deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) in all conformations, including single-stranded and double-stranded or linear and circularized substrates, without sequence specificity [1]. NSNs are ubiquitously distributed among all organisms and are of great potential for versatile biotechnological and clinical applications [2–4].

Enzymes that are highly indiscriminate towards different substrates are generally considered as potential evolutionary starting points for developing novel or more specific catalytic activities [5–7]. Members of the phospholipase D (PLD; Enzyme Commission number (EC) 3.1.4.4) superfamily are known to accept a wide range of ester substrates, including nucleic acids [8–10]. PLDs are mainly represented in eukaryotes and predominately catalyze the hydrolysis of phosphatidylcholine to produce choline and phosphatidic acid [11]. Moreover, PLDs act as important key players in various physiological processes, including cell migration and membrane trafficking [10]. This family of enzymes usually encodes two copies of the conserved $HxK(x)_4D(x)_6GSxN$ motif in one gene.

A structurally related bacterial enzyme (Nuc) has been initially described from the human pathogenic microorganism *Salmonella enterica* subsp. *enterica* serovar *Typhimurium*. Nuc contains a single $HxK(x)_4D(x)_6GSxN$ motif, but forms a homodimer, and is capable of degrading nucleic acids in a

non-specific manner [12,13]. This enzyme is among the very few known nucleases that are not dependent on a metal ion in its catalytic region, and is therefore of potential interest for biotechnological applications that take place in buffers supplemented with metal ion chelators, such as ethylenediaminetetraacetic acid (EDTA) or ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA). The group of metal-ion-independent nucleases mainly consists of two PLD-like, site-specific restriction endonucleases from *Bacillus firmus* and *Bacillus megaterium*, WSV191 from the white spot syndrome virus and GBSV1-NSN from a thermophilic bacteriophage, as well as the restriction glycosylase R.*Pab*I from the hyperthermophilic archeon *Pyrococcus abyssi* [14–18].

So far, only three isozymes of bacterial PLD-like NSNs, beside Nuc from *S. enterica* subsp. *enterica* serovar *Typhimurium*, have been investigated and described in detail with regard to their biochemical and biophysical properties: (1) *Ec*Nuc from *Escherichia coli* has been shown to be applicable during cell lysis and protein purification in EDTA-containing buffers; and (2) two isozymes from the plant pathogenic competitor bacterium *Pantoea agglomerans* were shown to be the result of an ancient gene duplication event followed by diversification [19,20]. These enzymes are completely devoid of catalytic performance towards lipids and exclusively degrade nucleic acids in a non-specific manner.

In this study, another metal-ion-independent NSN (DNase/D157G) from *Pseudomonas syringae* was mutagenized using a semi-rational strategy to gain a deeper insight into the composition of the substrate groove. Homology modeling was applied to identify amino acid residues on the surface of the NSN in the surrounding of the catalytic site, which is buried at the bottom of the putative substrate groove. It has been shown before that positively charged amino acid residues that can interact with the proximal negatively charged phosphate groups in nucleic acids had a stimulating impact on the catalytic activity of human DNase I [21]. Therefore, positively charged amino acids were introduced at positions on the surface of DNase/D157G within the substrate groove. Two regions were defined that were either directly adjacent to the catalytic site (inner ring) or at the border of the substrate groove (outer ring). Substitutions in the inner ring dramatically impaired or completely abolished the catalytic activity, while mutagenesis in the outer ring had no or little effect. DNase variant K119R/D157G displayed increased activity, the temperature optimum of variant S143R/D157G was shifted from 50 °C to 40 °C, and N95K/D157G and S143K/D157G exhibited an increased tolerance towards 50 mM of EDTA.

2. Results

2.1. Identification of Amino Acid Residues within the Substrate Groove of a NSN from Pseudomonas sp.

Phylogenetic analyses revealed a highly conserved NSN within the genomes of *Pseudomonas* species. These enzymes are related to Nuc from *S. enterica* subsp. *enterica* serovar *Typhimurium* (~57% identity in 159 amino acids overlap) and are highly active at neutral pH, in the presence of salt concentrations up to 250 mM, and in a temperature range between 4 and 50 °C (Supplementary Materials Figure S1). It has been shown that amino acid residues D157, E157, and G157 occur naturally in homologues from the genus *Pseudomonas*. A comparison of the catalytic activities in our laboratory revealed that an enzyme variant containing G157 is superior over a variant with a negatively charged amino acid at position 157 (unpublished results). Therefore, the natural amino acid sequence from a NSN of *P. syringae* containing a single amino acid substitution at position 157 (DNase/D157G) was used in our study as the starting variant for semi-rational mutagenesis to generate double-mutants.

A homology model of the enzyme DNase/D157G was produced based on the crystal structure from *S. enterica* subsp. *enterica* serovar *Typhimurium* (Figure 1). The enzyme is modelled as a hypothetical homodimer, with the catalytic site buried within a putative substrate groove at the dimeric interface. The catalytic site is composed of amino acids H122, K124, G136, S137, and N139, that are part of the HxK(x)$_4$D(x)$_6$GSxN motif, while D129 has been proposed to be of structural relevance [12]. Fifteen different amino acid residues were identified that are present on the protein surface within the substrate groove. These amino acid residues were either assigned to be part of an inner ring that is directly

adjacent to the catalytic site or to an outer ring that surrounded the inner ring amino acid residues. The following amino acids were identified as being located on the protein surface close to the catalytic site: Y63, S64, T66, I120, and S141, while P68, H91, G92 D94, N95, A97, A101, K119, A142, and S143 are defined as being part of the outer ring (Figure A1 Appendix A).
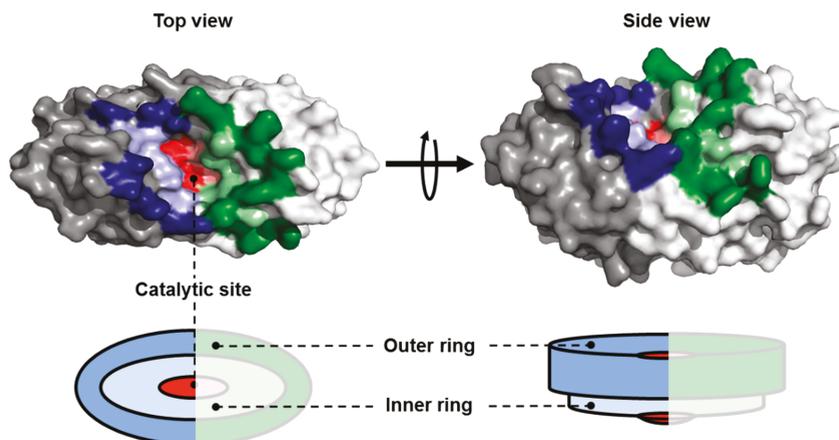

Figure 1. Homology model of non-specific nucleases (NSNs) from *Pseudomonas syringae* in top and side views. Low resolution model of DNase/D157G using Nuc (PDB: 1BYS_A) as a template. PyMOL was used to highlight a hypothetical dimeric structure with dark and light grey monomers. The conserved residues of the HxK(x)$_4$D(x)$_6$GSxN motif are given in dark and light red. Naturally occurring amino acids of the outer and inner rings are highlighted in dark and light blue, while substituted positively charged amino acid residues are indicated in dark and light green. Cartoon illustrations at the bottom are used to simplify the orientation of amino acid residues of the inner and outer ring within the potential substrate groove surrounding the catalytic site.

Surface-presented amino acid residues were exchanged for lysine or asparagine to improve substrate binding and modulate the enzymatic performance or to identify amino acids essential for catalytic activity. Histidine was not considered due to its bulkiness, aggravating the risk for interference with the structure of the protein. Potential steric hindrance was determined by in silico mutagenesis, and the following substitutions were selected: Y63K, S64K, T66R, P68R, H91R, D94K, N95K, K119R, I120K, S141K, A142R, and S143R. Due to the fact that K119 is the only positively charged amino acid within the substrate groove, this lysine was replaced with asparagine. Amino acid residues G92, A97, and A101 were not mutagenized due to expected clashes with adjacent amino acids.

2.2. Production and Purification of Recombinant Nuclease Mutants

Recombinant DNase/D157G double-mutants were produced in *Escherichia coli* Veggie BL21 (DE3), except for mutant P68R/D157G, because this expression strain could not be transformed with the corresponding expression plasmid. Therefore, *Escherichia coli* Veggie BL21 (DE3) pLysS was used for the production of this mutant. All recombinant mutant enzymes were produced in soluble form and could be purified using a two-step approach combining affinity and ion exchange chromatography. Purification strategy was optimized using the Äkta purifier (Figure 2). The purification level of all recombinant nucleases was visualized with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting analyses using an anti-HIS antibody.
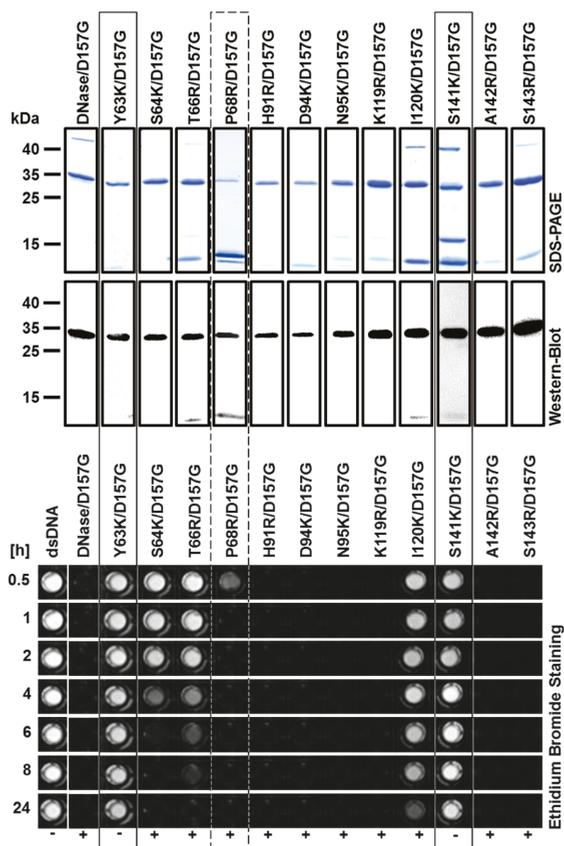
Figure 2. Biochemical and enzymatical analysis of purified nuclease mutants. SDS-PAGE results showed the purity of HIS-tagged nucleases (29 kDa) after a two-step purification approach composed of affinity and ion exchange chromatography (upper row). Western blotting analyses confirmed the presence of HIS-tagged nucleases (middle row). Qualitative activity assays used ethidium bromide staining in 96-well plates (lower row): white colored dots indicate the presence of ethidium bromide intercalating into DNA and grey dots indicate incomplete degradation, while a complete loss of fluorescence is due to complete DNA degradation. Solid line boxes—inactive variants Y63K/D157G and S141K/D157G were purified in non-optimized gravity flow experiments. Dashed line box—variant P68K/D157G was produced in expression strain *E. coli* Veggie BL21(DE3) pLysS. Note: "+" indicates activity and "−" indicates inactivity. SDS-PAGE results, including all purification steps, are shown in Figure A2 Appendix A.

A qualitative assay using ethidium bromide to visualize non-degraded, double-stranded DNA (dsDNA) was applied to demonstrate that mutants Y63K/D157G and S141K/D157G were inactive, while mutant S64K/D157G completely degraded dsDNA after 6 h, mutant T66R/D157G after 24 h, and mutant I120K/D157G partly hydrolyzes DNA after 24 h. These five mutants contain amino acid substitutions within the inner ring of the substrate groove. The remaining mutants as well as the positive control DNase/D157G completely hydrolyzed dsDNA after 30 min (Figure 2).

To ensure that the correct mutants were purified and characterized, the molecular masses of the purified enzymes were determined by mass spectrometry in addition to sequence verification of generated plasmids. Measured molecular weights are in accordance with predicted masses (Table 1). Variant P68R/D157G could not be properly identified due to impurities.

Variant	Theoretical MW (Da) [1]	Measured MW (Da)
DNase/D157G	28,900.30	28,900.61
Y63K/D157G	28,780.19	28,780.30
S64K/D157G	28,856.29	28,856.19
T66R/D157G	28,870.28	28,871.01
P68R/D157G	28,874.27	n.d. [2]
H91R/D157G	28,834.24	28,834.07
D94K/D157G	28,828.28	28,828.56
N95K/D157G	28,829.26	28,829.40
K119R/D157G	28,843.21	28,842.25
I120K/D157G	28,830.21	28,829.26
S141K/D157G	28,856.29	28,856.10
A142R/D157G	28,900.30	28,900.58
S143R/D157G	28,884.30	28,884.27

[1] Theoretical molecular weights (MW) of monomers were calculated using *Compute pI/MW*. [2] n.d.—not detectable.

2.3. Biochemical Properties of Nucleases

In vitro activity assays using dsDNA at 25 °C without the addition of ethidium bromide were conducted to confirm the qualitative plate assays. The reaction was stopped after 1 h of incubation and samples were loaded onto an agarose gel to investigate the level of hydrolysis of sheared dsDNA (Figure 3). As expected, inactive mutant enzymes Y63K/D157G and S141K/D157G were not capable of degrading dsDNA, while recombinant enzymes S64K/D157G, T66R/D157G, and I120K/D157G, respectively, exhibited reduced activity levels compared to the initial nuclease variant DNase/D157G and the remaining mutants.
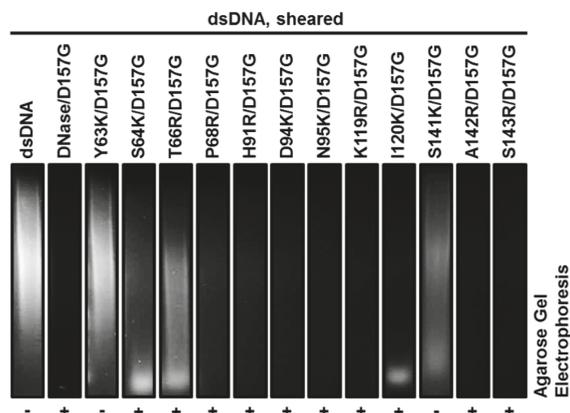
Figure 3. Digestion of sheared dsDNA by nuclease mutants. Sheared dsDNA (UltraPure[TM] Salmon Sperm DNA Solution) exhibited an average size of ≤ 2000 bps. Negative control containing substrate without enzymes is indicated as "dsDNA". The positive control DNase/D157G and mutants S64K/D157G, H91R/D157G, D94K/D157G, N95K/D157G, K119R/D157G, I120K/D157G, A142R/D157G, and S143R/D157G were able to completely hydrolyze sheared dsDNA. S64K/D157G, T66R/D157G, and I120K/D157G partially degraded sheared dsDNA within 1 h at 25 °C. Y63K/D157G and S141K/D157G did not exhibit activity towards sheared dsDNA. Note: "+" indicates activity and "−" indicates inactivity, without any quantification of the activity level.

After 1 h of incubation, mutant enzyme T66R/D157G only initiated the hydrolysis of dsDNA with some low molecular weight fragments visible at the bottom of the agarose gel. Therefore, the catalytic activity of this mutant was investigated with regard to its degradation velocity. Identical concentrations

of mutant enzyme T66R/D157G were incubated with substrate for 30 min, 1 h, 2, h, 4 h, and 24 h, respectively. Reactions were stopped and the samples were loaded onto an agarose gel, revealing that the substrate becomes slowly degraded and is still not completely digested after 4 h of incubation (Figure 4).
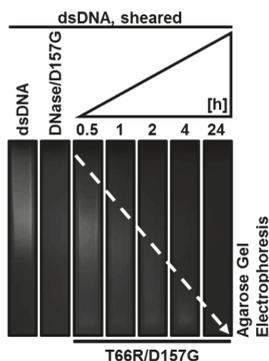

Figure 4. Degradation of sheared dsDNA by mutant T66R/D157G over time. Sheared dsDNA (UltraPureTM Salmon Sperm DNA Solution) exhibited an average size of ≤ 2000 bps. Reaction was stopped after indicated times (between 0.5 and 24 h). White dashed arrow indicates level of nucleic acid molecular weights. The dsDNA is completely hydrolyzed after 24 h of incubation at 25 °C.

To investigate the substrate promiscuity of active recombinant nuclease mutants, the enzymatic hydrolysis was studied towards the following substrates: unsheared dsDNA, single-stranded genomic DNA (ssDNA), circularized DNA, and RNA from bacteriophage MS2. Mutant enzymes with amino acid substitutions in the outer ring of the substrate groove that were active towards sheared dsDNA (Figure 3) also degraded unsheared dsDNA, ssDNA, circularized DNA, and RNA (Figure 5, Figure A3).
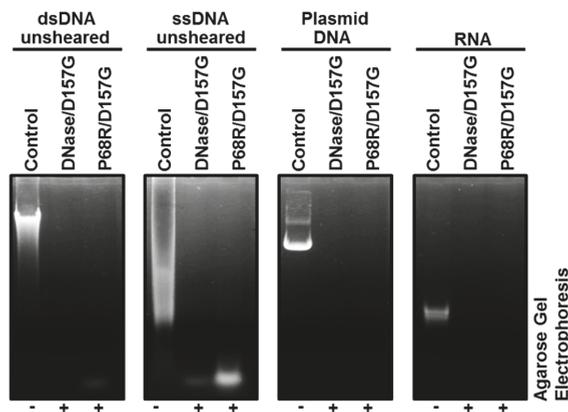

Figure 5. Substrate specificity of mutants with amino acid substitutions in the outer ring. Positive control DNase/D157G and mutant P68R/D157G were incubated for 1 h at 25 °C with different types of DNA and RNA. Further outer ring mutants H91R/D157G, D94K/D157G, N95K/D157G, K119R/D157G, A142R/D157G, and S143R/D157G were also able to degrade all types of nucleic acids (Figure A3 Appendix A). "Control" indicates negative controls containing substrate but no enzyme in the reaction mixture.

Nuclease mutant enzyme Y63K/D157G was also inactive towards unsheared dsDNA, ssDNA, and circularized DNA, while S141K/D157G showed some activity on all substrates except sheared

dsDNA (Figures 3 and 6). However, mutant Y63K/D157G exhibited some activity towards RNA. In good agreement with previous results, mutant enzyme T66R/D157G exhibited reduced activity compared to control DNase/D157G and active mutant enzymes when incubated with both DNA and RNA (Figure 6). In contrast to sheared dsDNA, mutant S64K/D157G completely digested unsheared dsDNA, ssDNA, and circularized DNA, and partially digested RNA, while I120K/D157G was also active on every substrate, but only completely degraded RNA within 1 h at 25 °C (Figure 6).
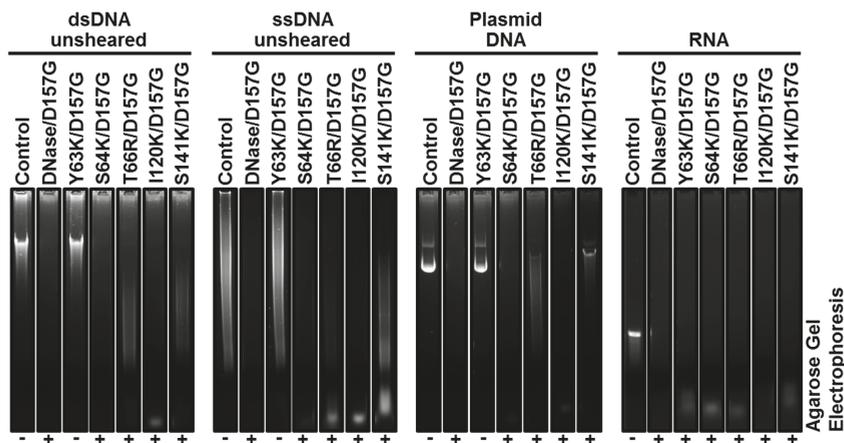
Figure 6. Substrate specificity of mutants with amino acid substitutions in the inner ring. Mutant Y63K/D157G was not able to degrade any type of nucleic acid. Mutants S64K/D157G, T66R/D157G, and I120K/D157G partially degraded all types of nucleic acids within 1 h at 25 °C, while mutant S141K/D157G exhibited low levels of degradation activity. "Control" indicates negative controls containing substrate but no enzyme in the reaction mixture.

Furthermore, pH and temperature optima of active mutants were determined. Every mutant as well as the initial variant DNase/D157G displayed optimal activity at pH 7.0. The temperature optima were around 50 °C for the initial variant DNase/D157G and all active mutants, except for mutants S64K/D157G and S143R/D157G, which showed optimal activities at 60–70 °C and 40 °C, respectively.

2.4. Enzyme Kinetics

Michaelis–Menten kinetics using sheared dsDNA confirmed the result of the quality activity assays: mutant enzymes Y63K/D157G and S141K/D157G did not exhibited any activity in these assays towards sheared dsDNA and the catalytic performances of mutant enzymes S64K/D157G, T66R/D157G, and I120K/D157G were lower compared to the initial nuclease variant DNase/D157G and remaining active mutant enzymes at 25 °C. Relative turnover rates were below 10% compared with DNase/D157G and the catalytic efficiency (k_{cat}/K_M) was below 0.2 s^{-1} μM^{-1} (Table 2). The turnover number of mutant enzyme K119R/D157G (1034 s^{-1}) exclusively surpassed the catalytic activity of DNase/D157G (924 s^{-1}), but the latter mutant enzyme also exhibited the lowest substrate affinity of all mutants (K_M = 428 μM).

Table 2. Kinetic characteristics of nuclease mutants [1].

Variant	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} μM^{-1})	Relative Turnover Rate (%)
DNase/D157G	357	924	2.58	100
Y63K/D157G	n.d. [2]	n.d.	n.d.	-
S64K/D157G	115	8.0	0.07	0.9
T66R/D157G	335	62	0.19	6.7
P68R/D157G	358	106	0.3	11
H91R/D157G	284	290	1.02	31
D94K/D157G	224	170	0.76	18
N95K/D157G	273	455	1.66	49
K119R/D157G	428	1034	2.42	112
I120K/D157G	115	3.0	0.03	0.3
S141K/D157G	n.d.	n.d.	n.d.	-
A142R/D157G	352	707	2.01	77
S143R/D157G	83	153	1.84	17

[1] All experiments were done at 25 °C. [2] Not detectable, below the detection limit.

2.5. EDTA Tolerance of Nuclease Mutants

As metal-ion-independent proteins, NSNs of the PLD-like family are known to be tolerant in the presence of chelating agents, such as EDTA. DNase/D157G displayed > 80% relative activity at a concentration of 20 mM EDTA and still preserved > 60% residual activity in the presence of 50 mM EDTA. Active positively charged mutant enzymes were also tested in the presence of EDTA in concentrations between 1 mM and 50 mM, without any abnormalities, except that mutant enzymes P68R/D157G, N95K/D157G, I120K/D157G, and S143K/D157G were almost not affected by any concentration of EDTA (> 80% residual activity at concentrations up to 50 mM), while mutant enzymes T66R/D157G and K119R/D157G only showed residual activities of 25% and 34% in the presence of 50 mm EDTA, respectively.

3. Discussion

Commercially available NSNs are of high potential for the elimination of nucleic acids during protein downstream processing, to reduce the viscosity, or for prevention of cell clumping in cell sorting approaches [1,20,22]. The prototype non-specific nuclease is the metal-ion-dependent NSN from *Serratia marcescens* that is commercially sold under the trademark "Benzonase® Nuclease" (Merck KGaA, Darmstadt, Germany). This enzyme has been investigated in detail with regard to protein maturation, secretion, catalytic mechanism, and biotechnological applications [2]. Further NSNs, especially metal-ion-independent enzymes, have only been rarely investigated to date [13,16,18–20,23].

In this study, a semi-rational approach was used to select suitable mutation sites within the substrate groove of an NSN from *Pseudomonas syringae*, which were substituted by site-directed mutagenesis against positively charged amino acid residues to modulate the affinity for negatively charged substrates. Critical amino acid residues for the enzymatic performance of different nucleases were routinely identified by site-directed mutagenesis in previous studies [21,24–26].

Under natural conditions, it has been hypothesized that either positive mutations are first installed in a protein, while negative and neutral mutations are accumulated over time (Neo-Darwinian hypothesis), or neutral mutations pave the way for a flexible evolution and positive or negative mutations are installed as a response to certain conditions (competing hypothesis) [27]. By comparing homology models of proteins with singular amino acid substitutions, it is difficult to determine which of the specific mutations evokes an advantageous, neutral, or even deleterious effect on the protein performance, with certain effects on secondary and tertiary structures often being totally unpredictable [28]. Therefore, in silico mutagenesis was exclusively used to predict steric hindrance between introduced charged amino acid residues and amino acids of the catalytic site, or the predicted substrate groove followed by experimental testing of produced mutants.

Fifteen amino acid residues on the protein surface near the catalytic site were identified in a NSN that is highly conserved within the genus *Pseudomonas*. In silico mutagenesis revealed that twelve amino acids could be substituted against either arginine or lysine, without any steric effects in the molecular model. Eleven mutants could be produced in recombinant form in *E. coli* Veggie BL21 (DE3). However, transformation of *E. coli* Veggie BL21 (DE3), with an expression plasmid encoding for mutant P68R/D157G, did not result in any clones, but the recombinant enzyme could be produced in expression strain *E. coli* Veggie BL21 (DE3) pLysS with a very low yield. The additional plasmid pLysS encodes for T7 lysozyme to lower the background expression level of genes under the control of the T7 promoter. Therefore, it can be hypothesized that background expression of the gene-encoding mutant P68R/D157G in pLysS-less expression strains leads to lethality of the host strain. It is worth mentioning that in another experiment, the same results were monitored when proline at position 68 was replaced with negatively charged aspartate in our control. Nevertheless, mutant P68R/D157G only exhibited a turnover rate of 11% compared with the original variant DNase/D157G. However, the reduced activity may also be dependent on protein impurities that were still present after application of a two-step purification approach.

Artificially increasing the positive charge of the putative substrate groove in the NSN did not accelerate the catalytic performance of the enzyme. In another study, the activity of human DNase I also dropped with the addition of basic amino acids compared with the wild-type enzyme. However, suboptimal conditions for the wild-type enzyme, such as increased salt concentrations, accelerated the performance of the mutated enzyme variants [21]. Eleven out of twelve mutants in our portfolio displayed reduced activity at optimal conditions, while K119R/D157G was the only variant exhibiting an increased turnover number compared with DNase/D157G. However, this enzyme was the sole exception with a basic amino acid replaced by another basic amino acid. Two mutations within the inner ring of the substrate groove, namely Y63K and S141K, completely abolished the ability to hydrolyze sheared dsDNA, while mutant S141K/D157G was still capable of partially hydrolyzing unsheared dsDNA and circular plasmid DNA. Both mutants were also active towards RNA, but it is important to note that although all buffers were prepared under sterile conditions, the possibility that the observed activity is an artefact due to contamination with RNase cannot be excluded. Interestingly, asparagine at position 95 within the outer ring of the substrate groove was slightly impaired with regard to k_{cat}. A comparable effect has also been observed in another mutant, in which asparagine was replaced by serine. The latter amino acid occurs naturally in the homologous NSN in some members of the genus *Pseudomonas*.

Mutants were also probed for stability and activity effects in the presence of chelating agents. A slight performance improvement was detected for mutants N95K/D157G and S143K/D157G with regard to the tolerance of high concentrations of EDTA, which were not affected by concentrations up to 50 mM. Metal-ion-dependent DNases are usually completely inhibited by low concentrations of EDTA (1–5 mM), and even some metalloproteins that tolerate EDTA are dramatically impaired by concentrations of 50 mM [29,30]. However, the related NSNs from *Escherichia coli* and *Pantoea agglomerans* were already inhibited at concentrations above 20 mM of EDTA, which is similar to the results obtained with mutant enzymes T66R/D157G and K119R/D157G [19,20]. Nevertheless, these data are in line with the crystal structure of Nuc from *S. enterica* subsp. *enterica* serovar *Typhimurium*, demonstrating that NSNs of the superfamily of PLD proteins are not metal-ion-dependent [12].

The optimal growth temperature of *Pseudomonas syringae* is at 28 °C, but the highest activity of the nuclease and its mutants was determined to take place between 40 and 70 °C [31], These results are in line with previous observation of enzymes derived from psychro-, meso-, and thermophiles that displayed a temperature optimum, which is above their preferred growth temperatures [19,20,29,32,33].

Detailed enzyme characterizations are always a prerequisite for understanding the functionality of enzymes and to enable the modulation of their catalytic performances [34]. It has been shown that the enzyme family of NSNs from the genus *Pseudomonas* is a promising model protein for modifying the catalytic performance with regard to turnover number, temperature optimum, or EDTA tolerance

by single amino acid substitutions. Due to their enzymatic properties, PLD-like NSNs from bacteria are of great potential for versatile biotechnological applications, and for this reason the discovery of novel enzymes, the optimization of available candidates, and the development of further applications are all highly needed [4,20,35–37]. Finally, the discovery of novel candidates and the extensive characterization in combination with straight-forward protein engineering techniques will lead to the production of more tailor-made enzymes for specific biotechnological applications.

4. Materials and Methods

4.1. Strain and Culture Conditions

Escherichia coli strains Veggie BL21 (DE3) and Veggie BL21(DE3) pLysS (both from Merck KGaA, Darmstadt, Germany) were used for gene expression and protein production. *E. coli* strain NEB® 5-alpha (New England Biolabs, Frankfurt/Main, Germany) was used for plasmid propagation and maintenance.

4.2. Computational Sequence Analysis and Structure Modelling

Protein sequence data of a non-specific nuclease (WP_050543862.1) from the gram-negative, ice-nucleating bacterium *Pseudomonas syringae* was identified and biochemically characterized in our laboratory (unpublished results). The naturally occurring amino acid substitution D157G in related homologous sequences was shown to be beneficial for catalytic activity of the enzyme compared to the wild-type polypeptide sequence. The three-dimensional model of the bacterial nuclease DNase/D157G was generated by the SWISS-MODEL online server, using the crystal structure of a homologous nuclease (PDB ID:1BYS_A) from *Salmonella enterica* subsp. *enterica* serovar *Typhimurium* as a template. The structure model was analyzed and visualized using the PyMOL software package (PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC, New York, NY, USA). Amino acid residues that are located on the surface of the predicted substrate groove were identified and chosen to be substituted against positively charged amino acids (lysine or asparagine). Fifteen amino acid residues were identified to be orientated towards the protein surface close to the catalytic region. These amino acids were assigned to two groups: (1) inner ring (closely located to the catalytic site): Y63, S64, T66, I120, S141, and S143; (2) outer ring (distantly located to the catalytic region): P68, H91, G92, D94, N95, A97, A101, K119, and A142 (Figure A1 Appendix A). In silico mutagenesis was performed to discriminate between lysine and asparagine residue, replacing amino acid residues that are located on the surface of the predicted substrate groove. The substitution with the preferred amino acid residue, either asparagine or lysine, did not result in any steric clashes with adjacent amino acids in the protein for 12 out of 15 amino acids. Therefore, amino acid residues G92, A97, and A101 were excluded from site-directed mutagenesis.

4.3. Cloning of Nuclease Variants

Genes-encoding nuclease variants with amino acid substitutions were codon-optimized for expression in *E. coli* and synthesized by ATUM (Newark, CA, USA). Flanking *Nco*I and *Aat*II restriction sites were used for unidirectional ligation into linearized vector pET24d(+) (Merck KGaA, Darmstadt, Germany), equipped with a double HIS tag. Sequence verification of inserted genes was done by Eurofins Genomics (Ebersberg, Germany).

4.4. Gene Expression and Protein Purification

Expression of genes in *E. coli* Veggie BL21 (DE3) was performed as described earlier [20]. In brief, nuclease variants were produced with *E. coli* Veggie BL21(DE3) in 1 L cultures in 2 L Erlenmeyer shaking flasks. It was not possible to transform *E. coli* Veggie BL21 (DE3) with a plasmid coding for the P68K/D157G mutant in our control. Therefore, *E. coli* BL21 (DE3) pLys was used as an alternative expression host for this NSN variant. Cells were grown under constant shaking (250 rpm) at 37 °C until an optical density at 600 nm (OD_{600}) of 0.6–0.8 was reached. Afterwards, gene expression was induced

by the addition of 0.4 mM isopropyl β-d-1-thiogalactopyranoside (IPTG). Cells were harvested 4 h post-induction by centrifugation for 15 min at 4 °C and 2880× *g*.

Cells were disrupted by high-pressure homogenization (constant cell disruption systems, Constant Systems Limited, Northants, UK) at 5 °C and 1250 bar. Pelleted cells were dissolved in lysis buffer (50 mM NaPO$_4$, pH 7.3): 1 g per 5 mL with a minimal volume of 20 mL. Crude protein extract was incubated for 30 min at 37 °C to enable digestion of nucleic acids by the recombinantly expressed nuclease. Afterwards, the sample was centrifuged at 4000× *g* for 30 min at 4 °C. HIS-tagged fusion enzymes were purified in a two-step approach using a combination of affinity (AC) and ion-exchange chromatography (IEX). Initially, gravity flow experiments using Ni sepharose 6 Fast Flow and SP Sepharose Fast Flow cation exchange chromatography resins (both GE Healthcare, Munich, Germany) were done to test activity of partly purified mutant enzymes. Afterwards, the Äkta purifier (GE Healthcare, Munich, Germany) was used to optimize the purification strategy. Supernatant from cell disruption was loaded onto a HisTrap FF Crude histidine-tagged protein purification column (GE Healthcare, Munich, Germany) equilibrated with 50 mM NaPO$_4$, 50 mM NaCl, 5 mM imidazole, pH 7.3. The loaded column was connected to an Äkta protein purifier system and washed with 10 column volumes of 50 mM NaPO$_4$, 50 mM NaCl, 50 mM imidazole, pH 7.3, prior to the elution with 50 mM NaPO$_4$, 50 mM NaCl, 500 mM imidazole, pH 7.3. For the purification by IEX, the eluate from the Ni sepharose affinity purification was diluted 1:4 with IEX running buffer (25 mM NaPO$_4$, pH 6.0) to reach a conductivity of 7–8 mS/cm. A HiTrap-SP FF column (GE Healthcare, Munich, Germany) was equilibrated with 10 column volumes of IEX running buffer, before the sample was loaded and washed with 10 column volumes of 50 mM sodium phosphate, 300 mM NaCl, pH 6.0, prior to the elution with 50 mM sodium phosphate, 700 mM NaCl, pH 6.0. All steps were performed at a flow velocity of 1 mL/min. Finally, a PD-10 desalting column (GE Healthcare, Munich, Germany) was used to replace the IEX elution buffer with storage buffer (50 mM NaPO$_4$, 25 mM NaCl, pH 7.3). SDS-PAGE (ProGel Tris Glycin 4–20%, Anamed Elektrophorese GmbH, Groß-Bierberau/Rodau, Germany) in combination with a Western blot using a nitrocellulose blotting membrane (GE Healthcare, Munich, Germany) was used for visualization of recombinant nuclease mutants. An anti-HIS horseradish peroxidase (HRP) antibody (Miltenyi Biotec B.V. & Co. KG, Bergisch Gladbach, Germany) and the Immobilon™ Western HRP substrate (Merck, Darmstadt, Germany) were used to detect HIS-tagged proteins.

Enzyme concentrations were measured using a (micro-) Bradford approach in 96-well plate format. A plate reader (EMax, Molecular Devices, San Jose, CA, USA) was used to determine absorbances at 590 nm, which were evaluated with the software SoftmaxPro V5 (Molecular Devices, San Jose, CA, USA). The identities of the nuclease mutants based on their molecular masses were verified with the micrOTOF-Q II Benchtop Mass Spectrometer (Bruker, Billerica, MA, USA).

4.5. Enzyme Activity Assays

Enzyme activity was determined both qualitatively and quantitatively based on the depolymerization of nucleic acids. All activity assays were done at pH 7, which has been determined to be the optimal pH. Due to the limited stability of nuclease mutants at elevated temperatures, all assays were done at 25 °C. No thermal effect on the stability was observed for any mutant at this temperature. (1) Qualitative ethidium bromide staining: this assay was adopted from [4]. Recombinant enzyme was incubated with 5 μg sheared double-stranded genomic DNA (dsDNA), namely UltraPure™ Salmon Sperm DNA Solution (Thermo Fisher Scientific, Darmstadt, Germany). This method is based on quenching of fluorescence of ethidium bromide intercalated into DNA. Repeated fluorescence recordings were taken using a VWR® imager (VWR international, Radnor, PA, USA) over a duration of up to 24 h. (2) Qualitative visualization by agarose gel electrophoresis: this assay was adopted from [20]. Substrate specificity was tested with sheared dsDNA, unsheared dsDNA, namely deoxyribonucleic acid from calf thymus (Sigma-Aldrich, St. Louis, MO, USA), single-stranded genomic DNA (ssDNA) from calf thymus (Sigma-Aldrich, St. Louis, USA), RNA from bacteriophage MS2 (Sigma-Aldrich, St. Louis, USA), and circularized plasmid DNA. Qualitative levels of activity are exclusively interpreted

as "+" (active enzyme) and "−" (inactive enzyme) in Figure 3, Figure 5, and Figure 6, and do not allow any quantification of activities. (3) Quantitative measurements were done in Corning® 96-well UV-transparent plates (Merck KGaA, Darmstadt, Germany) using the Victor™ X4 Multilabel Plate Reader (PerkinElmer, Rodgau, Germany), as described previously [20]. Standard activity assays were conducted in 50 mM sodium phosphate buffer at pH 7.3. For the generation of pH profiles, reactions were performed in 50 mM sodium acetate buffer at pH 5 and 6, in 50 mM sodium phosphate buffer at pH 6, 7, and 8, and in Tris/HCl buffer at pH 7, 8, and 9. Temperature profiles were conducted in the range of 10 to 90 °C. EDTA tolerance was tested at the following concentrations: 0, 1, 2, 5, 10, 20, and 50 mM. (4) Kinetic parameters were determined with V_{max} and K_M obtained from Michaelis–Menten technique by non-linear regression, as described previously [20,38]. It has been speculated that the initial reaction rate of high molecular weight substrates is reduced at high substrate concentrations due to the extension of the lag phase [39]. Therefore, maximum reaction rates were evaluated for each substrate concentration to determine defined kinetic parameters. All experiments were done in triplicate. The error level was below 10%.

5. Conclusions

The composition of the substrate groove was investigated by a combination of structural modelling, multiple sequence alignment, and site-directed mutagenesis. Amino acid residues that are in close proximity to the catalytic site (inner ring of the substrate groove) are of tremendous importance for proper activity of NSN, while amino acids at the border of the substrate groove (outer ring) are promising targets for modulation of the enzymatic properties with regard to turnover number, EDTA tolerance, and temperature preference.

6. Patents

A patent application describing the utilization of non-specific nucleases from the genus *Pseudomonas* and their application potential in cell-sorting approaches has been submitted by Miltenyi Biotec B.V. & Co. KG.
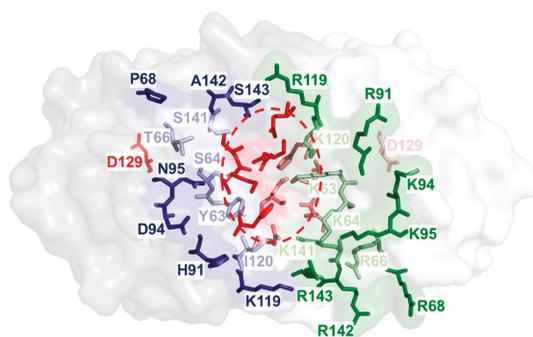
Appendix A

Figure A1. Homology model of NSN from *Pseudomonas syringae* in top view. Amino acid residues of the conserved HxK(x)$_4$D(x)$_6$GSxN motif are given as sticks in dark and light red in the respective monomer. Numbering of amino acids that are part of the catalytic site was omitted for clarity (encircled in red, dashed line), except for amino acid residue D129, which is part of the HxK(x)$_4$D(x)$_6$GSxN motif, but not part of the catalytic site. Naturally occurring amino acids of the outer and inner rings are also given as sticks and highlighted in dark and light blue, while substituted positively charged amino acid residues are indicated in dark and light green.
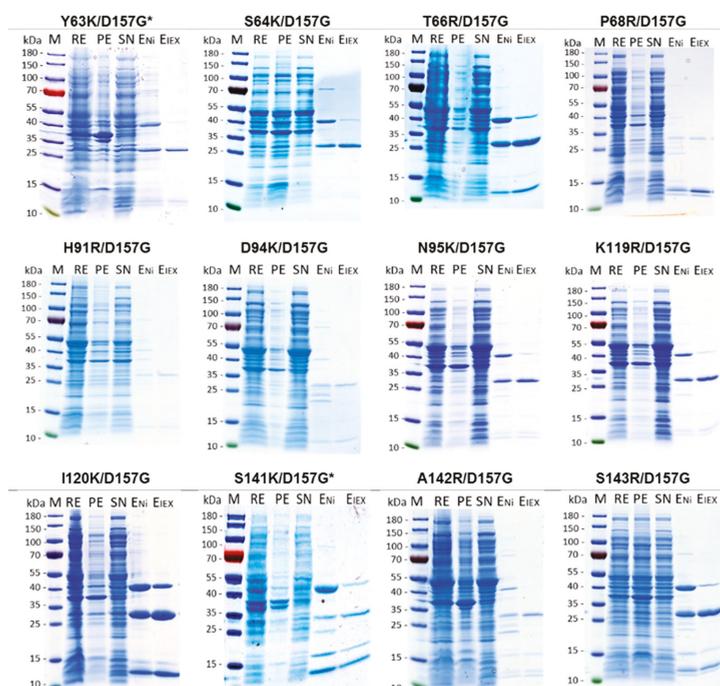
Figure A2. SDS-PAGE results illustrating the purification of all mutant enzymes. M—protein marker, RE—crude extract, PE—pellet, SN—supernatant, E$_{Ni}$—elution fraction Ni-agarose, E$_{IEX}$—elution fraction ion exchange chromatography. Asterisks (*) indicate that inactive mutants Y63K/D157G and S141K/D157G were purified using non-optimized gravity flow purification approaches, resulting in lower purities.
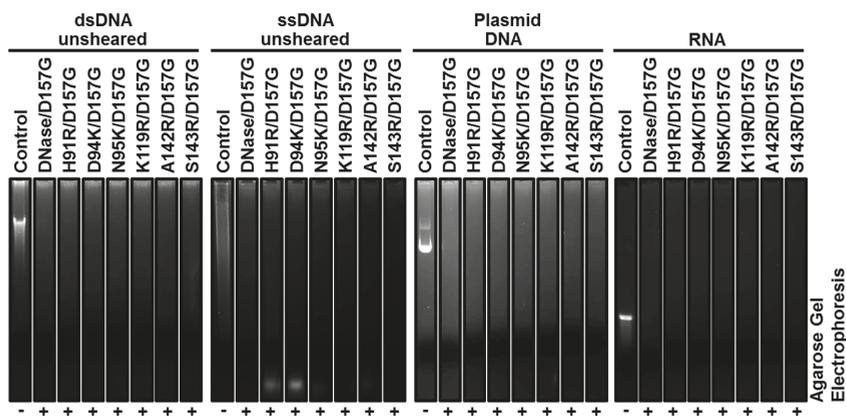
Figure A3. Substrate specificity of mutants with amino acid substitutions in the outer ring. All incubations were done for 1 h at 25 °C. "Control" indicates negative controls containing substrate but no enzyme in the reaction mixture.

References

1. Rangarajan, E.S.; Shankar, V. Sugar non-specific endonucleases. *FEMS Microbiol. Rev.* **2001**, *25*, 583–613. [CrossRef]
2. Benedik, M.J.; Strych, U. *Serratia marcescens* and its extracellular nuclease. *FEMS Microbiol. Lett.* **1998**, *165*, 1–13. [CrossRef] [PubMed]
3. Dang, G.; Cui, Y.; Wang, L.; Li, T.; Cui, Z.; Song, N.; Chen, L.; Pang, H.; Liu, S. Extracellular sphingomyelinase Rv0888 of *Mycobacterium tuberculosis* contributes to pathological lung injury of *Mycobacterium smegmatis* in mice via inducing formation of neutrophil extracellular traps. *Front. Immunol.* **2018**, *9*, 677. [CrossRef] [PubMed]
4. Vafina, G.; Zainutdinova, E.; Bulatov, E.; Filimonova, M.N. Endonuclease from gram-negative Bacteria *Serratia marcescens* is as effective as Pulmozyme in the hydrolysis of DNA in sputum. *Front. Pharmacol.* **2018**, *9*, 114. [CrossRef] [PubMed]
5. Khersonsky, O.; Tawfik, D.S. Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **2010**, *79*, 471–505. [CrossRef] [PubMed]
6. Yan, X.; Wang, J.; Sun, Y.; Zhu, J.; Wu, S. Facilitating the evolution of esterase activity from a promiscuous enzyme (Mhg) with catalytic functions of amide hydrolysis and carboxylic acid perhydrolysis by engineering the substrate entrance tunnel. *Appl. Environ. Microbiol.* **2016**, *82*, 6748–6756. [CrossRef]
7. Rao, S.J.; Shukla, E.; Bhatia, V.; Lohiya, B.; Gaikwad, S.M.; Kar, A.; Pal, J.K. The *Leishmania donovani* IMPACT-like protein possesses non-specific nuclease activity. *Int. J. Biol. Macromol.* **2018**, *119*, 962–973. [CrossRef] [PubMed]
8. Ponting, C.P.; Kerr, I.D. A novel family of phospholipase D homologues that includes phospholipid synthases and putative endonucleases: Identification of duplicated repeats and potential active site residues. *Protein Sci. A Publ. Protein Soc.* **1996**, *5*, 914–922. [CrossRef]
9. Rudolph, A.E.; Stuckey, J.A.; Zhao, Y.; Matthews, H.R.; Patton, W.A.; Moss, J.; Dixon, J.E. Expression, characterization, and mutagenesis of the *Yersinia pestis* murine toxin, a phospholipase D superfamily member. *J. Biol. Chem.* **1999**, *274*, 11824–11831. [CrossRef]
10. Nelson, R.K.; Frohman, M.A. Physiological and pathophysiological roles for phospholipase D. *J. Lipid Res.* **2015**, *56*, 2229–2237. [CrossRef]
11. Liscovitch, M.; Czarny, M.; Fiucci, G.; Tang, X. Phospholipase D: Molecular and cell biology of a novel gene family. *Biochem. J.* **2000**, *345*, 401–415. [CrossRef] [PubMed]
12. Stuckey, J.A.; Dixon, J.E. Crystal structure of a phospholipase D family member. *Nat. Struct. Biol.* **1999**, *6*, 278–284. [CrossRef] [PubMed]

13. Zhao, Y.; Stuckey, J.A.; Lohse, D.L.; Dixon, J.E. Expression, characterization, and crystallization of a member of the novel phospholipase D family of phosphodiesterases. *Protein Sci.* **1997**, *6*, 2655–2658. [CrossRef] [PubMed]

14. Bao, Y.; Higgins, L.; Zhang, P.; Chan, S.H.; Laget, S.; Sweeney, S.; Lunnen, K.; Xu, S.Y. Expression and purification of *Bmr*I restriction endonuclease and its N-terminal cleavage domain variants. *Protein Expr. Purif.* **2008**, *58*, 42–52. [CrossRef] [PubMed]

15. Grazulis, S.; Manakova, E.; Roessle, M.; Bochtler, M.; Tamulaitiene, G.; Huber, R.; Siksnys, V. Structure of the metal-independent restriction enzyme *Bfi*I reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15797–15802. [CrossRef] [PubMed]

16. Song, Q.; Zhang, X. Characterization of a novel non-specific nuclease from thermophilic bacteriophage GBSV1. *BMC Biotechnol.* **2008**, *8*, 43. [CrossRef] [PubMed]

17. Wang, D.; Miyazono, K.I.; Tanokura, M. Tetrameric structure of the restriction DNA glycosylase R.*Pab*I in complex with nonspecific double-stranded DNA. *Sci. Rep.* **2016**, *6*, 35197. [CrossRef]

18. Li, L.; Lin, S.; Yanga, F. Functional identification of the non-specific nuclease from white spot syndrome virus. *Virology* **2005**, *337*, 399–406. [CrossRef]

19. Schmitz, S.; Börner, P.; Nölle, V.; Elleuche, S. Comparative analysis of two non-specific nucleases of the phospholipase D family from the plant pathogen competitor bacterium *Pantoea Agglomerans*. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 2635–2648. [CrossRef]

20. Schmitz, S.; Nölle, V.; Elleuche, S. A non-specific nucleolytic enzyme and its application potential in EDTA-containing buffer solutions. *Biotechnol. Lett.* **2019**, *41*, 129–136. [CrossRef]

21. Pan, C.Q.; Lazarus, R.A. Hyperactivity of human DNase I variants. Dependence on the number of positively charged residues and concentration, length, and environment of DNA. *J. Biol. Chem.* **1998**, *273*, 11701–11708. [CrossRef] [PubMed]

22. Miltenyi, S.; Hübel, T.; Nölle, V. Process for Sorting Cells by Microfabricated Components Using a Nuclease. US Patent 10,018,541 B2, 10 July 2018.

23. Belkebir, A.; Azeddoug, H. Characterization of *Lla*KI, a new metal ion-independent restriction endonuclease from *Lactococcus lactis* KLDS4. *ISRN Biochem.* **2012**, *2012*, 287230. [CrossRef] [PubMed]

24. Friedhoff, P.; Kolmes, B.; Gimadutdinow, O.; Wende, W.; Krause, K.L.; Pingoud, A. Analysis of the mechanism of the *Serratia* nuclease using site-directed mutagenesis. *Nucleic Acids Res.* **1996**, *24*, 2632–2639. [CrossRef] [PubMed]

25. Zhang, Y.; Li, Z.H.; Zheng, W.; Tang, Z.X.; Shi, L.E. Enzyme activity and thermostability of a non-specific nuclease from *Yersinia enterocolitica supsp. palearctica* by site-directed mutagenesis. *Electron. J. Biotechnol.* **2016**, *24*, 32–37. [CrossRef]

26. Franke, I.; Meiss, G.; Pingoud, A. On the advantage of being a dimer, a case study using the dimeric *Serratia* nuclease and the monomeric nuclease from *Anabaena* sp. strain PCC 7120. *J. Biol. Chem.* **1999**, *274*, 825–832. [CrossRef] [PubMed]

27. Bommarius, A.S.; Paye, M.F. Stabilizing biocatalysts. *Chem. Soc. Rev.* **2013**, *42*, 6534–6565. [CrossRef] [PubMed]

28. Schaefer, C.; Rost, B. Predict impact of single amino acid change upon protein structure. *BMC Genom.* **2012**, *13*, S4. [CrossRef]

29. Elleuche, S.; Fodor, K.; Klippel, B.; von der Heyde, A.; Wilmanns, M.; Antranikian, G. Structural and biochemical characterisation of a NAD(+)-dependent alcohol dehydrogenase from *Oenococcus oeni* as a new model molecule for industrial biotechnology applications. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 8963–8975. [CrossRef]

30. Marcal, D.; Rego, A.T.; Fogg, M.J.; Wilson, K.S.; Carrondo, M.A.; Enguita, F.J. Crystallization and preliminary X-ray characterization of 1,3-propanediol dehydrogenase from the human pathogen *Klebsiella pneumoniae*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2007**, *63*, 249–251. [CrossRef]

31. Xin, X.F.; Kvitko, B.; He, S.Y. *Pseudomonas syringae*: What it takes to be a pathogen. *Nat. Rev. Microbiol.* **2018**, *16*, 316–328. [CrossRef]

32. Daniel, R.M.; Danson, M.J.; Eisenthal, R.; Lee, C.K.; Peterson, M.E. The effect of temperature on enzyme activity: New insights and their implications. *Extremophiles* **2008**, *12*, 51–59. [CrossRef] [PubMed]

33. Feller, G.; Narinx, E.; Arpigny, J.L.; Zekhnini, Z.; Swings, J.; Gerday, C. Temperature dependence of growth, enzyme secretion and activity of psychrophilic Antarctic bacteria. *Appl. Microbiol. Biotechnol.* **1994**, *41*, 477–479. [CrossRef]

34. Erb, T.J. Back to the future: Why we need enzymology to build a synthetic metabolism of the future. *Beilstein J. Org. Chem.* **2019**, *15*, 551–557. [CrossRef] [PubMed]

35. Fang, X.J.; Tang, Z.X.; Li, Z.H.; Zhang, Z.L.; Shi, L.E. Production of a new non-specific nuclease from *Yersinia enterocolitica subsp. palearctica*: Optimization of induction conditions using response surface methodology. *Biotechnol. Biotechnol. Equip.* **2014**, *28*, 559–566. [CrossRef] [PubMed]

36. Maciejewska, N.; Walkusz, R.; Olszewski, M.; Szymanska, A. New nuclease from extremely psychrophilic microorganism *Psychromonas ingrahamii* 37: Identification and characterization. *Mol. Biotechnol.* **2019**, *61*, 122–133. [CrossRef] [PubMed]

37. Anisimova, V.E.; Shcheglov, A.S.; Bogdanova, E.A.; Rebrikov, D.V.; Nekrasov, A.N.; Barsova, E.V.; Shagin, D.A.; Lukyanov, S.A. Is crab duplex-specific nuclease a member of the *Serratia* family of non-specific nucleases? *Gene* **2008**, *418*, 41–48. [CrossRef] [PubMed]

38. MacLellan, S.R.; Forsberg, C.W. Properties of the major non-specific endonuclease from the strict anaerobe *Fibrobacter succinogenes* and evidence for disulfide bond formation in vivo. *Microbiology* **2001**, *147*, 315–323. [CrossRef]

39. Friedhoff, P.; Meiss, G.; Kolmes, B.; Pieper, U.; Gimadutdinow, O.; Urbanke, C.; Pingoud, A. Kinetic analysis of the cleavage of natural and synthetic substrates by the *Serratia* nuclease. *Eur. J. Biochem.* **1996**, *241*, 572–580. [CrossRef]

Effects of Random Mutagenesis and *In Vivo* Selection on the Specificity and Stability of a Thermozyme

Giuseppe Perugino [1,†], **Andrea Strazzulli** [2,3,†], **Marialuisa Mazzone** [1], **Mosè Rossi** [1] and **Marco Moracci** [1,2,3,*]

[1]   Institute of Biosciences and BioResources - National Research Council of Italy, Via P. Castellino 111, 80131 Naples, Italy; giuseppe.perugino@ibbr.cnr.it (G.P.); marialuisam76@yahoo.it (M.M.); mose.rossi@ibbr.cnr.it (M.R.)

[2]   Department of Biology, University of Naples "Federico II", Complesso Universitario di Monte S. Angelo, Via Cintia 21, 80126 Napoli, Italy; andrea.strazzulli@unina.it

[3]   Task Force on Microbiome Studies, University of Naples Federico II, 80126 Naples, Italy

*   Correspondence: marco.moracci@unina.it; Tel.: +39-081-679046; Fax: +39-081-679233

†   These authors contributed equally to the paper as First Authors.

Abstract: Factors that give enzymes stability, activity, and substrate recognition result from the combination of few weak molecular interactions, which can be difficult to study through rational protein engineering approaches. We used irrational random mutagenesis and in vivo selection to test if a β-glycosidase from the thermoacidophile *Saccharolobus solfataricus* (*Ss*β-gly) could complement an *Escherichia coli* strain unable to grow on lactose. The triple mutant of *Ss*β-gly (S26L, P171L, and A235V) was more active than the wild type at 85 °C, inactivated at this temperature almost 300-fold quicker, and showed a 2-fold higher k_{cat} on galactosides. The three mutations, which were far from the active site, were analyzed to test their effect at the structural level. Improved activity on galactosides was induced by the mutations. The S26L and P171L mutations destabilized the enzyme through the removal of a hydrogen bond and increased flexibility of the peptide backbone, respectively. However, the flexibility added by S26L mutation improved the activity at T > 60 °C. This study shows that random mutagenesis and biological selection allowed the identification of residues that are critical in determining thermal activity, stability, and substrate recognition.

Keywords: carbohydrate active enzymes; archaea; glycosidase; *Sulfolobus solfataricus*; *Saccharolobus solfataricus*

1. Introduction

The ability of enzymes to recognize specific substrates and catalyzing reactions remaining stable even at extreme conditions is driven by molecular factors that are only poorly understood. In this regard, enzymes from thermophiles and hyperthermophiles, microorganisms growing at temperatures up to 80 °C or above, respectively, have been of long standing research interest, because of their ability to promote chemical reactions at temperatures close to the boiling point of water when the majority of enzymes from mesophiles are denatured. Consequently, extensive biochemical and structural studies on thermozymes (enzymes that function at high temperatures) have shown that their stability is not governed by peculiar amino acids, covalent modifications, or structural motifs, but rather by subtle redistributions of the same intramolecular interactions that are required to stabilize proteins from mesophiles (for reviews, see the literature [1–5]).

In thermozymes, these interactions allow both a balance between molecular stability to avoid unfolding at very high temperatures and the flexibility to promote catalysis at a metabolically appropriate rate. However, thermozymes persist for a long time during incubation at high temperatures

(thermal stability), and increase their catalytic rate with temperature (thermal activity or thermophilicity). In particular, thermozymes need high temperatures to have specific activity comparable to that of mesozymes at 37 °C. Comparisons of homologous enzymes from organisms spanning a wide range of thermal habitats show that adaptive mutations maintain a balance between thermal stability and activity, regardless of the temperature at which the enzyme operates [1].

Despite extensive studies, the identification of the molecular interactions that are responsible for these properties remained elusive. Most studies aimed to increase the thermal stability of mesozymes through rational design and enzyme engineering or random mutagenesis and appropriate selection methods to obtain new catalysts for biotechnological applications [6–9]. Even fewer examples exist of the laboratory evolution of thermozymes adapted to operational temperatures lower than their normal range of physiological conditions [10,11].

The objective of this work was to analyze whether and how the substrate specificity, thermal activity, and resistance to heat of a thermozyme are bound at the molecular level by combining random mutagenesis and in vivo selection of mutants in a mesophilic host. Thus, the selective pressure might mimic that encountered during natural evolution more closely. To this aim, we focused on a glycoside hydrolase from an Archaeon, as these thermozymes are easy to assay and have interesting biotechnological applications [12–15].

The model system in this study is the β-glycosidase from *Saccharolobus solfataricus*—previously *Sulfolobus solfataricus*—(Ssβ-gly). This enzyme, belonging to family 1 (GH1) of carbohydrate active enzymes (CAZy) classification (www.cazy.org [16]), which was firstly discovered by our group and extensively characterized [17–20], is a well-known model system for studies of molecular stabilization and catalysis engineering. In fact, studies on this enzyme allowed the identification of general structural and functional elements that make thermozymes more stable than mesozymes [21,22]. In particular, ion-pair networks between protein domains and subunits were identified as major players of the molecular stabilization and resilience of thermozymes [23–28]. In addition, studies on Ssβ-gly allowed us to demonstrate that the activity and stability of enzymes can be restored by small molecules that bind to allosteric sites [29,30].

Ssβ-Gly catalyzes a single-substrate reaction by promoting the hydrolysis of β-D-O-glycosides, including disaccharides (laminaribiose, cellobiose, and lactose) and aryl-glycosides (substituted nitrophenyl-β-D-gluco-, galacto-, fuco, and xylosides), thereby showing wide substrate specificity [31]. The simple *retaining* reaction mechanism followed by this enzyme and its versatility has allowed for extensive protein engineering studies that have led to mutants with different substrate specificity [32,33]. The modification of Ssβ-gly active site residues changed the reaction mechanism [34] and allowed the production of a novel class of enzymes, named *glycosynthases*, that promote the synthesis, rather than then hydrolysis, of oligosaccharides [35–37]. These fundamental studies led to the development of many other *glycosynthases* of technological interest [38–43].

The number of catalytic studies on Ssβ-gly, its well-known reaction mechanism, and the availability of mutants and of several three-dimensional (3D) structures at a high resolution [22,23,34,44,45] prompted us to study the relationship between substrate specificity, catalysis at a high temperature, and structural stability. We addressed this issue by following a non-rational approach exploiting random mutagenesis and in vivo selection of mutants. This was made possible because the *lacS* gene, encoding for Ssβ-gly, cannot promote the growth on lactose minimal medium of an *Escherichia coli* strain lacking the endogenous *lacZ* gene. We wanted to test whether random mutations on a thermozyme from a hyperthermophilic Archaeon (T_{growth} = 80 °C), easily and quickly produced by a chemical mutagen, could lead to novel functional β-galactosidase(s) able to promote growth of a bacterium at 37 °C on lactose as a unique carbon source. The in vivo selection is expected to affect, simultaneously, the substrate specificity and thermal activity of the enzyme, as lactose is a non-preferred substrate of Ssβ-gly, the catalyze reaction of which occurs at T > 80 °C [20,31]. The remarkable Ssβ-gly stability, about 500 min half-life at 75 °C [20], should not be directly affected by the selection method, but may be a side effect if thermal activity and stability are interconnected.

By following this approach, we report here that the selection of a *Ss*β-gly mutant library led to a triple mutant, named H7, able to complement a deficient strain of *E. coli*. The three mutations, S26L, P171L, and A235V, were in regions of the enzyme far from the active site, which could not be planned if one would rationally design an improved β-galactosidase activity. H7 catalyzed the hydrolysis of lactose 1.8-fold faster than the wild type, and, surprisingly, at 85 °C, was 2-fold more active than the wild type and less stable at 75 °C and 85 °C by 23-fold and 283-fold, respectively. The characterization of single and double mutants and the inspection of *Ss*β-gly 3D-structure models allowed hypothesizing some explanations to rationalize the effect of the mutations. The results are also discussed in term of the selective pressure produced by the molecular evolution method used.

2. Results and Discussion

2.1. Production of Ssβ-gly Complementing Growth on Lactose at 37 °C

The pGEX-K-Gly plasmid containing the *Saccharolobus solfataricus* β-glycosidase gene (*lacS*) was subjected to a reaction of chemical mutagenesis with hydroxylamine. This mutagen is known to cause transitions on the type C→T DNA [46]. The set of mutated genes was isolated from the plasmid and inserted in an untreated plasmid to avoid any mutagenic affects to the other regions (promoter, antibiotic resistance, replication origin, etc.).

This random library was used to transform the *E. coli* M5154 strain, which has a large deletion on the lactose operon that, unlike the commercial *E. coli* strains, affects only the β-galactosidase gene (*lacZ*) and not the sugar permease (*lacY*). This makes this strain unable to metabolize lactose as a sole carbon source, even in the presence of the sugar in the cytoplasm. Thus, *E. coli* M5154 can grow in lactose minimal medium, only if an active β-galactosidase is also expressed. The selection thus concerned only the clones bearing *Ss*β-gly mutants able to promote growth faster than the wild type enzyme that complemented this *E. coli* strain after 10 days at 37 °C in a lactose minimal medium.

The cells transformed with the plasmid library were plated on rich solid medium containing 5-bromo-4-chloro-3-indolyl-β-ᴅ-glucopyranoside (X-Glc) chromogenic substrate to verify the efficiency of the random mutagenesis. The presence of blue and white colonies ensured that the mutagen had an effect resulting in inactive *Ss*β-gly mutants. Then, the same cells transformed with the mutant library were plated on minimal solid medium containing lactose as the sole carbon source and incubated at 37 °C. The colonies appearing before the wild type control were identified and chosen. The screening of >6000 clones led to the selection of 39 colonies that were isolated and grown on rich liquid medium. Cells were assayed for β-galactosidase activity on 2-nitrophenyl-β-ᴅ-galacto- and glucopyranoside substrates (2Np-Gal and 2Np-Glc, respectively), and the calculated enzymatic units were compared to those of the wild type enzyme in the same conditions (see Section 3). This analysis led to a further selection of six clones from the starting 39, and one specific mutant, called *Ss*β-gly-H7, hereinafter H7, was characterized in detail.

The H7 gene was completely sequenced: as expected, using the hydroxylamine mutagenic agent, three C→T transitions were found, leading to three aminoacid substitutions; namely, S26L, P171L, and A235V. From the analysis of the three-dimensional structure of *Ss*β-gly, all of them were located far from the active site at 23 Å (S26), 25 Å (P171), and 18 Å (A235), respectively (Figure 1). S26L and P171L were on the surface of the protein, while A235V was completely buried. Wild type *lacS* and H7 genes were expressed in *E. coli* strain BL21(DE3)RIL, fused to glutathione S-transferase, as described in Materials and Methods, and the purified enzymes were characterized in detail. Similarly, S26L, P171L, and A235V single mutants, as well as S26L/P171L, S26L/A235V, and P171L/A235V double mutants, were prepared by subcloning procedures. The corresponding genes were expressed and the recombinant enzymes purified to determine the single contribution of each identified mutation (see **Materials and Methods**).
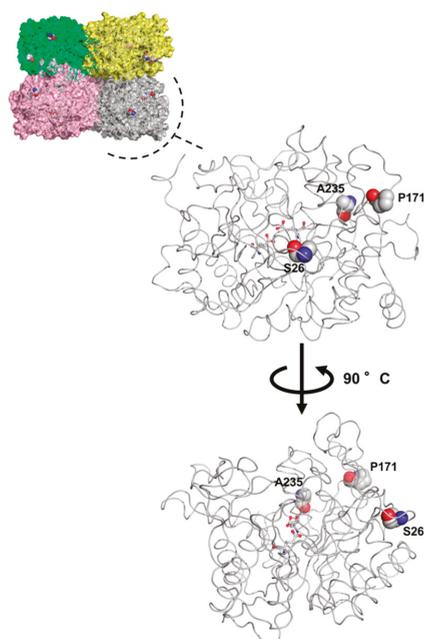
Figure 1. *The aminoacid residues mutated in the Saccharolobus solfataricus (Ssβ-Gly) H7 enzyme.* The position of S26, P171, and A235 residues (in spherical format) are shown on the tetrameric and monomeric 3D structures of *Ss*β-Gly, represented in ribbon format. Catalytic residues in the active site are highlighted in ball and stick format, and all atoms are colored by the CPK convention.

2.2. Steady State Kinetic Constants of the H7 and Relative Single and Double Mutants

The kinetic parameters at 65 °C reported in Table 1 show an increase of activity (k_{cat}) by about 2-fold toward the galactose-derived substrates for the H7 mutant. With 4Np-Gal, the K_M values are moderately affected with respect to the wild type; consequently, a relative increase in the catalytic efficiency (k_{cat}/K_M) is also observed. The turnover number values of the mutant on 4Np-Glc and cellobiose were similar to the wild type. However, the k_{cat}/K_M on 4Np-Glc substrate was 3.5-fold higher than the wild type. By contrast, the increased catalytic efficiency observed on 4Np-Glc is mainly attributable to the K_M, which is 4-fold lower than that of the wild type. The H7 mutant shows an almost 2-fold decreased affinity toward lactose; consequently, k_{cat}/K_M values are unchanged. Therefore, H7 offered to *E. coli* M5154 strain the ability of growing faster on lactose thanks to an enhanced turnover number rather than catalytic efficiency. Similar results have been obtained with an evolved indoleglycerol phosphate synthase (IGPS), also from *Sulfolobus solfataricus*, yielding mutants that complemented *E. coli* growth thanks to an improved turnover number [10].

As shown in Table 1, only the S26L mutant showed properties comparable to the H7 triple mutant, with both k_{cat} and k_{cat}/K_M values for 4Np-Gal being 2-fold higher than the wild type and specificity constants for 4Np-Glc, and cellobiose improved by 2.5-fold and 1.6-fold, respectively. The other single mutants showed catalytic constant values on all substrates tested similar to or lower than the wild type. Among the single mutants, S26L showed the highest catalytic efficiency on disaccharides, although the k_{cat}/K_M value on lactose is comparable to that of the wild type.

Table 1. Steady state kinetic constants at 65 °C of *Saccharolobus solfataricus* (*Ss*β-gly) wild type and mutants.

	4Np-Gal			4Np-Glc		
	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
Wild type (S26/P171/A235)	1.17 ± 0.15	389.8 ± 13.25	333.16	0.94 ± 0.22	437.2 ± 23.4	465.1
H7 (S26L/P171L/A235V)	1.71 ± 0.32	908.7 ± 43.3	530.0	0.23 ± 0.03	376.8 ± 13.9	1629.0
S26L	1.56 ± 0.23	938.9 ± 33.9	601.8	0.36 ± 0.09	423.8 ± 20.1	1177.2
P171L	1.39 ± 0.25	444.5 ± 18.8	319.8	0.44 ± 0.11	221.6 ± 10.8	503.6
A235V	1.00 ± 0.11	310.1 ± 7.9	310.0	0.41 ± 0.08	168.5 ± 6.1	414.9
S26L/P171L	2.85 ± 0.23	93.2 ± 2.2	32.8	0.31 ± 0.10	34.3 ± 2.0	109.0
S26L/A235V	1.20 ± 0.18	795.7 ± 27.2	662.0	0.44 ± 0.11	393.1 ± 19.2	885.7
P171L/V235V	1.46 ± 0.13	150.5 ± 3.2	102.9	0.47 ± 0.11	75.7 ± 3.6	159.0
	Lactose			Cellobiose		
	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
Wild type (S26/P171/A235)	137.7 ± 6.2	710.8 ± 9.6	5.2	33.2 ± 5.12	274.7 ± 12.9	8.3
H7 (S26L/P171L/A235V)	245.5 ± 20.4	1270.8 ± 40.7	5.2	33.6 ± 6.7	372.5 ± 22.1	11.1
S26L	118.2 ± 5.1	505.0 ± 7.2	4.3	18.3 ± 3.2	252.2 ± 11.0	13.8
P171L	140.4 ± 20.4	308.6 ± 15.7	2.2	20.6 ± 2.7	122.2 ± 4.0	5.9
A235V	149.6 ± 10,3	226.6 ± 5,5	1.5	18.3 ± 1.0	95.1 ± 1.4	5.2
S26L/P171L	144.8 ± 7.6	84.2 ± 1.5	0.6	30.7 ± 4.2	36.5 ± 1.4	1.2
S26L/A235V	120.0 ± 12.6	510.2 ± 17.6	4.3	80.4 ± 21.4	374.8 ± 36.5	4.7
P171L/V235V	161.9 ± 8.7	119.0 ± 2.3	0.7	55.6 ± 10.0	68.0 ± 4.4	1.2

The β-glycosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* (CelB), also belonging to family GH1, once evolved on 4Np-Glc at room temperature, showed a three-fold improved rates of hydrolysis by single mutations in the active site [11]. In the *Ss*β-gly H7 mutant, instead, S26L, P171L, and A235V mutations were all far from the active site. The superimposition of the 3D-structure of *Ss*β-gly and the derived model of the triple mutant did not show any difference in the residues of the active site (not shown). The distance from the active site did not allow us to find easy explanations on the molecular reasons of the ability of H7 to promote lactose hydrolysis faster than the wild type. Potentially, increasing the flexibility of the structure in the mutation sites led to conformational changes that were transferred to the active site and improved the conversion of galactose-derived substrates. Mutations influencing enzyme function even at a distance from the catalytic residues may be not uncommon, as proteins often undergo conformational changes during the reaction [47] and long distance effects play important roles in enzyme function, as recently reviewed in the work of [48]. Rational protein engineering experiments on the amino acids located in the active site region broadened the substrate specificity of *Ss*β-gly [32,33], suggesting that active site modifications would be needed to change the substrate specificity of the enzyme. Therefore, it would be easy to predict that further mutations for improved activity could be found in the same area. Instead, we showed here that substitutions not restricted to the active site had also some influence, allowing in vivo complementation and significantly enhancing the turnover number and the catalytic efficiency of the enzyme on galactose-derived substrates and 4-Np-Glc, respectively. The different effect observed for *Ss*β-gly and CelB could be the result of the different approaches used: in vivo selection instead of in vitro screening, respectively. In the former approach, biology exerts a combination of forces not determined only by kinetic and chemical recognition. Our study demonstrates that, at the in vivo selection conditions used, enzyme catalysis of *Ss*β-gly was determined by the enzyme as a whole.

2.3. Temperature Dependence of the H7 and Relative Single and Double Mutants

The effect of temperature on the specific activity on 4Np-Gal by all the mutants followed the same trend of the wild type, increasing continuously up to the maximum of 85 °C, with S26L and the H7 mutants showing the highest U mg^{-1} at all temperatures (Figure 2a,b). The specific activity of the S26L mutant, as single or in combination with A235V, was maximal at 85 °C, as observed in the triple mutant. Instead, at this temperature, the specific activity values were similar or lower than

those of the wild type, for the single mutants A235V and P171L and the double mutants P171L/A235V and P26L/P171L, respectively (Figure 2a,b). This is someway surprising. For the known rule 'you get what you screen for [49], and based on previous studies on CelB [11], it might be expected that the selection at 37 °C could result in increasing specific activity at mesophilic temperatures, possibly reducing the activity at 85 °C. Instead, selective pressure increased the specific activity of the mutant at temperatures >50 °C, where the upward shift of the curve was not constant, but increased with temperature (Figure 2a). This resembles the behaviour of the mutants of a thermophilic esterase after several directed evolution generations [50].
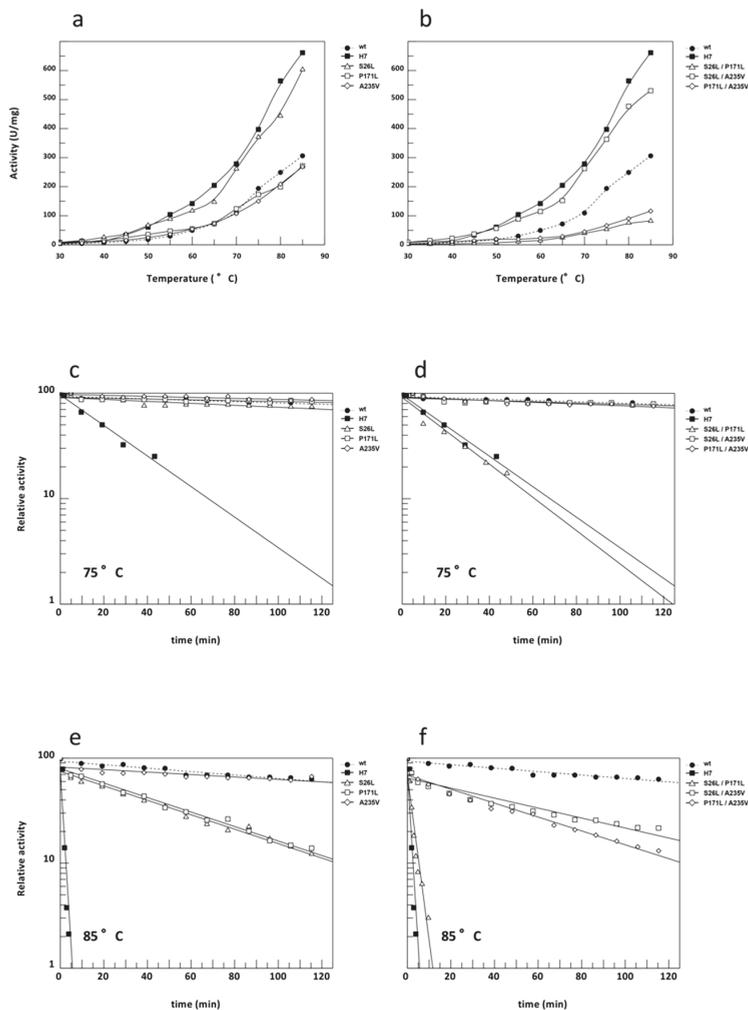
Figure 2. Thermal activity and stability of *Ss*β-Gly H7 and relative single and double mutants. The dependence on temperature of the wild type is compared with the H7 mutant; single (**a**) and double mutants (**b**). Similarly, the residual activity at 75 °C and 85 °C of the wild type, in comparison with H7 mutant, single (**c,e**) and double mutants (**d,f**), is reported. *Ss*β-Gly (closed circles on dotted line); *Ss*β-Gly H7 (closed squares); L26 (open triangles); L171 (open squares); V235 (open diamonds); L26/L171 (open triangles); L26/V235 (open squares); L171/V235 (open diamonds).

The effect of the mutations on *Ss*β-gly was detrimental on the thermal stability. The H7 mutant was as stable as the wild type at T ≤ 65 °C (data not shown), but 23-fold and 283-fold less stable at 75 °C and 85 °C, respectively (Figure 2c–f, Table 2). The half-lives measured at 75 °C showed that all single mutations did not significantly alter the enzyme stability with the exception of S26L, which was inactivated 1.5-fold faster than the wild type (Table 2). This is even clearer in the S26L/P171L double mutant, showing the same low stability of H7 (Figure 2c,d). Instead, the same analysis at 85 °C revealed that only the A235V mutant was as stable as the wild type (Figure 2e). All the other mutants showed a $t_{\frac{1}{2}}$ from 5.5-fold to 284-fold lower than the wild type *Ss*β-gly under these conditions (Table 2). Again, the S26L/P171L double mutant is similar to H7 (Figure 2f).

Table 2. Half-life of *Ss*β-gly wild type and mutants at 75 °C and 85 °C.

Mutants	$t_{\frac{1}{2}}$ (min)		Symbols Used in Figure 2
	75 °C	**85 °C**	
Wild type (S26/P171/A235)	461.3 ± 78.0	170.4 ± 10.7	···●···
H7 (S26L/P171L/A235V)	20.5 ± 1.5	0.6 ± 0.1	▬■▬
S26L	299.4 ± 68.2	26.6 ± 1.2	─△─
P171L	681.1 ± 175.7	30.7 ± 1.4	─□─
A235V	580.8 ± 10.5	200.9 ± 37.1	─◇─
S26L/P171L	17.4 ± 2.4	0.9 ± 0.1	─△─
S26L/A235V	465.7 ± 164.2	24.4 ± 2.7	─□─
P171L/A235V	335.8 ± 42.3	21.8 ± 1.4	─◇─

The inspection of the H7 mutant 3D-model (Figure 3) allowed the explanation of these results. The mutations increased the entire hydrophobicity of the protein and removed specific stabilizing interactions. Residues with low hydrophobicity index (Ser = −0.8; Pro = −1.8; Ala = 1.8) were substituted with amino acids that are highly hydrophobic (Leu = 3.8; Val = 4.2) [51]. The substitution of Ser26 with a leucine increased the hydrophobicity in the loop region encompassing Met21 and Gly22 (Figure 3a,b) and removed a hydrogen-bond with the peptide nitrogen of Gly22 and the carbonyl oxygen of Thr23. In addition, the replacement of the polar S26 residue exposed to the solvent with a non-polar leucine led to a local destabilization of the protein structure, observable at temperatures >80 °C. However, interestingly, this increased flexibility explains the high thermophilicity of the S26L mutant (Figure 2a).
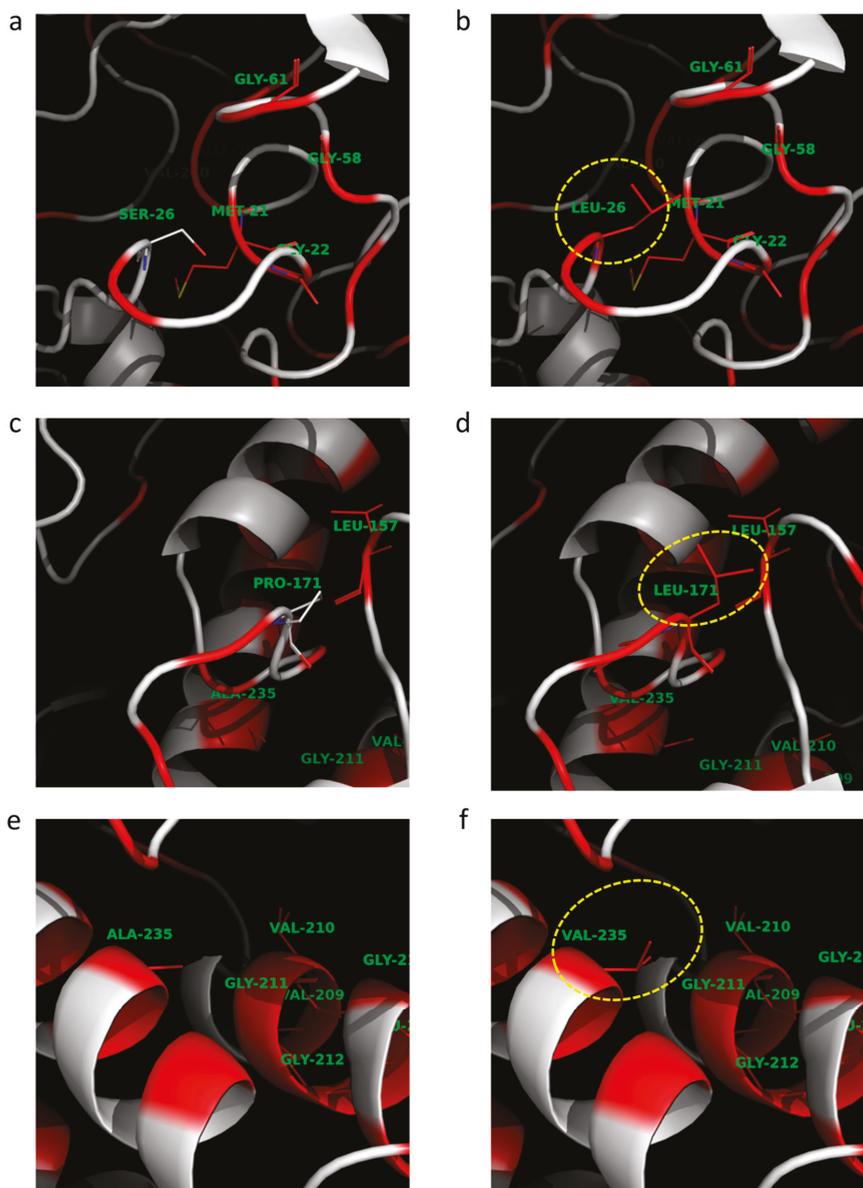
Figure 3. *Ss*β-Gly 3D structure in comparison with the H7 mutant model. Close-up of Ssβ-Gly wild type structure (PDB ID: 1GOW) (**a**,**c**,**e**) and *Ss*β-Gly single mutations in the H7 mutant, S26L, P171L, and A235V in (**b**,**d**,**f**), respectively. The hydrophobic residues are indicated in red. The mutant residues are highlighted in yellow.

The substitution of the Pro171 with a leucine residue (P171L) in the extended loop Arg165-Thr177 exposed to the solvent (Figure 3c,d) explained the destabilizing effect in the nearby hydrophobic environment involving the α4-helix (Ser228-Lys255), the small α-helix (Pro207-Val206), and the loop Asn147-Pro160 (Figure 3e,f). When we modelled in the structure the substituting leucine, the proline cis configuration was removed (Figure 3d) and the local flexibility increased, thus destabilizing the

structure at 85 °C (Figure 2e). Interestingly, the destabilizing effect of the S26L and P171L mutations is additive, indicating that they are responsible for the instability of the triple mutant H7 (Figure 2f).

According to the *Ss*β-gly 3D structure, A235 is located inside of the α4-helix and is very close to the small α-helix Pro207-Val216 [23]. The lack of destabilization by the A235V mutation could be explained by the similarity between these two residues, which does not affect the formation of the α-helix despite the greater steric hindrance of a valine residue. On the other hand, V235, showing a longer side chain when compared with the wild type alanine, might affect the optimal packing by protruding on the side of the helix. However, the increased local hydrophobicity could, in part, balance this negative effect, explaining why A235V showed the same thermostability at 85 °C as the wild type (Figure 3e).

The instability of the A235V-containing double mutants is clearly attributable to the detrimental contribution of the other mutations (Table 2; Figure 2d,f). The presence of the three mutations on the surface of *Ss*β-gly, and the fact that they do not seem to interact with each other, may indicate that the mutations act independently on different regions of the protein promoting the unfolding at T > 75 °C, thereby affecting the overall stability. In the case of P171L mutation, possibly, an increased flexibility may have made the H7 triple mutant more active than the wild type at all temperatures tested.

It is worth noting that the triple mutant is more active than its wild-type progenitor at all temperatures, but rapidly became inactivated at 85 °C. Directed evolution studies in which enzymes were adapted to different temperatures often led to the conclusion that there is a trade-off between catalytic activity at low temperatures and thermostability, explained as the balance to maintain the rigidity necessary for stability and the flexibility required for enzyme activity. However, this is not always the case; indeed, the evolved CelB with optimized substrate specificity and gained catalytic activity at 20 °C showed reduced activity at T > 80 °C and destabilization [11], and an evolved glucokinase from a thermophilic fungus showed both improved stability under heat and thermophilicity when compared with the wild type [52]. Laboratory evolution experiments can identify different adaptive mechanisms, but these can be achieved by exploring a limited number of protein fitness pathways often dependent on the selection method used [49,53]. In this study, we applied an in vivo selection, in which mutants complementing an *E. coli* mutant were isolated. The same approach on IGPS from *S. solfataricus* yielded mutants that improved growth appreciably, with activity increasing with temperature and accelerated thermal denaturation as observed here for evolved *Ss*β-gly [10]. Selecting enzyme properties through survival and growth of the host organism, unlike in vitro screenings, exerts a biological function that could be dictated not only by physical and chemical laws, but also by biological natural selection [53]. Possibly, the reduced stability observed for evolved *Ss*β-gly may allow its quicker protein turnover within the *E. coli* mesophilic cell that could be required to complement the lacking β-galactosidase activity. Conversely, the improved activity of the H7 mutant at temperatures where *S. solfataricus* grows optimally (85 °C) could merely be the consequence of *E. coli* being more active at 37 °C and might not fit in archaeal cells by interfering with the normal physiology of this organism.

3. Materials and Methods

3.1. Reagents

All commercially available substrates were purchased from Sigma. Restriction endonucleases and T4 DNA ligase were from New England BioLabs, USA. Synthetic oligonucleotides were from PRIMM (Milan, Italy). *Escherichia coli* M5154 strain was from the Coli Genetic Stock Center (Yale University, USA).

3.2. Random Mutagenesis

The pGEX-2TK-derivative (GE Healthcare, USA) plasmid (pGEX-K-Gly) containing the wild type *S. solfataricus* β-glycosidase gene (*lacS*) was described previously [54]. Random mutagenesis was performed by chemical treatment of the pGEX-K-Gly plasmid with hydroxylamine. In particular,

~15.0 μg of the plasmid was denatured with NaOH 0.2 M at room temperature for 5 min, followed by a fast neutralization step with ammonium acetate at pH 7.5. Finally, the plasmid was exposed to hydroxylamine (H_3NO) 0.5 M at 37 °C for 8.0 h in phosphate buffer 50.0 mM pH 6.5. The ~1500 bp DNA fragment relative to the *lacS* gene was purified by gel extraction, after digestion of the treated pGEX-K-Gly plasmid using the Bam HI and Sma I restriction enzymes. A library of random mutants was obtained by a ligation of the *lacS* gene with a Bam HI-Sma I linearized pGEX-2TK plasmid.

3.3. Selection of Ssβ-gly Mutants for Enhanced Lactose Hydrolysis

A ligation mixture was utilized for the transformation of the *E. coli* M5154 strain, which harbours a deletion of almost one-third of the *lacZ* gene (ΔlacZ39 deletion), making this strain completely inactive for the hydrolysis of galactose-derived β-substrates. For the screening, we exploited the β-glucosidase activity of the product of the *lacS* gene. *E. coli* M5154 cells transformed with the library of random mutants were first plated on 50 μg/mL ampicillin (Amp) LB Agar medium containing the 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside (X-Glc) chromogenic substrate, and the mutagenic impact of the hydroxylamine treatment was thus determined by calculating the blue/white colony ratio. Afterwards, transformed cells were plated on M9 Amp Trp minimal medium, containing lactose as the sole carbon source, and incubated at 37 °C. Wild type *lacS* expressing clones led to colonies after about 10 days of incubation because of the basal β-galactosidase activity of the enzyme. All colonies appearing before, namely after six to seven days of incubation, were identified and picked. Selected colonies were grown in rich medium and analyzed for their β-galacto- and β-glucosidase activity, by using a colorimetric enzymatic assay and 2Np-Gal and 2Np-Glc, respectively, as substrates [55].

3.4. Production of H7-Derived Single and Double Mutants

The in vivo selected H7 clone was sequenced and three mutations were identified (S26L, P171L, and A235V). The opportunity that unique restriction sites in the *lacS* gene separate these mutations allowed for the preparation of specific DNA fragments containing single and/or double mutations. These were combined and ligated with the opportunely digested pGEX-K-Gly plasmids to finally achieve the Ssβ-gly single and double mutants.

3.5. Enzyme Expression and Purification

Wild-type Ssβ-gly and relative mutants were expressed and purified as fusions of glutathione S-transferase (GST), as previously reported [36,54,56]. Briefly, transformed *E. coli* BL21(DE3)RIL, carrying extra copies of the *argU*, *ileY*, and *leuW* tRNA genes cells, were grown in SB medium at 37 °C and induced by the addition of 1.0 mM IPTG, when cultures reached an OD_{600} of 1.0. After 16 h of incubation, cells were harvested by centrifugation at 5000× *g* and stored at −20 °C. Then, the pellet was thawed and resuspended in 2.0 mL/g cells of PBS buffer (50.0 mM sodium phosphate buffer, 150.0 mM NaCl; pH 7.4). After French press cell disruption and centrifugation at 4 °C for 30 min at 30,000× g to discard cell debris, free cell crude extract containing the fusion protein was subjected to an affinity chromatography on a glutathione-Sepharose 4B™ resin (GE Healthcare, USA), followed by an incubation with thrombin protease in order to separate GST and Ssβ-gly.

Aliquots of pure proteins (>95.0%, as judged by SDS-PAGE analysis) were stored at −20 °C in PSB 1X containing 20.0% glycerol. The samples stored in these conditions are stable for several months. The protein concentration was determined with the method of Bradford, using bovine serum albumin as the standard [57].

3.6. Enzyme Characterization

All kinetic studies were performed by following spectrophotometrically the hydrolytic activity with a Cary 100 Scan spectrophotometer (Varian, Australia), coupled with a thermally controlled Peltier system. A total of 1.0 mL of reaction solution was preheated for 2.0 min, keeping the temperature constant during all measurements. β-glycosidase activity performed at 65 °C in 50 mM sodium

phosphate buffer at pH 6.5 was conventionally defined as the standard reaction. Kinetic parameters at standard conditions were determined using artificial and natural substrates, whose concentration ranged from 0.1 to 30.0 and from 1.25 to 750.0 mM, respectively. In order to correct the spontaneous hydrolysis of the substrates, mixtures containing all the reactants except enzymes were prepared and referred to as blank reactions. Steady-state kinetic constants were obtained by plotting initial velocity values versus 15 different substrate concentrations for each substrate. In each assay, amounts ranging from 0.5 to 20.0 μg of wild type *Ss*β-gly and mutants were used. Measuring the hydrolytic activity on 4-nitrophenol based substrates at standard conditions, a molar extinction coefficient (ε_M) value of 9340 M^{-1} cm^{-1} at 405 nm was used, whereas the activity on lactose and cellobiose was determined by measuring the released glucose with the glucose oxidase–peroxidase enzymatic assay GOPOD (Megazyme, Chicago, Illinois, USA), taking into account that one molecule of cellobiose leads to the release of two glucose units. One unit of enzyme activity was conventionally defined as the amount of the enzyme that hydrolyses 1.0 μmol of substrate in 1.0 min under standard conditions. All data were plotted and refined using the program GraFit 5.0 (Erithacus Software Ltd., Horley, U.K.), in order to determine the steady-state kinetic parameters.

3.7. Thermal Stability and Temperature Dependence

The thermal stability was analyzed by incubating all the enzymes at the concentration of 0.01 μg/mL in 50.0 mM phosphate buffer pH 6.5, at the indicated temperatures and times. The hydrolytic activity of each sample was determined at 405 nm in standard conditions, using a molar extinction coefficient (ε_M) value of 9340 M^{-1} cm^{-1} for the release of 4-nitrophenol. The activity of the sample measured without any incubation was taken as 100%.

Enzymes were assayed in standard conditions from 30 to 85 °C, in order to evaluate their temperature dependence. Enzymatic units were measured using different molar extinction coefficient values, as follows: 6130 M^{-1} cm^{-1} (30 °C); 6570 M^{-1} cm^{-1} (35 °C); 6910 M^{-1} cm^{-1} (40 °C); 7530 M^{-1} cm^{-1} (45 °C); 7970 M^{-1} cm^{-1} (50 °C); 8420 M^{-1} cm^{-1} (55 °C); 8890 M^{-1} cm^{-1} (60 °C); 9340 M^{-1} cm^{-1} (65 °C); 9700 M^{-1} cm^{-1} (70 °C); 10120 M^{-1} cm^{-1} (75 °C); 10610 M^{-1} cm^{-1} (80 °C); 10900 M^{-1} cm^{-1} (85 °C).

3.8. 3D-Model of Ssβ-gly H7 Mutant

Homology modeling of *Ss*β-gly H7 mutant structure was performed by SWISS-MODEL server [58] using the *Ss*β-gly wild type structure (PDB entry 1gow) as a template. Stereochemical quality of the model was analyzed by the PROCHECK program and PyMol 1.0 was used to analyze and visualize the structure [59].

4. Conclusions

In this study, we identified residues that are critical in determining the thermostability, thermal activity, and substrate specificity of the studied thermozyme *Ss*β-gly. In previous rationally designed protein engineering studies on *Ss*β-gly, mutants in the active site showed broadened substrate specificity. Here, three mutations far from the active site obtained randomly after an in vivo selection produced a more than two-fold enhancement of the specific activity at 85 °C and of the k_{cat} on galactoside substrates. Interestingly, the activity of the enzyme on cellobiose remained the same as the wild type or was even enhanced on 4-Np-Glc. This remarkable result was paid off with a 283-fold reduced thermal stability, when compared with the wild type. The change of Ser26 into Leu enhanced the thermophilicity because of the disruption of stabilizing interactions that, by increasing the flexibility of the enzyme, facilitated catalysis at a high temperature, but also weakened its resistance to heat. Our study shows that it is not always successful to improve the enzyme activity following a "rational" approach based on the known and understood physicochemical laws that determine the conformations and interactions of the amino acids and enzyme catalysis. When possible, "irrational" approaches, which exploit in vivo selection, can produce enzyme mutants of interest.

References

1. Fields, P.A. Review: Protein function at thermal extremes: Balancing stability and flexibility. *Comp. Biochem. Physiol. A: Mol. Integr. Physiol.* **2001**, *129*, 417–431. [CrossRef]
2. Littlechild, J.A.; Guy, J.; Connelly, S.; Mallett, L.; Waddell, S.; Rye, C.A.; Line, K.; Isupov, M. Natural methods of protein stabilization: Thermostable biocatalysts. *Biochem. Soc. Trans.* **2007**, *35*, 1558–1563. [CrossRef]
3. Imanaka, T. Molecular bases of thermophily in hyperthermophiles. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **2011**, *87*, 587–602. [CrossRef] [PubMed]
4. Pica, A.; Graziano, G. Shedding light on the extra thermal stability of thermophilic proteins. *Biopolymers* **2016**, *105*, 856–863. [CrossRef]
5. Vieille, C.; Zeikus, G.J. Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1–43. [CrossRef]
6. Ito, Y.; Ikeuchi, A.; Imamura, C. Advanced evolutionary molecular engineering to produce thermostable cellulase by using a small but efficient library. *Protein Eng. Des. Sel.* **2013**, *26*, 73–79. [CrossRef]
7. Ruller, R.; Alponti, J.; Deliberto, L.A.; Zanphorlin, L.M.; Machado, C.B.; Ward, R.J. Concommitant adaptation of a gh11 xylanase by directed evolution to create an alkali-tolerant/thermophilic enzyme. *Protein Eng. Des. Sel.* **2014**, *27*, 255–262. [CrossRef]
8. Morais, S.; Stern, J.; Kahn, A.; Galanopoulou, A.P.; Yoav, S.; Shamshoum, M.; Smith, M.A.; Hatzinikolaou, D.G.; Arnold, F.H.; Bayer, E.A. Enhancement of cellulosome-mediated deconstruction of cellulose by improving enzyme thermostability. *Biotechnol. Biofuels* **2016**, *9*, 164. [CrossRef]
9. Ribeiro, A.L.; Sanchez, M.; Hidalgo, A.; Berenguer, J. Stabilization of enzymes by using thermophiles. *Methods Mol. Biol.* **2017**, *1645*, 297–312.
10. Merz, A.; Yee, M.C.; Szadkowski, H.; Pappenberger, G.; Crameri, A.; Stemmer, W.P.; Yanofsky, C.; Kirschner, K. Improving the catalytic activity of a thermophilic enzyme at low temperatures. *Biochemistry* **2000**, *39*, 880–889. [CrossRef] [PubMed]
11. Lebbink, J.H.; Kaper, T.; Bron, P.; van der Oost, J.; de Vos, W.M. Improving low-temperature catalysis in the hyperthermostable pyrococcus furiosus beta-glucosidase celb by directed evolution. *Biochemistry* **2000**, *39*, 3656–3665. [CrossRef]
12. Sunna, A.; Moracci, M.; Rossi, M.; Antranikian, G. Glycosyl hydrolases form hyperthermophiles. *Extremophiles* **1997**, *1*, 2–13. [CrossRef]
13. Elleuche, S.; Schroder, C.; Sahm, K.; Antranikian, G. Extremozymes-biocatalysts with unique properties from extremophilic microorganisms. *Curr. Opin. Biotechnol.* **2014**, *29*, 116–123. [CrossRef]
14. Blumer-Schuette, S.E.; Brown, S.D.; Sander, K.B.; Bayer, E.A.; Kataeva, I.; Zurawski, J.V.; Conway, J.M.; Adams, M.W.W.; Kelly, R.M. Thermophilic lignocellulose deconstruction. *FEMS Microbiol. Rev.* **2014**, *38*, 393–448. [CrossRef] [PubMed]
15. Botha, J.; Mizrachi, E.; Myburg, A.A.; Cowan, D.A. Carbohydrate active enzyme domains from extreme thermophiles: Components of a modular toolbox for lignocellulose degradation. *Extremophiles* **2018**, *22*, 1–12. [CrossRef] [PubMed]

16. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (cazy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490–D495. [CrossRef] [PubMed]

17. Pisani, F.M.; Rella, R.; Raia, C.A.; Rozzo, C.; Nucci, R.; Gambacorta, A.; De Rosa, M.; Rossi, M. Thermostable beta-galactosidase from the archaebacterium sulfolobus solfataricus. Purification and properties. *Eur. J. Biochem.* **1990**, *187*, 321–328. [CrossRef] [PubMed]

18. Moracci, M.; Lavolpe, A.; Pulitzer, J.F.; Rossi, M.; Ciaramella, M. Expression of the thermostable beta-galactosidase gene from the archaebacterium sulfolobus-solfataricus in saccharomyces-cerevisiae and characterization of a new inducible promoter for heterologous expression. *J. Bacteriol.* **1992**, *174*, 873–882. [CrossRef] [PubMed]

19. Moracci, M.; Ciaramella, M.; Nucci, R.; Pearl, L.H.; Sanderson, I.; Trincone, A.; Rossi, M. Thermostableβ-glycosidase from sulfolobus-solfataricus. *Biocatalysis* **1994**, *11*, 89–103. [CrossRef]

20. Moracci, M.; Nucci, R.; Febbraio, F.; Vaccaro, C.; Vespa, N.; Lacara, F.; Rossi, M. Expression and extensive characterization of a beta-glycosidase from the extreme thermoacidophilic archaeon sulfolobus-solfataricus in escherichia-coli - authenticity of the recombinant enzyme. *Enzyme Microb. Tech.* **1995**, *17*, 992–997. [CrossRef]

21. Pouwels, J.; Moracci, M.; Cobucci-Ponzano, B.; Perugino, G.; van der Oost, J.; Kaper, T.; Lebbink, J.H.G.; de Vos, W.M.; Ciaramella, M.; Rossi, M. Activity and stability of hyperthermophilic enzymes: A comparative study on two archaeal beta-glycosidases. *Extremophiles* **2000**, *4*, 157–164. [CrossRef]

22. Leon, M.; Isorna, P.; Menendez, M.; Sanz-Aparicio, J.; Polaina, J. Comparative study and mutational analysis of distinctive structural elements of hyperthermophilic enzymes. *Protein J.* **2007**, *26*, 435–444. [CrossRef] [PubMed]

23. Aguilar, C.F.; Sanderson, I.; Moracci, M.; Ciaramella, M.; Nucci, R.; Rossi, M.; Pearl, L.H. Crystal structure of the beta-glycosidase from the hyperthermophilic archeon sulfolobus solfataricus: Resilience as a key factor in thermostability. *J. Mol. Biol.* **1997**, *271*, 789–802. [CrossRef]

24. Cobucci-Ponzano, B.; Moracci, M.; Di Lauro, B.; Ciaramella, M.; D'Avino, R.; Rossi, M. Ionic network at the c-terminus of the beta-glycosidase from the hyperthermophilic archaeon sulfolobus solfataricus: Functional role in the quaternary structure thermal stabilization. *Proteins* **2002**, *48*, 98–106. [CrossRef]

25. Ausili, A.; Di Lauro, B.; Cobucci-Ponzano, B.; Bertoli, E.; Scire, A.; Rossi, M.; Tanfani, F.; Moracci, M. Two-dimensional ir correlation spectroscopy of mutants of the beta-glycosidase from the hyperthermophilic archaeon sultolobus soltataricus identifies the mechanism of quaternary structure stabilization and unravels the sequence of thermal unfolding events. *Biochem. J.* **2004**, *384*, 69–78. [CrossRef]

26. Tehei, M.; Franzetti, B.; Madern, D.; Ginzburg, M.; Ginzburg, B.Z.; Giudici-Orticoni, M.T.; Bruschi, M.; Zaccai, G. Adaptation to extreme environments: Macromolecular dynamics in bacteria compared in vivo by neutron scattering. *Embo. Rep.* **2004**, *5*, 66–70. [CrossRef] [PubMed]

27. Kalimeri, M.; Rahaman, O.; Melchionna, S.; Sterpone, F. How conformational flexibility stabilizes the hyperthermophilic elongation factor g-domain. *J. Phys. Chem. B* **2013**, *117*, 13775–13785. [CrossRef] [PubMed]

28. Matsuura, Y.; Takehira, M.; Joti, Y.; Ogasahara, K.; Tanaka, T.; Ono, N.; Kunishima, N.; Yutani, K. Thermodynamics of protein denaturation at temperatures over 100 degrees c: Cuta1 mutant proteins substituted with hydrophobic and charged residues. *Sci. Rep.* **2015**, *5*, 15545. [CrossRef]

29. Deckert, K.; Budiardjo, S.J.; Brunner, L.C.; Lovell, S.; Karanicolas, J. Designing allosteric control into enzymes by chemical rescue of structure. *J. Am. Chem. Soc.* **2012**, *134*, 10055–10060. [CrossRef]

30. Budiardjo, S.J.; Licknack, T.J.; Cory, M.B.; Kapros, D.; Roy, A.; Lovell, S.; Douglas, J.; Karanicolas, J. Full and partial agonism of a designed enzyme switch. *ACS Synth. Biol.* **2016**, *5*, 1475–1484. [CrossRef]

31. Nucci, R.; Moracci, M.; Vaccaro, C.; Vespa, N.; Rossi, M. Exo-glucosidase activity and substrate-specificity of the beta-glycosidase isolated from the extreme thermophile sulfolobus-solfataricus. *Biotechnol. Appl. Bioc.* **1993**, *17*, 239–250.

32. Corbett, K.; Fordham-Skelton, A.P.; Gatehouse, J.A.; Davis, B.G. Tailoring the substrate specificity of the beta-glycosidase from the thermophilic archaeon sulfolobus solfataricus. *Febs. Lett.* **2001**, *509*, 355–360. [CrossRef]

33. Hancock, S.M.; Corbett, K.; Fordham-Skelton, A.P.; Gatehouse, J.A.; Davis, B.G. Developing promiscuous glycosidases for glycoside synthesis: Residues w433 and e432 in sulfolobus solfataricus beta-glycosidase are important glucoside- and galactoside-specificity determinants. *Chembiochem* **2005**, *6*, 866–875. [CrossRef] [PubMed]

34. Iglesias-Fernandez, J.; Hancock, S.M.; Lee, S.S.; Khan, M.; Kirkpatrick, J.; Oldham, N.J.; McAuley, K.; Fordham-Skelton, A.; Rovira, C.; Davis, B.G. A front-face 'sni synthase' engineered from a retaining 'double-sn2' hydrolase. *Nat. Chem. Biol.* **2017**, *13*, 874–881. [CrossRef]

35. Mackenzie, L.F.; Wang, Q.P.; Warren, R.A.J.; Withers, S.G. Glycosynthases: Mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.* **1998**, *120*, 5583–5584. [CrossRef]

36. Moracci, M.; Trincone, A.; Perugino, G.; Ciaramella, M.; Rossi, M. Restoration of the activity of active-site mutants of the hyperthermophilic beta-glycosidase from sulfolobus solfataricus: Dependence of the mechanism on the action of external nucleophiles. *Biochemistry* **1998**, *37*, 17262–17270. [CrossRef] [PubMed]

37. Trincone, A.; Perugino, G.; Rossi, M.; Moracci, M. A novel thermophilic glycosynthase that effects branching glycosylation. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 365–368. [CrossRef]

38. Moracci, M.; Trincone, A.; Rossi, M. Glycosynthases: New enzymes for oligosaccharide synthesis. *J. Mol. Catal. B-Enzym* **2001**, *11*, 155–163. [CrossRef]

39. Cobucci-Ponzano, B.; Strazzulli, A.; Rossi, M.; Moracci, M. Glycosynthases in biocatalysis. *Adv. Synth. Catal.* **2011**, *353*, 2284–2300. [CrossRef]

40. Cobucci-Ponzano, B.; Moracci, M. Glycosynthases as tools for the production of glycan analogs of natural products. *Nat. Prod. Rep.* **2012**, *29*, 697–709. [CrossRef] [PubMed]

41. Wei, J.H.; Lv, X.; Lu, Y.; Yang, G.Z.; Fu, L.F.; Yang, L.; Wang, J.J.; Gao, J.H.; Cheng, S.H.; Duan, Q.; et al. Glycosynthase with broad substrate specificity-an efficient biocatalyst for the construction of oligosaccharide library. *Eur. J. Org. Chem.* **2013**, 2414–2419. [CrossRef]

42. Planas, A. Glycosynthase technology for enzymatic synthesis of functional polysaccharides. *Abstr. Pap. Am. Chem. Soc.* **2015**, *249*.

43. Goddard-Borger, E.D.; Tysoe, C.; Withers, S.G. Glycosynthase mediated synthesis of psychosine. *Carbohydr. Res.* **2016**, *435*, 97–99. [CrossRef]

44. Gloster, T.M.; Roberts, S.; Ducros, V.M.A.; Perugino, G.; Rossi, M.; Hoos, R.; Moracci, M.; Vasella, A.; Davies, G.J. Structural studies of the beta-glycosidase from sulfolobus solfataricus in complex with covalently and noncovalently bound inhibitors. *Biochemistry* **2004**, *43*, 6101–6109. [CrossRef]

45. Gloster, T.M.; Roberts, S.; Perugino, G.; Rossi, M.; Moracci, M.; Panday, N.; Terinek, M.; Vasella, A.; Davies, G.J. Structural, kinetic, and thermodynamic analysis of glucoimidazole-derived glycosidase inhibitors. *Biochemistry* **2006**, *45*, 11879–11884. [CrossRef]

46. Hill, C.M.; Duggleby, R.G. Mutagenesis of escherichia coli acetohydroxyacid synthase isoenzyme ii and characterization of three herbicide-insensitive forms. *Biochem. J.* **1998**, *335*, 653–661. [CrossRef]

47. Shimotohno, A.; Oue, S.; Yano, T.; Kuramitsu, S.; Kagamiyama, R. Demonstration of the importance and usefulness of manipulating non-active-site residues in protein design. *J. Biochem.* **2001**, *129*, 943–948. [CrossRef] [PubMed]

48. Agarwal, P.K. A biophysical perspective on enzyme catalysis. *Biochemistry* **2019**, *58*, 438–449. [CrossRef]

49. Romero, P.A.; Arnold, F.H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866–876. [CrossRef]

50. Giver, L.; Gershenson, A.; Freskgard, P.O.; Arnold, F.H. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12809–12813. [CrossRef]

51. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132. [CrossRef]

52. Zhou, W.; Huang, R.; Zhu, Z.; Zhang, Y.P.J. Coevolution of both thermostability and activity of polyphosphate glucokinase from thermobifida fusca yx. *Appl. Environ. Microbiol.* **2018**, *84*, e01224-18. [CrossRef]

53. Arnold, F.H.; Wintrode, P.L.; Miyazaki, K.; Gershenson, A. How enzymes adapt: Lessons from directed evolution. *Trends Biochem. Sci* **2001**, *26*, 100–106. [CrossRef]

54. Moracci, M.; Capalbo, L.; Ciaramella, M.; Rossi, M. Identification of two glutamic acid residues essential for catalysis in the beta-glycosidase from the thermoacidophilic archaeon sulfolobus solfataricus. *Protein Eng.* **1996**, *9*, 1191–1195. [CrossRef]

55. Griffith, K.L.; Wolf, R.E., Jr. Measuring beta-galactosidase activity in bacteria: Cell growth, permeabilization, and enzyme assays in 96-well arrays. *Biochem. Biophys. Res. Commun.* **2002**, *290*, 397–402. [CrossRef] [PubMed]

56. Perugino, G.; Trincone, A.; Giordano, A.; van der Oost, J.; Kaper, T.; Rossi, M.; Moracci, M. Activity of hyperthermophilic glycosynthases is significantly enhanced at acidic ph. *Biochemistry* **2003**, *42*, 8484–8493. [CrossRef] [PubMed]
57. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]
58. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The swiss-model workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **2006**, *22*, 195–201. [CrossRef] [PubMed]
59. Laskowski, R.A.; Macarthur, M.W.; Moss, D.S.; Thornton, J.M. Procheck: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291. [CrossRef]

Change of the Product Specificity of a Cyclodextrin Glucanotransferase by Semi-Rational Mutagenesis to Synthesize Large-Ring Cyclodextrins

Christian Sonnendecker and Wolfgang Zimmermann *

Department of Microbiology and Bioprocess Technology, Institute of Biochemistry, Leipzig University, Johannisallee 23, 04103 Leipzig, Germany; christian.sonnendecker@uni-leipzig.de
* Correspondence: wolfgang.zimmermann@uni-leipzig.de; Tel.: +49-341-97-36781

Abstract: Cyclodextrin glucanotransferases (CGTases) convert starch to cyclodextrins (CD) of various sizes. To engineer a CGTase for the synthesis of large-ring CD composed of 9 to 12 glucose units, a loop structure of the protein involved in substrate binding was targeted for semi-rational mutagenesis. Based on multiple protein alignments and protein structure information, a mutagenic megaprimer was designed to encode a partial randomization of eight amino acid residues within the loop region. The library obtained encoding amino acid sequences occurring in wild type CGTases in combination with a screening procedure yielded sequences displaying a changed CD product specificity. As a result, variants of the CGTase from the alkaliphilic *Bacillus* sp. G825-6 synthesizing mainly CD9 to CD12 could be obtained. When the mutagenesis experiment was performed with the CGTase G825-6 variant Y183R, the same loop alterations that increased the total CD synthesis activity resulted in lower activities of the variant enzymes created. In the presence of the amino acid residue R183, the synthesis of CD8 was suppressed and larger CD were obtained as the main products. The alterations not only affected the product specificity, but also influenced the thermal stability of some of the CGTase variants indicating the importance of the loop structure for the stability of the CGTase.

Keywords: cyclodextrin glucanotransferases; large-ring cyclodextrins; semi rational mutagenesis

1. Introduction

Evolutionary methods can be applied to change amino acid residues of proteins to alter their properties as desired [1]. Iterative rounds of diversification and selection thereby mimic an evolutionary process in vitro [2]. Such directed evolution approaches on a molecular level are in in contrast to site-directed mutagenesis techniques, where rational considerations lead to specific and targeted alterations of the target protein [3]. A combination of both approaches can be described as semi-rational protein engineering where the size of the library of the created protein variants is reduced and its fitness is increased [4]. Libraries with higher fitness can be constructed using mutagenic primers [5–7]. Their design requires information obtained from multiple sequence alignments, consensus sequences, and from the structure and function of the protein [5–7]. By selecting distinct amino acid residues instead of total randomization, which would lead to an incongruous number of combinations, a drastic decrease of the size of the variant library can be achieved allowing the alteration of further residues of the target protein without increasing its size [8].

Cyclodextrin glucanotransferases (CGTases), a part of the GH13 α-amylase superfamily, are starch-degrading enzymes produced by many bacteria [9,10]. They catalyze an intramolecular transglycosylation reaction resulting in the formation of cyclic α-1,4-linked glucans, cyclodextrins (CD) [11]. Due to their ability to form reversible complexes with guest molecules thereby increasing their solubility and stability, CD6 to CD8 composed of 6 to 8 glucose units have found many industrial

applications [12]. In contrast, limited information is available on the potential of larger ring CD as host compounds [13–15]. In particular, CD9 to CD12 could be of interest as hosts for the formation of complexes with bulky guest compounds or with pharmaceuticals not able to form complexes with CD6 to CD8 [16–18]. Although a mixture of CD with a degree of polymerization of 6 up to more than 60 is initially synthesized by the CGTase, the larger CD are rapidly reused as substrates in further intermolecular transglycosylation reactions [19]. Therefore, mostly CD6, CD7 and some CD8 with only small amounts of large-ring CD of more than eight glucose units are obtained as products of a reaction of CGTases with starch [20,21]. While efforts to engineer CGTases to increase the yield of CD 6 to CD8 have been published previously [22], only recently attempts to increase the yield of larger CD by site-directed mutagenesis and domain shuffling of CGTases have been reported [23,24].

In this paper, we engineered a CGTase to synthesize large-ring CD in high amounts. We used a semi-rational mutagenesis approach to generate a diverse library of enzyme variants with high fitness. Based on multiple sequence alignments and protein structure information, a loop region of the enzyme was selected for mutagenesis using a partially randomized megaprimer. A library adjusted to reduce sequence space and increase its fitness was designed using degenerated codons, followed by fitness screening of the obtained recombinant *E. coli* clones.

2. Results

2.1. Construction of Vector Libraries

A low-conserved loop region of the CGTase G825-6 (amino acid residues 81–89) representing the binding site for the glucan substrate at the subsite-3 (three glucose residues downstream from the glucan chain cleavage site towards the non-reducing end) was selected as the target for mutagenesis. Based on a multiple sequence alignment of 31 CGTases (Table S1), degenerated codons were used to design a DNA library encoding for a partially randomized loop sequence (Figure S1, Figure 1a). The mutagenic megaprimer with conserved flanking regions was amplified and used in a further PCR with CGTase G825-6 and the variant Y183R as templates. The obtained vector library displayed the desired randomization and was subsequently used for expression and activity screening in *E. coli* BL21(DE3) as host (Figure 1a). The constructs were designated Loop-3 (L3) with the CGTase G825-6 as a template and Loop-3 Y183R (L3YR) with the corresponding CGTase encoding the mutation Y183R as a template.
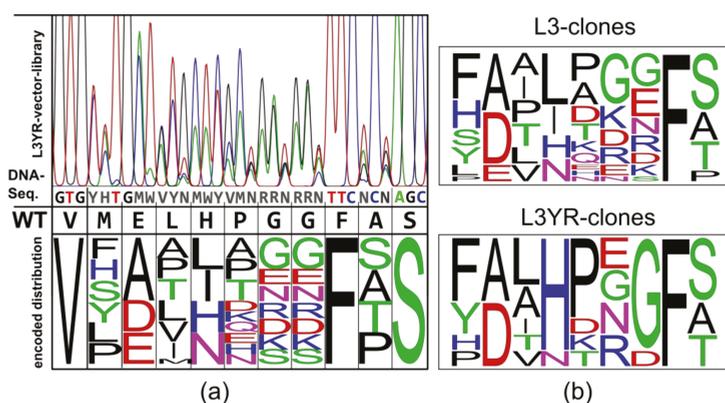

Figure 1. Cyclodextrin glucanotransferases (CGTase) library design and screening process. (**a**) Sequencing of the generated L3YR vector library. The protein sequence 80–90 of the CGTase G825-6 and the library of encoded residues and codon distributions are shown as the frequency of the sequences. (**b**) Logos of the sequence frequency of 22 variants from the L3, and of eight variants from the L3YR experiment.

2.2. Agar Plate Sceening of the CGTase Variants

About 1350 *E. coli* BL21 clones of the L3 and 1550 clones of the L3YR experiment were screened plates for the specific detection of CD7, CD8, and starch-degrading activity (Figure S2). The majority of the L3 clones (>80%) synthesized CD7 and CD8 and about 10% of them in larger amounts as compared to the wild type (WT) CGTase indicated by the formation of larger halos around the colonies. Twenty-two clones from the L3 and eight clones from the L3YR experiments were selected and further characterized by isolating and sequencing the vector DNA followed by the recombinant expression and purification of the CGTase variants. SDS-PAGE analysis confirmed the high purity of the CGTase (Figure S3).

2.3. Comparison of the CD Synthesis by the CGTase Variants

To compare the synthesis of CD by the created L3 variants, their cyclic glucan products up to a size of CD12 were analyzed after 1 h and 24 h of reaction at 40 °C and 50 °C (Figure 2). At a reaction time of 24 h, the WT CGTase produced the maximum yield of the larger CD. After 1 h of reaction at 50 °C, 20 of the 22 L3-variants showed higher yields of total CD (CD7 to CD12) compared to the WT CGTase. All variants except L3-8 also synthesized larger amounts of CD9 to CD12 (Figure 2a). After 24 h of reaction at 50 °C, nine variants (L3-1 to L3-9) produced between 6.9 to 11.5 mg CD7 to CD12, with a higher yield as the WT enzyme (6.4 mg total CD) corresponding to a substrate conversion of up to 58% (Figure 2b). While CD9 to CD12 made up 22% of the total CD synthesized by the WT enzyme, the variants L3-1, L3-3, L3-5 and L3-9 synthesized 38%, 43%, 47% and 51% CD9 to CD12 at 50 °C, respectively. The loop randomization step was repeated with the CGTase G825-6 variant Y183R, which synthesized CD8 as the smallest CD. The CD products were analyzed after 1 h (Figure 2c) and 24 h (Figure 2d) of reaction. The L3YR variants synthesized between 76% and 97% CD9 to CD12 after 1 h of reaction. CD9 to CD12 made up 58% of its products after 24 h of reaction (Figure 2d). In contrast, the variants L3YR-6, L3YR-7 and L3YR-8 synthesized between 91% to 94% CD9 to CD12 at the same reaction conditions, however with a concomitant very low total CD yield of 0.7 to 1.5 mg. At a reaction temperature of 40 °C, these three L3YR variants showed a similar total CD yield as Y183R after 1 h of reaction, but with a higher proportion of CD9 to CD12 (Figure 2e). After 24 h of reaction, the total amount of CD produced by L3YR-8 at 40 °C was three times higher compared to a synthesis performed at 50 °C indicating a destabilizing effect of the alterations at the loop structure of the enzyme (Figure 2f). The variant L3-2-M constructed by incorporating the mutation Y183R in the highly active variant L3-2 resulted in similar CD synthesis properties as displayed by L3YR-6, L3YR-7 and LRYR-8 (Figure 2d).
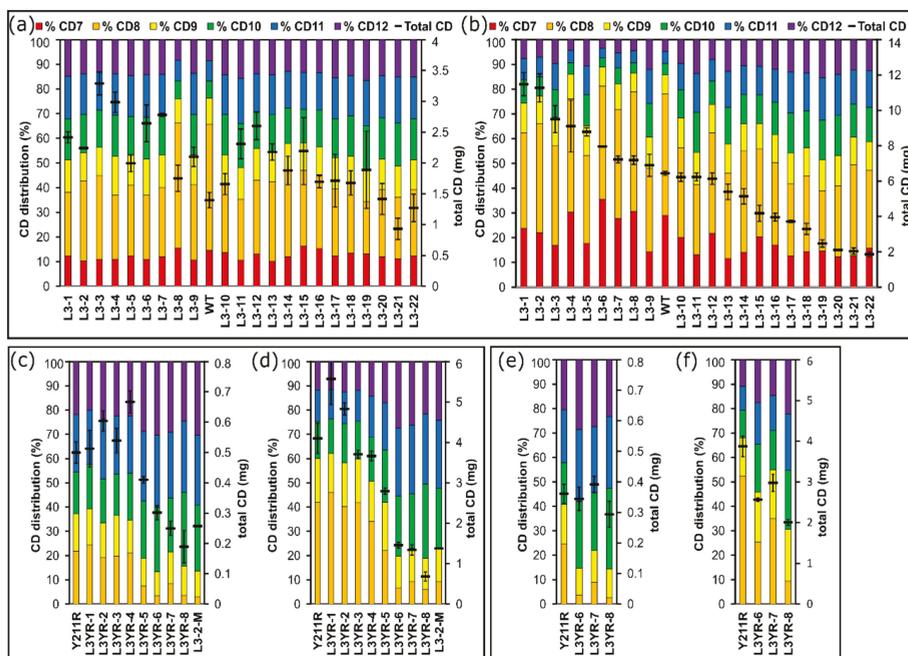
Figure 2. Synthesis of CD by the CGTase variants. The distribution of synthesized CD7 to CD12 is shown at the primary y-axis. The total amount of CD synthesized (CD6 to CD12) is shown at the secondary y-axis. The CD products of the CGTase G825-6 and the L3 variant enzymes obtained after 1 h (**a**) and 24 h (**b**) of reaction are shown. The CD products of the Y183R variant and of eight L3YR variant CGTases obtained after 1 h (**c**) and 24 h (**d**) of reaction at 50 °C are also shown. Selected enzymes were compared to Y183R in a synthesis reaction at 40 °C for 1 h (**e**) and 24 h (**f**). Mean values ($n = 3$) ± S.D. are presented.

2.4. Thermostability of the CGTase Variants

The WT CGTase and the L3-6, L3-8 and L3-18 variants showed a temperature optimum for starch-degrading activity at 60 °C while L3-1, L3-2, L3-21, L3-22 displayed a lower temperature optimum at 50 °C (Figure 3a). The L3YR variants L3YR-2 and L3YR-5 had a similar temperature optimum compared to Y183R at 60 °C, while the other L3YR variants showed an optimum temperature of 50 °C (Figure 3b). Among those, L3YR-6 displayed a residual activity of 81% after 2 h of incubation at 50 °C while no activity loss was observed for Y183R (Figure S4). A determination of the protein melting temperature by nano-DSF confirmed the significant lower thermostability of the L3 variants with a low total CD yield in contrast to highly active variants (L3-1) (Figure S5) (Table 1). While L3YR-1 displayed a similar thermostability as the Y183R CGTase, it was reduced in L3YR6, L3YR-7, L3YR-8 and L3-2M.
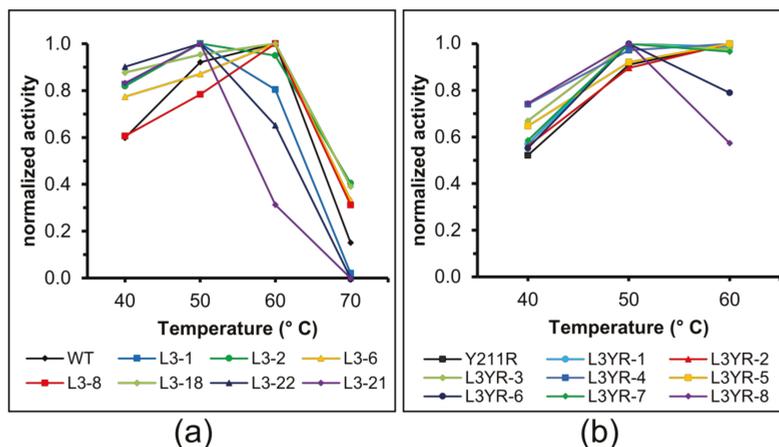
Figure 3. Temperature optimum of the CGTase variants. The starch-degrading activity of L3 variant enzymes (**a**) and L3YR variant enzymes (**b**) was determined between 40 °C and 70 °C. Data represent mean values, $n = 3$.

Table 1. Melting temperature (T_m) of L3 and L3YR CGTase variants determined by nano-DSF. Mean values ($n = 3$) are shown.

CGTase	T_m (°C)
WT	58.7 ± 0.19
L3-1	59.7 ± 0.06
L3-18	58.3 ± 0.08
L3-21	52.5 ± 0.08
Y211R	67.3 ± 0.04
L3YR-1	66.3 ± 0.08
L3YR-6	64.8 ± 0.05
L3YR-7	62.7 ± 0.05
L3YR-8	63.5 ± 0.03
L3-2M	59.9 ± 0.14

3. Discussion

Previously we had engineered the CGTase G825-6 variant Y183R, which produced a high yield of large-ring CD without concomitant formation of CD6 and CD7 [23]. However, its total CD synthesis activity was strongly decreased. To reconstitute its activity, we aimed to modify the enzyme near its substrate binding site by a semi-rational mutagenesis approach. We focused on a loop structure of the protein involved in substrate binding composed of the amino acid residues 81−89 (Figure 4). This region has been previously suggested to control the CD product specificity of CGTases [25,26]. Indeed, the loop structure of CGTases mainly synthesizing CD6 or CD7 (Figure 4a) was different from CGTases with CD8 as the main product (Figure 4b,c). Since the residue F88 in the CGTase G825-6 plays an important role in its product specificity for CD8, it was kept constant while the surrounding residues were targeted for mutagenesis (Figure S6).
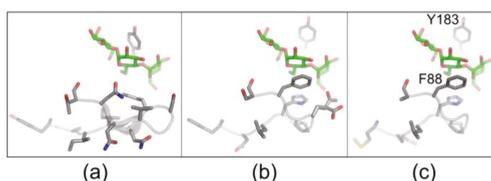
Figure 4. Structure of the L3 loop in different CGTases. X-ray structures from (**a**) *B. circulans* 251 CGTase (PDB: 1CXK), (**b**) *B. clarkii* CGTase (PDB: 4JCM) and (**c**) a model of the CGTase G825-6 based on the PDB structure 4JCM are shown.

3.1. High Fitness of the CGTase Variant Library

Based on the multiple sequence alignments performed, the limited selection of residues used for the design of the variant library reduced the number of possibly translated protein sequences by a factor of ~6 × 10^5 to 0.9 million sequences. The resulting library not only encoded for the majority of amino acid residues occurring in WT CGTases, but also included a set of residues not found in CGTases within the corresponding loop structure (Figure S1, Figure 1a). The residues D, E, K and R were encoded to allow for the formation of salt bridges with other loop regions to stabilize the binding pocket, since previous studies have shown that a crosslinking of CGTases increases their catalytic efficiency [27–29]. The residues D, E and P were encoded to destabilize a β-sheet formation of the loop (Figure 5) with the aim to increase its flexibility [30,31]. Since stabilizing the loop or making it more flexible was expected to influence the CD product specificity, the library was designed to allow both possibilities.
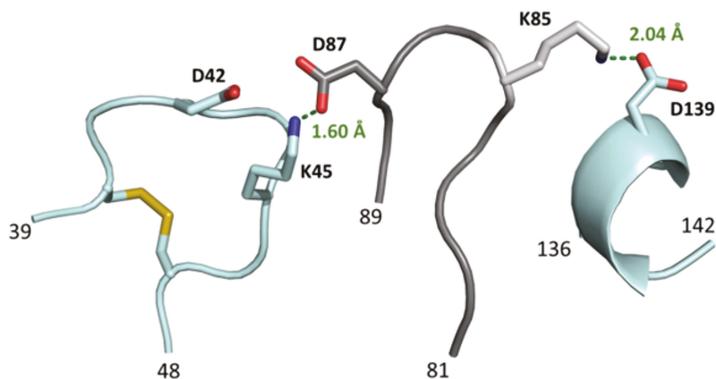

Figure 5. Potential salt bridges between loop elements in the CGTase variants. A model of residues encoded in the L3 library (positions 81−89) predicted to form salt bridges with residues from the adjacent loops 39−48 and 136−142 is shown. The residues D, E, R and K encoded in the library design at position 87 are in close range to K45 and D42 and could form a salt bridge between these loops. K85 may form a salt bridge with D139, located on a neighboring loop.

Instead of a classical directed evolution approach performing small adaptive walks in an imaginary fitness landscape [32], the library design set a high mutation frequency within a narrow sequence window to cover a large sequence space simultaneously. By this strategy, a large set of combinations could be tested for a small sequence element like a loop, while keeping the remaining structure intact. This allowed screening for combinatorial effects that had a beneficial outcome regarding the catalytic activity or product specificity of the CGTase.

In comparison to other loop saturation experiments [33,34], the high fitness of the L3 semi-randomized library can be attributed to several factors, which ensured loop compositions with a

high proportion of active variants: (i) The sufficient distance of the mutagenesis target site to avoid a direct interference with the catalytic triad, (ii) the choice of encoded residues deduced from multiple sequence alignments and (iii) the low conservation of this loop. Furthermore, the introduction of ionic residues should promote synergistic effects by electrostatic interactions between different residues of the loop [35]. Three of the randomized codons encoded amino acid residues with an equal distribution (YHT, MWY, NCN), while for the other randomized codons the ratio of the most abundant to the rarest amino acid encoded was 2:1, except for VYN (4:1). Therefore, only a minor encoded bias occurred in the library.

3.2. Selection of CGTase Variants with Changed CD Product Specificity

By screening for either CD7 or CD8 and of starch-degrading activity in a semi-quantitative agar plate assay, CGTase variants synthesizing CD7 or CD8 in amounts relative to their starch-degrading activity could be selected in parallel (Figure S2). Despite coverage of less than 0.1% of the L3 library, this screening method resulted in the identification of variants with a changed CD product specificity demonstrating the validity of the library design. The L3 variants showed a distribution of amino acid residues similar to the encoded library indicating that certain residues were not preferred among the 22 L3 variants (Figure 1a,b). In the L3YR clones (*n* = 8), residues H84 and G87 occurred in 7 out of 8 protein sequences of the variants, indicating their importance for the synthesis of CD in the presence of the substitution Y183R. Interestingly, L3YR-1, L3YR-2 and L3YR-3 had the sequence 83-LHPXG-87 with X = E or G, similar to the CGTase G825-6 and a CGTase of *Bacillus clarkii*. In fact, the sequences of the variants L3YR-1 and L3-8 were identical to the CGTase from *B. clarkii* (Figure S1). The sequence 81-YALHP of variant L3YR-2 was also found in the protein sequences of other WT CGTases (Figure S1). This suggests that the naturally evolved loop sequence efficiently compensated alterations of the centrally located Y183 residue resulting in active enzyme variants. While the number of investigated variants was too small to draw definitive conclusions, we consider these sequences as interesting targets for further site-directed mutagenesis experiments.

3.3. Variants Synthesizing High Yields of Large-Ring CD

The incorporated mutations resulted in a set of variants, which produced high proportions of CD9 to CD12 (Figure 2). The variant L3-8, carrying the *B. clarkii* CGTase/G825-6 CGTase loop consensus sequence 83-LHP-85 showed the same composition of its CD products as the WT enzyme, indicating that this sequence could be involved in the suppression of the synthesis of larger CD. L3YR variants with the composition 83-LHPGG or LHPEG, corresponding to the sequences of the *B. clarkii* CGTase/G825-6, synthesized a similar proportion of large-ring CD compared to Y183R. In contrast, L3YR-6, L3YR-7, L3YR 8 and L3-2M showed a strongly decreased synthesis of CD8. Whether these changes were caused by the alteration of the sequence 83-LHP, the position 86 or by a combination of both remains ambiguous. The CD yield of these variants increased at lower reaction temperatures, indicating that they were less thermostable compared to the WT CGTase.

From 18 possible combinations of FA, FD and YA of the loop amino acid residues 81 and 82, 16 were present in the 30 variant sequences and represent combinations also found in WT CGTases (Table S1, Figure S1). When the altered sequence of variant L3-2 was cloned into Y183R, the resulting variant L3-2M showed similar properties as L3YR-6, L3YR-7 and L3YR-8 (Figure 2d). The replacement of Y at position 183 with R resulted in the synthesis of almost only large-ring CD composed of CD9 to CD12, however concomitant with a decreased overall CD yield.

The most efficient variants from the L3 experiment displayed E, R or K at position 87, amino acids able to electrostatic interaction with D42 and K45 of the nearby loop. Variant L3-3 with K85 and E87 possibly formed salt bridges between loop 81-89, loop 39-48 and loop 136-142 (Figure 5). All three loops were positioned near the binding site and have been reported to contribute to the activity and CD product specificity of the CGTase [11,36,37]. Accordingly, our results also indicate that a stabilization of the loop resulted in an increased synthesis of large-ring CD. However, less efficient L3 variants

likewise had E, D and K at position 87, indicating that surrounding residues also contribute to the observed changes in CD synthesis and enzyme stability.

3.4. The L3 Loop Contributes to the Thermal Stability of the CGTase

Several L3YR variants showed high CD synthesis activities within 1 h of reaction, which were significantly lower after 24 h, suggesting a thermal inactivation of the enzyme during longer reaction times, due to the altered loop (Figure S4). A determination of the thermal stability and optimum temperature of the variant enzymes confirmed this assumption. The variants L3YR-6, L3YR-7 and L3YR-8 showed a lower T_M and optimum temperature compared to the WT enzyme (Figure 3). When the CD synthesis reaction with these variants was performed at 40 °C, these L3YR variants indeed showed a 1.8 to 2.9-fold higher yield of total CD compared to 50 °C (Figure 2e,f).

4. Materials and Methods

4.1. Generation of the Vector Library

A CGTase expression system in *E. coli* based on a pET20b+ vector harboring an expression cassette encoding for a mature CGTase from the alkaliphilic *Bacillus* sp. G825-6 with an N-terminal DacD signal peptide was used [38]. A partially randomized megaprimer 5'-agc ccg ccg att gaa aat gtg yht gmw vyn nmy vmn rrn rrn ttc ncn agc tat cat ggc tat tgg ggc was amplified by PCR (Phusion HF Kit, New England Biolabs, Frankfurt, Germany,) using primers hybridizing to the flanking sites to generate a dsDNA megaprimer. The randomized ssDNA primer (1 pmol) was added as a template for a 50 µL PCR reaction for 30 cycles according to the supplier's manual. The product was purified by agarose gel electrophoresis. The megaprimer DNA and the template vector pET20b(+):dacD-cgt was used in a molar ratio of 10:1 in a second PCR reaction (18 cycles, T_{Anneal}: 60 °C). The product designated pET20b(+):dacD-cgt-L3 was digested with DpnI and ultracompetent *E. coli* XL10-Gold cells (New England Biolabs, Frankfurt, Germany) were used for transformation [39]. After 1 h, an aliquot (50 µL) was plated on LB agar plates containing 70 µg mL^{-1} ampicillin (LB-amp), and the residual cells were further incubated for 12 h before harvesting the cells. Plasmids were extracted to obtain the L3 vector library and used for the transformation of *E. coli* BL21 (DE3). The procedure was repeated with a mutated version of the vector encoding the CGTase G825-6 with the substitution Y183R to generate the L3YR vector library.

4.2. Agar Plate Screening

Single colonies of *E. coli* BL21(DE3) pET20b(+):dacD-cgt-L3 were transferred to LB-amp, congo red [40], and phenolphthalein (LB agar pH 7.4, 40 mg L^{-1} phenolphthalein) agar plates [41]. Both congo red and phenolphthalein agar plates were supplemented with 10 g L^{-1} soluble starch and 70 µg mL^{-1} ampicillin. Plates for the L3 experiment were incubated at 37 °C for 24 h. For the screening of the L3YR clones the congo red plates were transferred after 14 h from 37 °C to 50 °C and further incubated for 8 h. The size of the halos formed on the congo red plates was estimated after 14 h and 24 h of incubation. Subsequently, starch-degrading activity was visualized by covering the congo red plates with 1% (*w/w*) Lugol's solution and halo formation were estimated after 1 min. After 24 h of incubation, the phenolphthalein plates were overlayed with a solution containing 1 M NaOH, 0.1 M glycine, 1 M NaCl. After 1 min of incubation the size of the halos was estimated. 28 clones were picked per plate. Each plate contained two clones for the positive control *E. coli* BL21(DE3) pET20b(+):dacD-cgt, a second positive control encoding the G825-CGTase variant D358R with low CD7- and CD8-synthesizing activity, and a negative control (*E. coli* BL21(DE3) pET20b(+):tfcut, encoding a cutinase [42]).

4.3. Recombinant Protein Production and Analysis

Positive clones were repeatedly screened, and selected clones were used for expression of the proteins in 50 mL cultures. The enzyme in the extracellular fraction was purified by starch adsorption [38]. The starch-degrading activity and the protein concentration of the purified enzyme variants were determined, as previously described [38].

4.4. CD Synthesis and Analysis

For the determination of the product spectrum synthesized by the CGTase variants, 0.2 µg purified protein (0.4 µg for the L3YR variants) was added to a 20 g L^{-1} soluble starch substrate (soluble potato starch, CAS9005-84-9; Merck KGaA, Darmstadt, Germany) in a total volume of 1 mL. Reactions were performed at 40 °C and 50 °C in 25 mM Tris-HCl pH 8.5 containing 10 mM KCl and 5 mM MgCl$_2$. Samples were analyzed by high pressure anion exchange chromatography with pulsed amperometric detection after 1 h and 24 h of reaction [23].

4.5. Determination of Temperature Optimum and Thermostability of the Variants

The temperature optimum of the variants was determined by measuring their starch-degrading activity between 40 °C and 70 °C [38]. The thermostability of Y211R and L3YR-6 was analyzed by determining its residual activity after incubation of the enzyme solution at 50 °C for 2 h. Nano differential scanning fluorimetry (nano-DSF) (Prometheus NT.48, Nanotemper Technologies, Munich, Germany) based on the tryptophan fluorescence ratio 350/330 nm (20 °C to 95 °C with 1 °C/min) was used to determine the melting temperature (T$_m$) of the proteins, which were calculated by first derivative analysis.

4.6. Protein Sequence Alignments and Molecular Modeling

Protein sequences from 30 CGTases were obtained from the National Center of Biotechnology Information (www.ncbi.nlm.nih.gov; accession numbers in Table S1) and compared to the CGTase G825-6 sequence [43] using a clustalW algorithm implemented in the homology software MEGA, version 6.06 [44]. Based on the X-ray structure of the CGTase from *Bacillus clarkii* (PDB:4JCM), obtained from the Protein Data Bank (PDB) (www.rcsb.org), structure homology modeling was performed using the SWISS-MODEL platform [45,46]. The PyMOL molecular graphic system (v0.99, Schrödinger, LCC) was used to model amino acid substitutions. The superimposed structure of a maltononaose substrate was derived from PDB:1CXK [47].

5. Conclusions

A strongly compressed library with high fitness was constructed using a semi-rational design for the alteration of the CD synthesis activity of a CGTase. The screening method based on three different markers allowed both the selection for variants with changed CD synthesis activity and CD product specificity. With this approach, we were able to screen a reasonable small number of clones to obtain CGTase variants synthesizing preferentially large-ring CD. This approach resulted in enzyme variants producing three times higher total amounts of CD7 to CD12 with a high proportion of CD9 and CD12. Some of the variants even synthesized almost solely CD10, CD11 and CD12. The results demonstrate that by semi-rational design, CGTase variants specifically producing these large-ring CD can be generated providing a previously difficult to access group of novel host molecules in supramolecular complexing reactions.

References

1. Cobb, R.E.; Chao, R.; Zhao, H. Directed evolution: Past, present and future. *AIChE J.* **2013**, *59*, 1432–1440. [CrossRef] [PubMed]

2. Dalby, P.A. Strategy and success for the directed evolution of enzymes. *Curr. Opin. Struct. Biol.* **2011**, *21*, 473–480. [CrossRef] [PubMed]

3. Hsieh, P.-C.; Vaisvila, R. Protein engineering: Single or multiple site-directed mutagenesis. *Methods Mol. Biol.* **2013**, *978*, 173–186. [CrossRef] [PubMed]

4. Chica, R.A.; Doucet, N.; Pelletier, J.N. Semi-rational approaches to engineering enzyme activity: Combining the benefits of directed evolution and rational design. *Curr. Opin. Biotechnol.* **2005**, *16*, 378–384. [CrossRef] [PubMed]

5. Patrick, W.M.; Firth, A.E. Strategies and computational tools for improving randomized protein libraries. *Biomol. Eng.* **2005**, *22*, 105–112. [CrossRef] [PubMed]

6. Lutz, S. Beyond directed evolution-semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* **2010**, *21*, 734–743. [CrossRef] [PubMed]

7. Chen, M.M.Y.; Snow, C.D.; Vizcarra, C.L.; Mayo, S.L.; Arnold, F.H. Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes. *Protein Eng. Des. Sel.* **2012**, *25*, 171–178. [CrossRef] [PubMed]

8. Tang, L.; Gao, H.; Zhu, X.; Wang, X.; Zhou, M.; Jiang, R. Construction of "small-intelligent" focused mutagenesis libraries using well-designed combinatorial degenerate primers. *BioTechniques* **2012**, *52*, 149–158. [CrossRef] [PubMed]

9. DePinto, J.A.; Campbell, L.L. Purification and properties of the amylase of Bacillus macerans. *Biochemistry* **1968**, *7*, 114–120. [CrossRef] [PubMed]

10. Qi, Q.; Zimmermann, W. Cyclodextrin glucanotransferase: From gene to applications. *Appl. Microbiol. Biotechnol.* **2005**, *66*, 475–485. [CrossRef] [PubMed]

11. Uitdehaag, J.C.; Kalk, K.H.; van der Veen, B.A.; Dijkhuizen, L.; Dijkstra, B.W. The cyclization mechanism of cyclodextrin glycosyltransferase (CGTase) as revealed by a gamma-cyclodextrin-CGTase complex at 1.8—A resolution. *J. Biol. Chem.* **1999**, *274*, 34868–34876. [CrossRef] [PubMed]

12. Del Valle, E. Cyclodextrins and their uses: A review. *Process Biochem.* **2004**, *39*, 1033–1046. [CrossRef]

13. Endo, T.; Nagase, H.; Ueda, H.; Kobayashi, S.; Nagai, T. Isolation, purification, and characterization of cyclomaltodecaose (.EPSILON.-Cyclodextrin), cyclomaltoundecaose (.ZETA.-Cyclodextrin) and cyclomaltotridecaose (.THETA.-Cyclodextrin). *Chem. Pharm. Bull.* **1997**, *45*, 532–536. [CrossRef]

14. Larsen, K.L.; Endo, T.; Ueda, H.; Zimmermann, W. Inclusion complex formation constants of α-, β-, γ-, δ-, ε-, ζ-, η- and θ-cyclodextrins determined with capillary zone electrophoresis. *Carbohydr. Res.* **1998**, *309*, 153–159. [CrossRef]

15. Ueda, H.; Endo, T. Large-ring cyclodextrins. In *Cyclodextrins and Their Complexes: Chemistry, Analytical Methods, Applications*; Dodziuk, H., Ed.; Wiley-VCH: Weinheim, Germany, 2006; pp. 370–380. ISBN 9783527608980.

16. Akasaka, H.; Endo, T.; Nagase, H.; Ueda, H.; Kobayashi, S. Complex formation of cyclomaltononaose delta-cyclodextrin (delta-CD) with macrocyclic compounds. *Chem. Pharm. Bull.* **2000**, *48*, 1986–1989. [CrossRef] [PubMed]

17. Dodziuk, H.; Ejchart, A.; Anczewski, W.; Ueda, H.; Krinichnaya, E.; Dolgonos, G.; Kutner, W. Water solubilization, determination of the number of different types of single-wall carbon nanotubes and their partial separation with respect to diameters by complexation with η-cyclodextrin. *Chem. Commun.* **2003**, *9*, 986–987. [CrossRef]

18. Assaf, K.I.; Gabel, D.; Zimmermann, W.; Nau, W.M. High-affinity host-guest chemistry of large-ring cyclodextrins. *Org. Biomol. Chem.* **2016**, *14*, 7702–7706. [CrossRef] [PubMed]

19. Terada, Y.; Yanase, M.; Takata, H.; Takaha, T.; Okada, S. Cyclodextrins are not the major cyclic alpha-1,4-glucans produced by the initial action of cyclodextrin glucanotransferase on amylose. *J. Biol. Chem.* **1997**, *272*, 15729–15733. [CrossRef] [PubMed]
20. Li, Z.; Wang, M.; Wang, F.; Gu, Z.; Du, G.; Wu, J.; Chen, J. Gamma-cyclodextrin: A review on enzymatic production and applications. *Appl. Microbiol. Biotechnol.* **2007**, *77*, 245–255. [CrossRef] [PubMed]
21. Terada, Y.; Sanbe, H.; Takaha, T.; Kitahata, S.; Koizumi, K.; Okada, S. Comparative study of the cyclization reactions of three bacterial cyclomaltodextrin glucanotransferases. *Appl. Environ. Microbiol.* **2001**, *67*, 1453–1460. [CrossRef] [PubMed]
22. Han, R.; Li, J.; Shin, H.-D.; Chen, R.R.; Du, G.; Liu, L.; Chen, J. Recent advances in discovery, heterologous expression, and molecular engineering of cyclodextrin glycosyltransferase for versatile applications. *Biotechnol. Adv.* **2014**, *32*, 415–428. [CrossRef] [PubMed]
23. Sonnendecker, C.; Melzer, S.; Zimmermann, W. Engineered cyclodextrin glucanotransferases from Bacillus sp. G-825-6 produce large-ring cyclodextrins with high specificity. *Microbiologyopen* **2018**, e757. [CrossRef] [PubMed]
24. Sonnendecker, C.; Zimmermann, W. Domain shuffling of cyclodextrin glucanotransferases for tailored product specificity and thermal stability. *FEBS Open Bio* **2019**, *9*, 384–395. [CrossRef] [PubMed]
25. Kim, Y.H.; Bae, K.H.; Kim, T.J.; Park, K.H.; Lee, H.S.; Byun, S.M. Effect on product specificity of cyclodextrin glycosyltransferase by site-directed mutagenesis. *Biochem. Mol. Biol. Int.* **1997**, *41*, 227–234. [CrossRef] [PubMed]
26. Goh, K.M.; Mahadi, N.M.; Hassan, O.; Rahman, R.N.Z.R.A.; Illias, R.M. A predominant β-CGTase G1 engineered to elucidate the relationship between protein structure and product specificity. *J. Mol. Catal. B Enzym.* **2009**, *57*, 270–277. [CrossRef]
27. Kumar, S.; Nussinov, R. Close-range electrostatic interactions in proteins. *ChemBioChem* **2002**, *3*, 604. [CrossRef]
28. Kaulpiboon, J.; Pongsawasdi, P.; Zimmermann, W. Molecular imprinting of cyclodextrin glycosyltransferases from Paenibacillus sp. A11 and Bacillus macerans with gamma-cyclodextrin. *FEBS J.* **2007**, *274*, 1001–1010. [CrossRef] [PubMed]
29. Kaulpiboon, J.; Pongsawasdi, P.; Zimmermann, W. Altered product specificity of a cyclodextrin glycosyltransferase by molecular imprinting with cyclomaltododecaose. *J. Mol. Recognit.* **2010**, *23*, 480–485. [CrossRef] [PubMed]
30. Chou, P.Y.; Fasman, G.D. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1978**, *47*, 45–148. [CrossRef] [PubMed]
31. Leemhuis, H.; Rozeboom, H.J.; Dijkstra, B.W.; Dijkhuizen, L. Improved thermostability of bacillus circulans cyclodextrin glycosyltransferase by the introduction of a salt bridge. *Proteins* **2004**, *54*, 128–134. [CrossRef] [PubMed]
32. Gillespie, J.H. Molecular evolution over the mutational landscape. *Evolution* **1984**, *38*, 1116–1129. [CrossRef] [PubMed]
33. Hörtnagel, K.; Voloshin, O.N.; Kinal, H.H.; Ma, N.; Schaffer-Judge, C.; Camerini-Otero, R.D. Saturation mutagenesis of the E. coli RecA loop L2 homologous DNA pairing region reveals residues essential for recombination and recombinational repair. *J. Mol. Biol.* **1999**, *286*, 1097–1106. [CrossRef] [PubMed]
34. Duenas-Decamp, M.; Jiang, L.; Bolon, D.; Clapham, P.R. Saturation mutagenesis of the HIV-1 envelope CD4 binding loop reveals residues controlling distinct trimer conformations. *PLoS Pathog.* **2016**, *12*, e1005988. [CrossRef] [PubMed]
35. Lehmann, M.; Wyss, M. Engineering proteins for thermostability: The use of sequence alignments versus rational design and directed evolution. *Curr. Opin. Biotechnol.* **2001**, *12*, 371–375. [CrossRef]
36. Parsiegla, G.; Schmidt, A.K.; Schulz, G.E. Substrate binding to a cyclodextrin glycosyltransferase and mutations increasing the gamma-cyclodextrin production. *Eur. J. Biochem.* **1998**, *255*, 710–717. [CrossRef] [PubMed]
37. Leemhuis, H.; Kelly, R.M.; Dijkhuizen, L. Engineering of cyclodextrin glucanotransferases and the impact for biotechnological applications. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 823–835. [CrossRef] [PubMed]
38. Sonnendecker, C.; Wei, R.; Kurze, E.; Wang, J.; Oeser, T.; Zimmermann, W. Efficient extracellular recombinant production and purification of a Bacillus cyclodextrin glucanotransferase in *E. coli*. *Microb. Cell. Fact.* **2017**, *16*, 87. [CrossRef] [PubMed]

39. Sambrook, J.; Russell, D.W. The inoue method for preparation and transformation of competent e. Coli: "ultra-competent" cells. *CSH Protoc.* **2006**, *2006*. [CrossRef] [PubMed]

40. Hamaker, K.; Tao, B.Y. Screening of Gamma Cyclodextrin-Producing Recombinant *E. coli* using congo red dye on solid complex media. *Starch/Stärke* **1993**, *45*, 181–182. [CrossRef]

41. Park, C.S.; Park, K.H.; Kim, S.H. A rapid screening method for alkaline BETA-cyclodextrin glucanotransferase using phenolphthalein-methyl orange-containing-solid medium. *Agric. Niol. Chem.* **1989**, *53*, 1167–1169. [CrossRef]

42. Oeser, T.; Wei, R.; Baumgarten, T.; Billig, S.; Föllner, C.; Zimmermann, W. High level expression of a hydrophobic poly (ethylene terephthalate)-hydrolyzing carboxylesterase from Thermobifida fusca KW3 in *E. coli* BL21(DE3). *J. Biotechnol.* **2010**, *146*, 100–104. [CrossRef] [PubMed]

43. Melzer, S.; Sonnendecker, C.; Föllner, C.; Zimmermann, W. Stepwise error-prone PCR and DNA shuffling changed the pH activity range and product specificity of the cyclodextrin glucanotransferase from an alkaliphilic Bacillus sp. *FEBS Open Bio* **2015**, *5*, 528–534. [CrossRef] [PubMed]

44. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef] [PubMed]

45. Guex, N.; Peitsch, M.C.; Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* **2009**, *30*, S162–S173. [CrossRef] [PubMed]

46. Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Gallo Cassarino, T.; Bertoni, M.; Bordoli, L.; et al. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **2014**, *42*, W252–W258. [CrossRef] [PubMed]

47. Uitdehaag, J.C.; Mosi, R.; Kalk, K.H.; van der Veen, B.A.; Dijkhuizen, L.; Withers, S.G.; Dijkstra, B.W. X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the alpha-amylase family. *Nat. Struct. Biol.* **1999**, *6*, 432–436. [CrossRef] [PubMed]

Article

Low-Level Organic Solvents Improve Multienzyme Whole-Cell Catalytic Synthesis of Myricetin-7-*O*-Glucuronide

Yan Yang [1,2], Min-Zhi Liu [2], Yun-Song Cao [2], Chang-Kun Li [3] and Wei Wang [1,2,*]

[1] State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China; yangyan@imm.ac.cn

[2] Key Laboratory of Biosynthesis of Natural Products of National Health Commission of the People's Republic of China, Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China; lmzluoxue@imm.ac.cn (M.-Z.L.); caoyunsong@imm.ac.cn (Y.-S.C.)

[3] Beijing Branch, Shimadzu (China) Co. Ltd., Beijing 100020, China; fxlck@shimadzu.com.cn

[*] Correspondence: wwang@imm.ac.cn; Tel.: +86-10-631-65196

Abstract: Multienzyme whole-cell biocatalysts are preferred in industrial applications, and two major concerns regarding the use of these biocatalysts, cell viability and cell membrane integrity, must be addressed. In this work, the transformation of myricetin to myricetin-7-*O*-glucuronide catalyzed by an engineered *Escherichia coli* strain was taken as the model reaction to examine the impacts of low-level organic solvents on whole-cell biocatalysis. Low-level organic solvents (2%, *v/v*) showed a significant increase (roughly 13-fold) in myricetin-7-*O*-glucuronide yields. No obvious compromises of cellular viability and integrity were observed by a flow cytometry assay or in the determination of extracellular protein leakage, suggesting the addition of low-level organic solvents accommodates whole *E. coli* cells. Furthermore, a scaled-up reaction was conducted to test the capability and efficiency of whole-cell catalysis in the presence of organic solvents. This study presents a promising and simple means to enhance the productivity of multienzyme whole-cell catalysis without losing the barrier functions of the cell membrane.

Keywords: engineered *Escherichia coli*; flavonoid glucuronides; multienzyme whole-cell biocatalyst; organic solvents
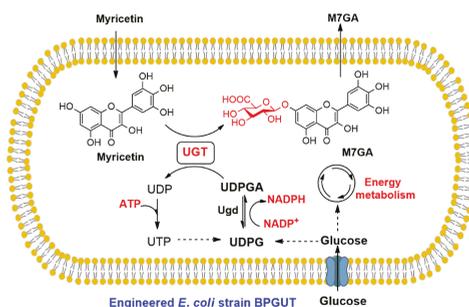
1. Introduction

Multistep whole-cell catalysis is becoming a fast-developing area in the field of biocatalysis [1,2]. A single, tailored, whole-cell biocatalyst containing all of the required enzymes not only offers a high degree of selectivity, including regio-, stereo-and enantio-selectivity [3,4], but is also superior to in vitro enzyme biotransformation, as it circumvents cell lysis, enzyme purification, and the addition of external cofactors (e.g., NAD(P)H and ATP). In addition to these capacities, microbial cells provide a protective environment for enzymes, preventing conformational changes from occurring in the protein structure in non-conventional media or under harsh reaction conditions [1,5].

Whole-cell biocatalysts have been widely used as an excellent alternative means to directly convert different substrates into fine chemicals [2,6], but they are generally less efficient than reactions catalyzed by isolated enzymes by a factor of one or two orders of magnitude [7,8]. The most prominent limitations are unimpeded transport of the substrate into and the final product out of the cell [8]. Since hydrophobic substrates can enter cells through passive diffusion [9], the solubility of substrates is generally considered a limiting factor, and the permeability of the cell membrane as a

mass transfer barrier is a most common concern for whole-cell bioprocesses. Although many chemical (e.g., by adding detergents or solvents) and physical (e.g., temperature shock) treatments are used to improve the permeability of cell membranes, they may compromise cell integrity, spur the leakage of cellular components (e.g., diffusional loss of NAD(P)H), and affect downstream processes [10]. Given the attractive advantages of conventional water-miscible organic solvents, including their simple management, high solubility for various organic substances, and capacities to improve the enantio-selectivity and activity of biocatalysts [11,12], organic solvents have been used as "cosolvents" and "permeability enhancers" to stimulate substrate transfer across cells [7,10], but the choice and amount of added organic solvent are highly empirical and generally undesirable for large-scale applications [13]. In most single-enzyme whole-cell catalytic processes, dimethyl sulfoxide (DMSO) and methanol (MeOH) are commonly used as solubility enhancers to promote the biosynthesis of glycosides and hydrophobic compounds [14,15]. The use of a recombinant single enzyme overexpressed in *Escherichia coli* is employed as a whole-cell biocatalyst treated with 60% acetone (*v/v*) [16] or 10% ethanol (EtOH) [11] to enhance the permeabilization of cell membranes and, further, to promote high yields of L-phenylalanine and L-menthol.

Although the above studies have demonstrated that organic solvents are effective in enhancing the yields of whole-cell reactions by affecting substrate solubility or cellular permeability, a widespread means to balance the contradiction between reducing barrier-like actions and maintaining the integrity of the cell membrane remain elusive. In addition, the presence of an intact cell is required for whole-cell catalysis, and especially for multienzyme whole-cell reactions coupled with the (re)generation of cofactors and energy molecules. While adaptation of bacterial cells to the organic solvent is an important issue to be addressed [17,18], little attention has been paid to the lesser negative effects on cell survival and membrane integrity resulting from the use of low-level organic solvents that can affect bioconversion efficiency. Indeed, biocatalytic systems, especially whole-cell ones, can be very sensitive to organic solvents.

In this study, an engineered *E. coli* BPGUT [19] performing the multienzyme cascade biosynthesis of a uridine diphosphate (UDP) sugar donor, which requires the (re)generation of redox (NAD(P)H) and energy ATP, was used as a multienzyme whole-cell model biocatalyst to test the effects of organic solvents on the conversion of myricetin into myricetin-7-*O*-glucuronide (M7GA) (Scheme 1). Myricetin is a natural plant-derived flavonoid that has been well recognized for its nutraceutical value, but which exhibits limited aqueous solubility [20]. M7GA, a structural analogue of scutellarein-7-*O*-glucuronide [21], exhibits the potential capacity to treat cardio–cerebral vascular diseases. Herein, we focused on the influence of different organic solvents on the one-pot whole-cell reaction. The effects of the addition of low-level water-miscible organic solvents on solubility and bioconversion rates and on microbial cellular viability and membrane integrity were investigated. Additionally, the whole-cell reaction for producing M7GA with the use of optimal organic solvent was scaled up.

Scheme 1. Production of myricetin-7-*O*-glucuronide (M7GA) from myricetin with engineered *Escherichia coli* BPGUT cells used as the whole-cell biocatalyst.

2. Results and Discussion

2.1. Effects of the Organic Solvents on Bioconversion

As many flavonoid aglycones have been suggested to be able to move into cells through passive diffusion [22–24], a high concentration gradient of flavonoid aglycones in a metabolic cell aqueous system may allow for a higher level of cross-membrane uptake. The screening test involving the addition of 2%–10% organic solvents to the reaction system (Figure 1A and Table S1) showed that all of the organic solvents tested (except for 10% acetone and 10% EtOH) triggered higher degrees of yield, with the highest value observed following the addition of 2% acetone, 2% EtOH and 2% MeOH (over 13-fold increase). M7GA yields underwent a steep decline when acetone and EtOH concentrations reached 10%, exhibiting 10%–27% decreases relative to the control. The conversion rate underwent an approximate 9-fold increase relative to the control when the *N,N*-dimethylformamide (DMF) and DMSO concentrations were 2%, and the product yield decreased slowly as the content of DMF and DMSO sequentially increased.

The solubilities of myricetin and M7GA coupled with working concentrations of organic solvent were evaluated. As shown in Figure 1B and Table S1, the addition of organic solvents can induce minor improvements (ranging from 16% to 200%) in the solubility of myricetin in the presence of DMF and DMSO (2%, *v/v*), of acetone and EtOH (2% or 5%, *v/v*) or of MeOH (below 10%, *v/v*) relative to the control. However, the solubility of myricetin with acetone or DMF or DMSO (10%, *v/v*) was dramatically stimulated by approximately 22- to 48-fold relative to the one obtained from the control. In contrast to the substrate myricetin, the product M7GA showed good solubility in modified M9 medium and those with low-level organic solvents (Figure 1C and Table S1). Almost 68% of M7GA dissolved in modified M9 medium, with only minor differences in the solubility of M7GA (less than 31%) among the tested solutions, which indicated that the inherent hydrophilic property of M7GA rather than the presence of organic solvent improved the solubility or the transport pace of the product.

A step-down relationship was found between yields and organic solvent content levels in the reaction system, with an optimal value attained when the volume of organic solvents was only 2% (Figure 1A). A step-up relationship was found between solubility levels and organic solvent content, with an optimal value obtained when the volumes of certain organic solvents reached 10% (*v/v*) (Figure 1B). Low concentrations of organic solvents exhibited accelerating potential during the biocatalytic conversion of myricetin into M7GA, with a slight increase in solubility. Moreover, when plotting production yields against the solubility of myricetin for the respective M9 solution with organic solvents, an obvious decrease in yields accompanied by an increase in myricetin solubility was found (Figure S1), possibly related to interactions between the cells and added organic solvents, which may have affected the surface tension, fluidity, permeability, or integrity of cellular membranes or further caused the leakage of intracellular small molecules (e.g., UDP-sugars, NAD(P)H, or ATP). DMSO can increase membrane fluidity, reduce membrane thickness, form transient water pores, or destroy the membrane structure with increases in concentrations [13,25]. Short-chain alkanols (MeOH and EtOH) can partition into cell membranes and reduce surface tension. As concentrations increase, the membrane becomes overly fluid, which further damages its integrity [26–28]. Thus, a low concentration of organic solvent reduces mass-transfer resistance from the membranes of whole cells, and a high concentration of organic solvent damages the membrane structure and integrity. Certainly, low or high concentrations of organic solvent should be correlated with cell density (OD_{600} = 6.0, 12 g cells L^{-1} in this study).
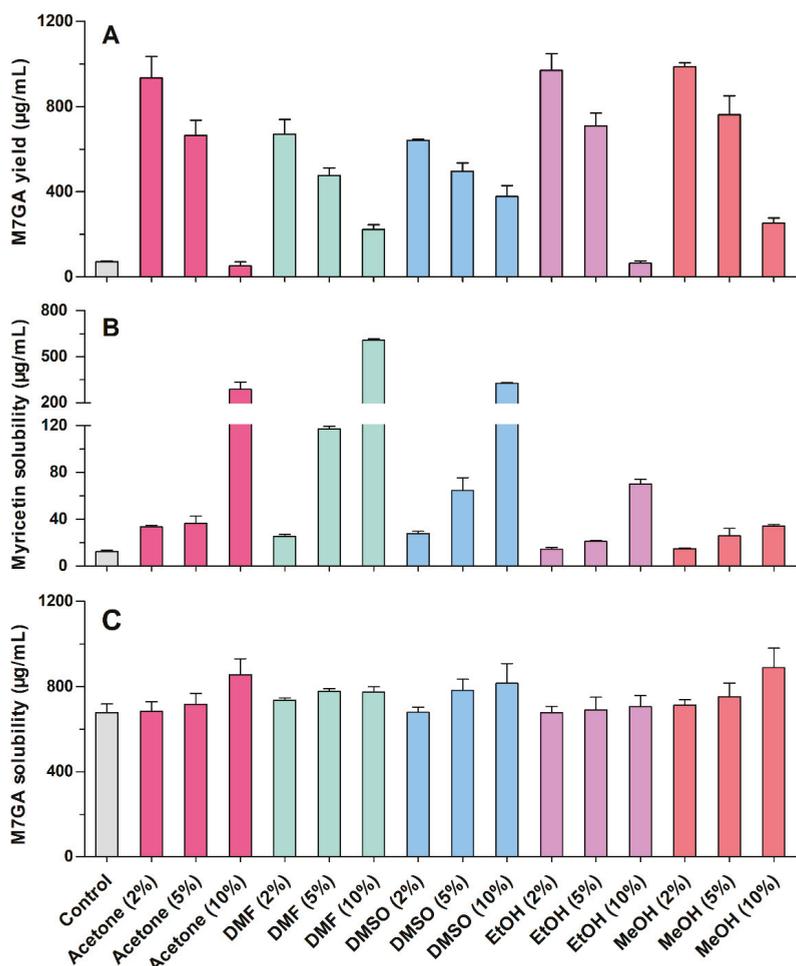
Figure 1. Effects of different organic solvents on the whole-cell reaction and the substrate and product solubilities. (**A**) Bioconversion of myricetin to myricetin-7-*O*-glucuronide (M7GA) catalyzed by *E. coli* BPGUT cells in the presence of organic solvents; (**B**) The solubility of myricetin and (**C**) the solubility of M7GA obtained in the modified M9 medium containing different concentrations of organic solvents.

2.2. Cell Viability and Its Impacts on Bioconversion

To clarify the effects of the addition of organic solvents on the catalytic activity of *E. coli* BPGUT in converting myricetin into M7GA, we assumed that these additives may primarily affect bioconversion through their interactions with cell membranes. Whether cell membranes are damaged or not can be determined from cell viability. In this study, the viability of the treated cells was assessed through the flow cytometry (FCM) test, as dead cells with damaged cytoplasmic membranes exhibit red fluorescence (due to propidium iodide (PI) staining), and live cells with intact membranes appear green (due to SYTO9 staining).

As can be observed from the FCM micrographs (Figure 2 and Table S2), the addition of 20% acetone and 20% EtOH seriously disrupted cell membranes, leading to cell death; the disruptions became much more severe when cells suffered from 20% concentrations of acetone and EtOH than from 2% concentrations, reflecting a result distinct from that observed for cells treated with DMSO

and DMF (Figure 2). Interestingly, although the distinction in production yield was obvious when different organic solvents (2%–10%) were added, the FCM test for cell viability did not show obvious differences (Figure 2). However, when the cells were treated with 20% acetone and 20% EtOH, the damage was the most severe, and coincidently, the yields of M7GA were the lowest when 10% acetone and 10% EtOH were used.
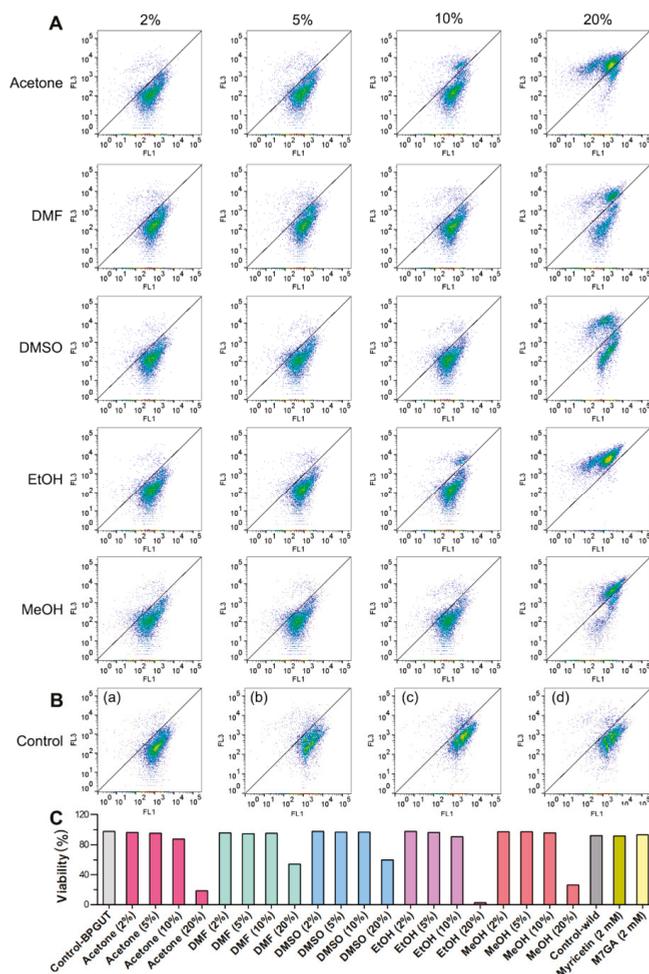

Figure 2. Viability of *E. coli* BPUGT cells or wild-type cells after being treated for 6 h at 30 °C in a modified M9 solution containing 2% (*v/v*), 5% (*v/v*), 10% (*v/v*), and 20% (*v/v*) of organic solvents, or 2 mM of myricetin and M7GA, respectively. (**A**) Flow cytometry (FCM) images of organic solvent-treated *E. coli* BPUGT cells; (**B**) FCM images of untreated *E. coli* BPUGT cells (**a**), untreated *E. coli* wild type cells (**b**), myricetin-treated *E. coli* wild type cells (**c**), and M7GA-treated *E. coli* wild type cells (**d**); (**C**) The cell viability as determined by FCM.

Cell protein leakage was observed after the removal of the treated cells, which serves as another indication of damage to the cell membrane on one hand and reflects the leakage of intracellular small molecules on the other. The levels of protein leakage increased with the increase of organic solvents, especially for acetone and EtOH (Figure 3 and Table S3). The 2%–10% organic solvents (DMF and

DMSO) only spurred a less than twofold increase with respect to the control reaction system, which might be due to the *E. coli* surface protein. Extracellular protein leakage resulted in a sharp increase when the cells were exposed to acetone (20%) and EtOH (20%)—containing aqueous solution. This finding was in accordance with that of the FCM tests showing a decrease in the viability of the *E. coli* cells. These results suggest that the reduced barrier function of cell membranes accompanied by an increase in organic solvents can result in the further leakage of intracellular components essential to multienzyme whole-cell biocatalysts.
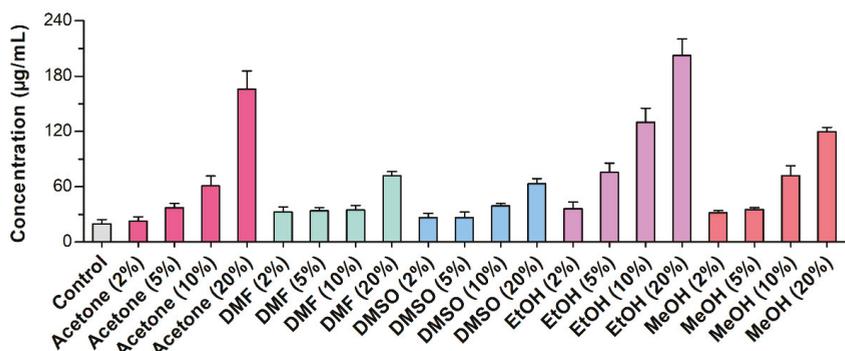

Figure 3. Concentrations of proteins released from *E. coli* BPGUT cells after the cells were exposed to 2% (*v/v*), 5% (*v/v*), 10% (*v/v*) and 20% (*v/v*) of acetone, DMF, DMSO, EtOH, and MeOH for 6 h, respectively.

Therefore, the above experiments demonstrate that organic solvents (for the type and concentration domains of organic solvents explored in this work) supported mild interactions with cells, especially in regard to DMSO. The minor improvement in solubility in the presence of low-level organic solvent (2%) led to an outstanding increase in yield, and no obvious compromises to cellular viability or integrity were observed from the FCM and protein leakage assay. It is reasonable to speculate that low-level organic solvents probably affected the membrane structure (including fluidity, tension, thickness, permeability, or water pore formation, and so on), which may be referred to as "micro-permeabilization", and the tiny change improved the transport pace of the myricetin. In contrast, the major improvement in solubility in the presence of 10% of organic solvent led to a major decrease in yield and no prominent compromises to cellular viability or integrity. The explanation for this finding might be the toxicity of organic solvents in the whole-cell reaction system.

2.3. Scaled-Up Synthesis of M7GA in a Fermenter System

To further evaluate the industrial potential of the whole-cell strategies, the one-pot bioconversion system was further scaled up to 1 L in a 3-L fermenter, in which some factors involved in the growth of *E. coli*, such as pH and dissolved oxygen, could be controlled. Based on the effects of organic solvents on cell integrity, viability, and stability, DMSO was selected as the co-solvent. A kinetic analysis of M7GA production over time is shown in Figure 4. The M7GA production dramatically increased within 12 h, and the production then incrementally leveled off over time. Finally, approximately 1322 ± 106 mg/L (roughly 2.7 mM) M7GA was produced after 48 h, and a 67% conversion yield was achieved.
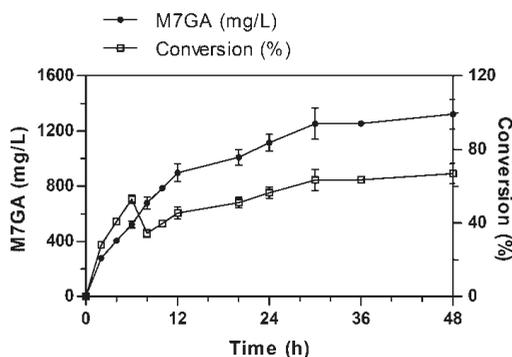
Figure 4. Time course for the production and conversion of myricetin-7-*O*-glucuronide (M7GA) obtained in a 3-L fermenter system containing 1 L of whole-cell catalyst. Two volumes of 20 mL myricetin stock solution (100 mM, dissolved in DMSO) was supplied in the reaction system at 0 h and 6 h, respectively.

3. Materials and Methods

3.1. Materials

Myricetin and M7GA were purchased from BioBioPha (Kunming, Yunnan, China). Organic solvents, acetone, DMF, DMSO, EtOH, MeOH, and all other reagents of analytical grade were purchased from local manufacturers (Beijing Chemical Factory, Beijing, China). The LIVE/DEAD BacLight Bacterial Viability Kit (# L7012), including solutions of 3.34 mM SYTO9 dye and 20 mM propidium iodide (PI) (both dissolved in DMSO), was purchased from Thermo Fisher Scientific (Waltham, MA, USA).

Transformed *E. coli* cells were grown on Luria–Bertani (LB) medium supplemented with antibiotics (100 mg/L ampicillin, 50 mg/L kanamycin and 25 mg/L chloramphenicol). The standard reaction system used for all shake-flask or fermenter production studies consisted of a modified M9 medium (pH 7.0, Na_2HPO_4 6.8 g/L; KH_2PO_4 2.7 g/L; NaCl 0.5 g/L; NH_4Cl 1.0 g/L; $MgSO_4$ 0.24 g/L; $CaCl_2$ 0.01 g/L) supplemented with glucose (10 g/L).

3.2. Preparation of a Certain Concentration of Organic Solvent

Myricetin was dissolved in the organic solvents such as acetone, DMF, DMSO, EtOH, and MeOH to form specific concentrations of myricetin stock solutions (100 mM, 40 mM, and 20 mM). Certain volumes of liquid solvents (acetone, DMF, DMSO, EtOH, and MeOH) were mixed with modified M9 medium to form final volumes of 4%, 20%, and 40% (*v/v*) (as 2× stocks used for the extracellular protein leakage assay).

3.3. General Procedure for the Whole-Cell Biocatalytic Conversion of Myricetin into M7GA

The engineered *E. coli* strain BPGUT was constructed as described in our previous work [19]. Transformed BPGUT cells were activated in LB medium, and the recombinant proteins were induced to express by the addition of 0.5 mM IPTG, and then incubated for an additional 15 h at 16 °C. To obtain the whole-cell biocatalyst, the bacterial cells were harvested by centrifugation, and the pellets were washed twice with modified M9 medium to remove the residual medium from the cell pellets. The washed pellets were resuspended in 10 mL of modified M9 medium, and the cell density was adjusted to an OD_{600} of 6 (12 g of cells per liter). Myricetin solid powder or stock solution was added to the above whole-cell biocatalyst system to form a working concentration of 2 mmol/L, which was then kept shaking at 30 °C for 6 h. The reaction mixture including cells was collected and lyophilized.

The dried mixture was dissolved in MeOH and analyzed by HPLC. All of the reactions were carried out in triplicate.

3.4. Determination of Myricetin and M7GA Solubility

Myricetin (0.64 mg) or M7GA (0.99 mg) solid powder or stock solution was added to 1 mL of modified M9 medium and then incubated in a shaker at 30 °C for 6 h. After centrifugation, the supernatant was analyzed by HPLC to determine the solubility of myricetin or M7GA. All assays were conducted in triplicate.

3.5. HPLC Analysis

A 20-μL sample of supernatant was used for analysis using a Thermo HPLC system (Thermo Fisher, Germering, Germany) equipped with a GL AQ-C18 reverse-phase column (GL Sciences Inc., Tokyo, Japan, 4.60×250 mm, 5 μm particle size). The mobile phase was composed of water with 0.05% TFA (A) and acetonitrile (B) with a gradient program. The analysis of myricetin and M7GA was conducted with a 20-min linear gradient of 20%–45% B. This was followed by isocratic flow with 100% B for 5 min to rinse off hydrophobic compounds and by isocratic development with 20% B for 5 min to equilibrate the column. The solubility of myricetin (μg/mL) or M7GA (μg/mL) and yields of M7GA (μg/mL) were calculated using the external standard curve method.

3.6. Flow Cytometry (FCM) Tests

The whole-cell biocatalyst cells or *E. coli* wild type cells were incubated in 1 mL of modified M9 medium containing a certain concentration of liquid organic solvents at 30 °C for 6 h. After being rinsed and appropriately diluted in 0.85% (w/v) NaCl solution to reach a final density of 1×10^7 cells/mL, 600 μL of the cell suspension was sampled and stained with 2 μL of the premixed dye solution (a 2:1 mixture of SYTO9 and PI stains generated by the LIVE/DEAD BacLight bacterial viability kit) for 15 min at room temperature in the dark before being subjected to FCM tests using a BD FACS melody flow cytometer (BD Biosciences, Franklin Lakes, NJ, USA). For FCM tests, the sample was excited at 488 nm, and the emission wavelengths were 515–545 nm for the SYTO9 stain and >650 nm for the PI stain, and data were analyzed using FlowJo software (Version 10, BD Biosciences, Franklin Lakes, NJ, USA, 2019).

3.7. Extracellular Protein Leakage Assay

The whole-cell biocatalyst was prepared as described above, and the final cell concentration was adjusted to $OD_{600} = 12$. The above cell suspension and different concentrations of organic solvents (2× stock) were then separately mixed at a volume ratio of 1:1. The mixture was then incubated at 30 °C for 6 h followed by a 16,000 g centrifugation for 20 min. The total content of extracellular proteins leaked from the treated cells in the supernatants was estimated using a protein quantitative kit (Bradford) (Bio-Rad, Hercules, CA, USA). The tests were conducted three separate times.

3.8. Scale-Up Synthesis of M7GA

To assess the scale-up whole-cell synthesis of M7GA with low-level organic solvent, two volumes of 20 mL of DMSO containing 100 mM myricetin were added at 0 h and 6 h. The scale-up reaction was carried out in a 3-L glass fermenter system (Biotech, Shanghai, China) containing 1 L of the 12 g/L whole-cell catalyst ($OD_{600} = 6.0$) at 30 °C and 350 rpm for 48 h. The pH was maintained at 6.8–7.0 throughout the process using ammonium hydroxide (28%) and phosphoric acid (4 M). Dissolved oxygen (DO) was maintained above 50% during the experiment. Samples were collected at given time points and lyophilized and dissolved in MeOH for quantification by HPLC.

Catalysts **2019**, 9, 970

4. Conclusions

In summary, the effects of five kinds of commonly used water-miscible organic solvents on the transformation of myricetin into M7GA were examined using engineered *E. coli* BPGUT as a multienzyme whole-cell biocatalyst. Higher M7GA yields were obtained from low-level organic solvents (2%, *v/v*), and in further correlating cell viability and bioconversion yields together with cell membrane integrity, we can conclude that the addition of low-level organic solvents accommodates whole *E. coli* cells. The fine physical or structural property changes to the cell membrane induced by low-level organic solvent, not the solubility of the substrate or product, were the main causes to improve the yields of the product. However, in spite of the encouraging results of the M7GA yields and cell viability, we have not gained enough information for cell membrane micro-permeabilization induced by low-level organic solvents and their corresponding molecular mechanisms. The application of organic solvents in multistep whole-cell catalysis should be a compromise between two factors: solvent properties and toxic effects. From the results of this research, the former aspect dealing with the solubility of substrates may be neglected. The majority of attention should be paid to the viability and integrity of the cells. Furthermore, whole-cell reactions involve a complex reaction system; other important factors, including the inhibition or denaturation of enzymes, the cellular tolerance to organic solvents, and the density of *E. coli* cells, may also have effects. Thus, the ways in which low-level organic solvents contribute to the biotransformation of a multienzyme whole-cell biocatalyst must be further investigated. This study highlights the relationships among solubility, cell viability, and performance of the whole-cell catalyst. To further increase stability, the catalytic capability of whole-cell catalysts for modification of other hydrophobic substrates with structure similar to myricetin can be tested.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4344/9/11/970/s1. Table S1: the yields of M7GA and the solubilities of myricetin or M7GA obtained in the modified M9 reaction system containing different concentrations of organic solvents; Table S2: the proportion of the live and dead cells as determined by FCM; Table S3: concentrations of proteins released from *E. coli* BPGUT cells after being exposed to organic solvents for 6 h; Table S4: concentrations of proteins released from *E. coli* BPGUT cells after being exposed to organic solvents for 12 h; Figure S1: correlations between the M7GA yield and the solubility of myricetin obtained in the modified M9 reaction system containing different concentrations of organic solvents; Figure S2: FCM images of *E. coli* BPUGT cells or wild type cells after being treated for 6 h at 30 °C in a modified M9 solution containing 2% (*v/v*), 5% (*v/v*), 10% (*v/v*), and 20% (*v/v*) organic solvents, or myricetin (2 mM) and M7GA (2 mM); Figure S3: viability of *E. coli* BPUGT cells after being treated for 12 h at 30 °C in a modified M9 solution containing 2% (*v/v*) organic solvents; Figure S4: concentrations of proteins released from *E. coli* BPGUT cells after the cells were exposed to 2% (*v/v*), 5% (*v/v*), 10% (*v/v*) and 20% (*v/v*) organic solvents for 12 h; Figure S5: time course for the production of M7GA obtained in a 100-mL shake-flask system containing 10 mL of whole-cell catalyst.

Author Contributions: Conceptualization, Y.Y. and W.W.; data curation, Y.Y.; formal analysis, Y.Y.; funding acquisition, W.W.; investigation, Y.Y.; methodology, Y.Y., and M.-Z.L.; project administration, W.W.; resources, W.W.; software, Y.Y. and C.-K.L.; validation, Y.Y., and Y.-S.C.; writing—original draft preparation, Y.Y., and M.-Z.L.; writing—review and editing, Y.Y. and W.W.

Funding: This research was funded by the CAMS Innovation Fund for Medical Science (CAMS-2016-I2M-3-012, CAMS-2017-I2M-1-011) and National Mega-project for Innovative Drugs (2018ZX09711001-006-001).

Acknowledgments: We are grateful to Liu Yang (Institute of Materia Medica, Chinese Academy of Medical Sciences) for providing technical assistance.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

DMF	*N,N*-dimethylformamide
DMSO	Dimethyl sulfoxide
EtOH	Ethanol
FCM	Flow cytometry
MeOH	Methanol
M7GA	Myricetin-7-*O*-glucuronide
PI	Propidium iodide

References

1. Lin, B.; Tao, Y. Whole-cell biocatalysts by design. *Microb. Cell Fact.* **2017**, *16*, 106. [CrossRef] [PubMed]
2. Wu, S.; Li, Z. Whole-cell cascade biotransformations for one-pot multistep organic synthesis. *ChemCatChem* **2018**, *10*, 2164–2178. [CrossRef]
3. Ricca, E.; Brucher, B.; Schrittwieser, J.H. Multi-enzymatic cascade reactions: Overview and perspectives. *Adv. Synth. Catal.* **2011**, *353*, 2239–2262. [CrossRef]
4. Shi, J.; Wu, Y.; Zhang, S.; Tian, Y.; Yang, D.; Jiang, Z. Bioinspired construction of multi-enzyme catalytic systems. *Chem. Soc. Rev.* **2018**, *47*, 4295–4313. [CrossRef] [PubMed]
5. De Carvalho, C.C. Whole cell biocatalysts: Essential workers from nature to the industry. *Microb. Biotechnol.* **2017**, *10*, 250–263. [CrossRef] [PubMed]
6. Kadisch, M.; Willrodt, C.; Hillen, M.; Bühler, B.; Schmid, A. Maximizing the stability of metabolic engineering-derived whole-cell biocatalysts. *Biotechnol. J.* **2017**, *12*, 1600170. [CrossRef] [PubMed]
7. Wachetmeister, J.; Rother, D. Recent advances in whole cell boicatalsysis. *Curr. Opin. Biotech.* **2016**, *42*, 169–177. [CrossRef]
8. Rudroff, F. Whole-cell based synthetic enzyme cascades-light and shadow of a promising technology. *Curr. Opin. Chem. Biol.* **2019**, *49*, 84–90. [CrossRef]
9. McAuliffe, J.C. Industrial enzymes and biocatalysis. In *Handbook of Industrial Chemistry and Biotechnology*; Kent, J.A., Ed.; Springer: Boston, MA, USA, 2012; pp. 1183–1227. ISBN 978-1-4614-4259-2.
10. Krauser, S.; Weyler, C.; Blaß, L.K.; Heinzle, E. Directed multistep biocatalysis using tailored permeabilized cells. *Adv. Biochem. Eng. Biotechnol.* **2013**, *137*, 185–234. [CrossRef]
11. Zheng, G.W.; Pan, J.; Yu, H.L.; Ngo-Thi, M.T.; Li, C.X.; Xu, J.H. An efficient bioprocess for enzymatic production of L-menthol with high ratio of substrate to catalyst using whole cells of recombinant *E. coli*. *J. Biotechnol.* **2010**, *150*, 108–114. [CrossRef]
12. Li, X.F.; Yuan, T.; Xu, H.; Xin, X.; Zhao, G.; Wu, H.; Xiao, X. Whole-cell catalytic synthesis of puerarin monoesters and analysis of their antioxidant activities. *J. Agric. Food Chem.* **2019**, *67*, 299–307. [CrossRef] [PubMed]
13. Gurtovenko, A.A.; Anwar, J. Modulating the structure and properties of cell membranes: The molecular mechanism of action of dimethyl sulfoxide. *J. Phys. Chem. B* **2007**, *111*, 10453–10460. [CrossRef] [PubMed]
14. Kim, M.J.; Seo, M.J.; Shin, K.C.; Oh, D.K. Production of 10*S*-hydroxy-8(*E*)-octadecenoic acid from oleic acid by whole recombinant *Escherichia coli* cells expressing 10*S*-dioxygenase from *Nostoc punctiforme* PCC 73102 with the aid of a chaperone. *Biotechnol. Lett.* **2017**, *39*, 133–139. [CrossRef] [PubMed]
15. Ma, L.; Du, L.; Chen, H.; Sun, Y.; Huang, S.; Zheng, X.; Kim, E.S.; Li, S. Reconstitution of the in vitro activity of the cyclosporine-specific P450 hydroxylase from *Sebekia benihana* and development of a heterologous whole-cell biotransformation system. *Appl. Environ. Microbiol.* **2015**, *81*, 6268–6275. [CrossRef]
16. Cui, J.D.; Jia, S.R.; Sun, A.Y. Influence of amino acids, organic solvents and surfactants for phenylalanine ammonia lyase activity in recombinant *Escherichia coli*. *Lett. Appl. Microbiol.* **2008**, *46*, 631–635. [CrossRef]
17. Janardhan Garikipati, S.V.; Peeples, T.L. Solvent resistance pumps of *Pseudomonas putida* S12: Applications in 1-naphthol production and biocatalyst engineering. *J. Biotechnol.* **2015**, *210*, 91–99. [CrossRef]
18. Honda, K.; Inoue, M.; Ono, T.; Okano, K.; Dekishima, Y.; Kawabata, H. Improvement of operational stability of *Ogataea minuta* carbonyl reductase for chiral alcohol production. *J. Biosci. Bioeng.* **2017**, *123*, 673–678. [CrossRef]
19. Yang, Y.; Wang, H.M.; Tong, Y.F.; Liu, M.Z.; Cheng, K.D.; Wu, S.; Wang, W. Systems metabolic engineering of *Escherichia coli* to enhance the production of flavonoid glucuronides. *RSC Adv.* **2016**, *6*, 33622–33630. [CrossRef]
20. Semwal, D.K.; Semwal, R.B.; Combrinck, S.; Viljoen, A. Myricetin: A dietary molecule with diverse biological activities. *Nutrients* **2016**, *8*, 90. [CrossRef]
21. Guo, C.; Zhu, Y.; Weng, Y.; Wang, S.; Guan, Y.; Wei, G.; Yin, Y.; Xi, M.; Wen, A. Therapeutic time window and underlying therapeutic mechanism of breviscapine injection against cerebral ischemia/reperfusion injury in rats. *J. Ethnopharmacol.* **2014**, *151*, 660–666. [CrossRef]
22. Ishisaka, A.; Kawabata, K.; Miki, S.; Shiba, Y.; Minekawa, S.; Nishikawa, T.; Mukai, R.; Terao, J.; Kawai, Y. Mitochondrial dysfunction leads to deconjugation of quercetin glucuronides in inflammatory macrophages. *PLoS ONE* **2013**, *8*, e80843. [CrossRef] [PubMed]

23. Spencer, J.P.; Abd-el-Mohsen, M.M.; Rice-Evans, C. Cellular uptake and metabolism of flavonoids and their metabolites: Implications for their bioactivity. *Arch. Biochem. Biophys.* **2004**, *423*, 148–161. [CrossRef] [PubMed]

24. Wong, C.C.; Barron, D.; Orfila, C.; Dionisi, F.; Krajcsi, P.; Williamson, G. Interaction of hydroxycinnamic acids and their conjugates with organic anion transporters and ATP-binding cassette transporters. *Mol. Nutr. Food Res.* **2011**, *55*, 979–988. [CrossRef] [PubMed]

25. Notman, R.; Noro, M.; O´Malley, B.; Anwar, J. Molecular basis for dimethylsulfoxide (DMSO) action on lipid membranes. *J. Am. Chem. Soc.* **2006**, *128*, 13982–13983. [CrossRef] [PubMed]

26. Sandoval, N.R.; Papoutsakis, E.T. Engineering membrane and cell-wall programs for tolerance to toxic chemicals: Beyond solo genes. *Curr. Opin. Microbiol.* **2016**, *33*, 56–66. [CrossRef] [PubMed]

27. Marquês, J.T.; Viana, A.S.; De Almeida, R.F. Ethanol effects on binary and ternary supported lipid bilayers with gel/fluid domains and lipid rafts. *Biochim. Biophys. Acta* **2011**, *1808*, 405–414. [CrossRef]

28. Weber, F.J.; de Bont, J.A. Adaptation mechanisms of microorganisms to the toxic effects of organic solvents on membranes. *Biochim. Biophys. Acta* **1996**, *1286*, 225–245. [CrossRef]

An Innovative Biocatalyst for Continuous 2G Ethanol Production from Xylo-Oligomers by *Saccharomyces cerevisiae* through Simultaneous Hydrolysis, Isomerization, and Fermentation (SHIF)

Thais S. Milessi-Esteves [1,†], Felipe A.S. Corradini [2,†], Willian Kopp [1], Teresa C. Zangirolami [1,2], Paulo W. Tardioli [1,2], Roberto C. Giordano [1,2] and Raquel L.C. Giordano [1,2,*]

1   Department of Chemical Engineering, Federal University of São Carlos, Rodovia Washington Luiz, km 235, 13565-905, São Carlos, SP, Brazil; thais.milessi@gmail.com (T.S.M.-E.); willkopp@gmail.com (W.K.); teresacz@ufscar.br (T.C.Z.); pwtardioli@ufscar.br (P.W.T.); roberto@ufscar.br (R.C.G.)
2   Graduate Program of Chemical Engineering, Federal University of São Carlos (PPGEQ-UFSCar), Rodovia Washington Luiz, km 235, 13565-905, São Carlos, SP, Brazil; eq.felipe.silva@gmail.com
*   Correspondence: raquel@ufscar.br; Tel.: +55-16-3351-8264
†   Both authors contributed equally to this work

Abstract: Many approaches have been considered aimed at ethanol production from the hemicellulosic fraction of biomass. However, the industrial implementation of this process has been hindered by some bottlenecks, one of the most important being the ease of contamination of the bioreactor by bacteria that metabolize xylose. This work focuses on overcoming this problem through the fermentation of xylulose (the xylose isomer) by native *Saccharomyces cerevisiae* using xylo-oligomers as substrate. A new concept of biocatalyst is proposed, containing xylanases and xylose isomerase (XI) covalently immobilized on chitosan, and co-encapsulated with industrial baker's yeast in Ca-alginate gel spherical particles. Xylo-oligomers are hydrolyzed, xylose is isomerized, and finally xylulose is fermented to ethanol, all taking place simultaneously, in a process called simultaneous hydrolysis, isomerization, and fermentation (SHIF). Among several tested xylanases, Multifect CX XL A03139 was selected to compose the biocatalyst bead. Influences of pH, Ca^{2+}, and Mg^{2+} concentrations on the isomerization step were assessed. Experiments of SHIF using birchwood xylan resulted in an ethanol yield of 0.39 g/g, (76% of the theoretical), selectivity of 3.12 $g_{ethanol}/g_{xylitol}$, and ethanol productivity of 0.26 g/L/h.

Keywords: 2G ethanol; hemicellulose usage; *S. cerevisiae*; enzyme immobilization; cell immobilization; SHIF

1. Introduction

Biofuels will have a significant role in the energetic matrix of the low-carbon economy, helping to meet the goals established at Conference of the Parties (COP 21) [1,2]. Among biofuels, bioethanol production from lignocellulosic materials has been intensively studied once it was shown that these byproducts had high availability, had a low cost, and did not compete with the production of food [3]. Lignocellulosic raw materials are mainly composed of cellulose and hemicellulose (up to 70%), which are polysaccharides that, after a hydrolysis step, generate fermentable sugars, mostly xylose from hemicellulose and glucose from cellulose [4]. The use of these two polysaccharides is important for the economic feasibility of the biofuel production process.

Some microorganisms that naturally ferment pentoses to ethanol have been tested for industrial use, such as *Scheffersomyces stipitis* and *Pachysolen tannophilus* [5,6]. However, these microorganisms

have a low tolerance to ethanol and slow fermentation rates and are inhibited by compounds generated during the biomass pretreatment step, such as furfural [7].

Saccharomyces cerevisiae is the most common microorganism used for ethanol production from hexoses, due to its high rate of fermentation and superior ethanol yield. In addition, this yeast exhibits unbeatable tolerance to ethanol, to inhibitors, and to high concentrations of sugar [8,9]. However, in its wild form *S. cerevisiae* is unable to efficiently metabolize D-xylose.

The genetic modification of *S. cerevisiae* aimed at xylose fermentation has been extensively studied [8,10–12]. However, the low specific growth rate, high xylitol production, reduced yeast tolerance, and possible genetic instability are still hindrances for the application of recombinant strains on an industrial scale [7].

In spite of the inability of *S. cerevisiae* to metabolize xylose, it is capable of fermenting its isomer, xylulose, to ethanol. Hence, an alternative for the utilization of the hemicellulose fraction for bioethanol production would be to isomerize xylose to xylulose ex vivo, followed by fermentation by *S. cerevisiae* [7]. The enzyme xylose isomerase (XI) (EC 5.3.1.5) is widely used in the industry for the production of fructose syrup from corn starch and also catalyzes the reversible isomerization of xylose to xylulose [13]. Although the xylose/xylulose chemical equilibrium is unfavorable (3.5:1 at 60 °C) [14], the reaction can be displaced by the simultaneous isomerization and fermentation (SIF) process, where the continuous conversion of xylulose to ethanol might allow the complete depletion of the available xylose [15].

The use of catalysts with immobilized enzymes may be crucial for the application of multi-enzymatic processes on an industrial scale. This approach allows the continuous operation of the reactor and facilitates the product recovery as well as the use of high loads of cells and enzymes [16]. The production of an active and stable enzyme derivative using a non-expensive support is also an important issue in enzyme immobilization [17]. The literature reports successful applications of immobilized enzymes on an industrial scale [18,19] and immobilized XI is one of the most successful and established examples [13]. Silva et al. [7] developed a biocatalyst containing chitosan-immobilized XI, co-immobilized with *S. cerevisiae* in calcium alginate gel. Calcium alginate gel was chosen for being a natural polymer widely studied as a support for the immobilization of viable cells [20]. However, this system showed to be susceptible to contamination by xylose-consuming bacteria. High concentrations of xylose in the medium disfavored the *S. cerevisiae* population, due to its low uptake rates of xylulose.

An alternative to tackling the contamination problem is to use a cultivation medium containing non-readily fermentable substrates, such as xylo-oligomers obtained by the solubilization of hemicellulose under mild conditions [21]. The hetero-polysaccharides that compose hemicellulose are polymers with about 100 units of monomers, mainly xylose, and their solubility depends on the number of monomeric units in the chain [22]. Thus, the extraction of xylan in the form of large oligomers must be carried out under conditions that allow a sufficient number of glycosidic bonds to be broken, so that soluble polymers with lower molecular weight (xylo-oligomers) are released.

Xylanases (β-1,4-D-xylanase) are enzymes that catalyze the hydrolysis of the glycosidic bonds between xylose units. The enzymatic complex is commonly composed of endoxylanases, exoxylanases, β-D-xylosidases as well as accessory enzymes such as glucuronidase and arabinofuranosidase that act on the ramifications of the xylan chain [23]. The addition of these enzymes to the biocatalyst proposed by Silva et al. [7] would allow the feeding of xylo-oligomers to the bioreactor. This substrate might decrease the probability of contamination during the operation of the bioreactor for long periods, which are typical in industry. Preliminary results showed the technical viability of this process [24].

The development of viable processes to increase ethanol yields from lignocellulosic materials is crucial, despite the challenges that still remain for the production of 2G ethanol from xylose. Considering the higher production cost of 2G ethanol (compared to 1G ethanol), the use of pentoses as raw material could make its production more profitable and might overcome the costs of 2G ethanol extra steps [25]. The integration of several biocatalytic transformations in a multi-enzymatic cascade system is particularly appealing to the development of cleaner and more efficient biochemical processes.

Multi-enzymatic cascade reactions offer advantages such as lower demand of time, reduced costs, easier recovery of products, completion of reversible reactions as well as concentrations of inhibitory compounds restrained to a minimum [26].

In this context, the simultaneous hydrolysis, isomerization, and fermentation (SHIF) process stands out for 2G ethanol production since unmodified *S. cerevisiae* remains the preferred microorganism in industry, due to its robustness, high ethanol tolerance, and production rates. The use of wild strains to produce ethanol from xylose is an important issue in countries like Brazil, where biosafety regulations are strict [15]. In addition, one advantage of this approach is that XI, along with amylases and proteases, is among the most widely and cheaply available commercial enzymes [27]. The present work reports the results of using this new biocatalyst for the simultaneous hydrolysis, isomerization, and fermentation of xylan derived from the hemicellulose fraction of biomass, aimed at the production of ethanol (Figure 1).
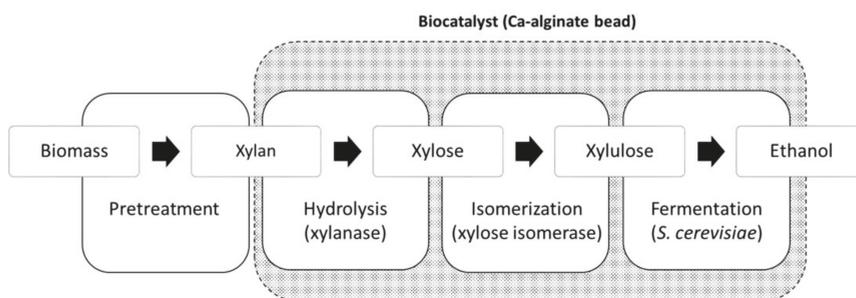

Figure 1. Xylan biomass simultaneous hydrolysis, isomerization, and fermentation (SHIF). Biocatalyst composed of xylanases, xylose isomerase, and co-immobilized *S. cerevisiae*.

2. Results

2.1. Application of the New Biocatalyst in the SHIF Process

The biocatalyst is designed for the industrial production of second-generation ethanol in continuous, fixed-bed reactors through long-term operation, by applying simultaneous hydrolysis, isomerization, and fermentation (SHIF) of the hemicellulosic fraction of biomass. First, xylanase and xylose isomerase were covalently immobilized on chitosan. The obtained XI derivative presented an activity of 252.5 ± 1.6 IU/g (immobilization yield of 93% and recovered activity of 91%), whereas the Accellerase derivative exhibited 346.3 ± 9.2 IU/g (immobilization yield of 54% and recovered activity of 12%). Both derivatives were co-encapsulated with *S. cerevisiae* in Ca-alginate gel and this biocatalyst was used to produce ethanol from commercial birchwood xylan.

Birchwood xylan, which is a heteropolymer composed of long chains, was first hydrolyzed to smaller xylo-oligomers by the action of recombinant endoxylanase of *Bacillus subtilis* (XynA) in order to increase the concentration of xylo-oligomers with smaller chains that may diffuse into the catalyst beads [28]. This step was carried out to make xylan more similar to lignocellulosic hydrolysates obtained from the pretreatment of biomass (data not shown). The composition of the substrate after xylan pre-hydrolysis is shown in Table 1.

As expected, there was no xylose production since XynA is a strict endoxylanase [28]. The solubilized fraction corresponded to 67% (w/w) of the offered xylan. According to Gray et al. [22], the solubility of the xylan oligomers depends on the degree of polymerization of each compound. Under the used conditions, 33% (w/w) of birchwood xylan is insoluble. Therefore, the substrate obtained for the SHIF process had 73 g/L of xylo-oligomers.

Table 1. Characterization of SHIF substrate: birchwood xylan, 108 g/L after 24 h hydrolysis by endoxylanase XynA (150 IU/g_{xylan}) at 50 °C and pH 5.6.

Component	Hydrolyzed Birchwood Xylan	
	(g/L)	(%)
Xylohexaose (X6) or bigger	62.3	57.4
Xylopentaose (X5)	0.0	0.0
Xylotetraose (X4)	1.2	1.1
Xylotriose (X3)	2.5	2.2
Xylobiose (X2)	7.1	6.5
Xylose	0.0	0.0
Total Soluble	73.0	67.2
Total Insoluble	35.6	32.8
Total	108.6	100.0

For the SHIF assays, the offered enzyme activity in the reactor was 1.7×10^4 IU/$L_{reactor}$ for xylanase (Accellerase XY) and 3.7×10^4 IU/$L_{reactor}$ of xylose isomerase. Accellerase XY was used due to the presence of β-xylosidase, which is necessary to xylose formation. Commercial baker's yeast (Itaiquara®) concentration was 50 $g_{dry\ mass}$/$L_{reactor}$ at the beginning of the SHIF assays. Results in Figure 2 show the production of ethanol through SHIF using the developed biocatalyst. Ethanol production was higher compared to xylitol, presenting a selectivity of 2.61 (2.2 g/L ethanol and 0.84 g/L xylitol). Ethanol productivity of 0.092 g/L/h and yield ($Y_{P/S}$) of 0.160 $g_{ethanol}$/$g_{potentialxylose}$ (32% of theoretical, calculated on the basis of potential xylose in the xylan) were achieved at the end of the SHIF run.
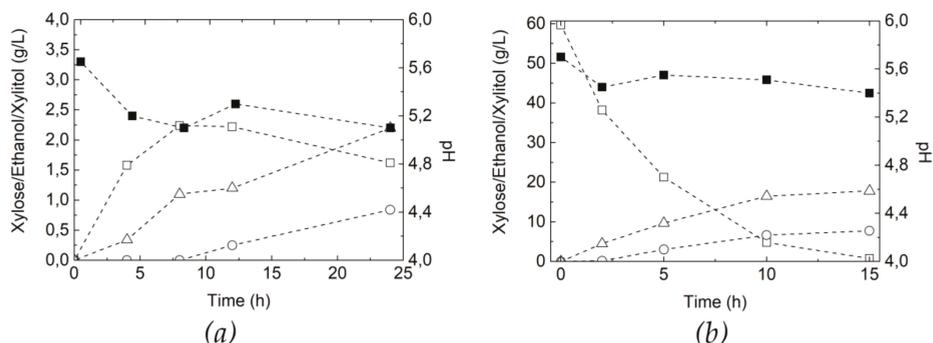

Figure 2. (**a**) SHIF of pre-hydrolyzed birchwood xylan using a biocatalyst containing 5 w % of Accellerase XY, 15 w % of XI, and 10 w % dry mass yeast (1.7×10^4 IU/L, 3.7×10^4 IU/L, and 50 g/L, respectively), pH 5.6, 150 rpm stirring, and 32 °C; (**b**) control experiment using xylose as substrate (SIF): pH (■), xylose (□), ethanol (△), and xylitol (○).

Figure 2 shows that ethanol is produced from xylo-oligomers. The slower velocity of SHIF compared to the control experiment (where the substrate is xylose, resulting in a simultaneous isomerization, and fermentation (SIF) process) was expected since xylose concentration in SHIF depends on the velocity of hydrolysis of the xylo-oligomers. However, a decrease in the xylose consumption was observed after eight hours of SHIF (Figure 2a), which indicates that the xylose isomerization was impaired. There are two possible reasons for this occurrence: XI is not catalyzing isomerization of xylose in the required velocity; or the yeast is not consuming the generated xylulose, which would be accumulating and consequently stopping the isomerization due to the xylose:xylulose equilibrium ratio. Since the yeast remained viable (initial and final cell viability unchanged: 96%) and there was no accumulation of xylulose in the medium, the accumulation of xylose seemed to be related to the isomerization step.

In the control experiment, using only xylose as substrate, XI catalyzed the isomerization reaction effectively, almost until depletion of the pentose (60 g/L of xylose were consumed in 12 h, a productivity of 1.2 g/L/h, with $Y_{P/S}$ of 0.303 g/g and selectivity of 2.3 with respect to xylitol). Thus, an increase of the isomerization velocity must be sought. Some of the possible causes for the hindrance of the isomerization step were then investigated. High calcium concentrations are known to inhibit XI action [29], and the hydrolysis reaction could be demanding a higher release of the calcium ion (increased hydrolysis of $CaCO_3$ to control the pH). However, even with a higher release of calcium, the results shown in Figure 2 indicate that the hydrolysis step led to a drop of pH to approximately 5.1. XI shows maximum activity at pH 8, being highly sensitive to a drop of pH to this range [15]. On the other hand, it is known that the magnesium ion is an important cofactor for XI, as an activator of this enzyme [30]. The influence of pH and Ca^{2+} and Mg^{2+} ions in the isomerization step was then investigated.

2.2. Influence of pH, Ca^{2+}, and Mg^{2+} on XI Activity

In order to investigate the influence of calcium and magnesium ions on XI, the activity of this enzyme to catalyze fructose–glucose isomerization was measured at different pHs (5.0 to 8.0) in the presence of different concentrations of Ca^{2+} and Mg^{2+} ions. The standard medium for assessing activity was 2 M fructose, pH 8.0 (50 mM tris-maleate buffer) supplemented with 50 mM of $MgSO_4$ and 2.5 mM of $CoCl_2$ at 60 °C. Both Co^{2+} and Mg^{2+} are essential for the activity of XI, however they play differentiated roles. Mg^{2+} is superior to Co^{2+} as an activator, while the latter is responsible for the stabilization of the enzyme and maintenance of its conformation, especially the quaternary structure [31]. Table 2 shows the measured activities of XI in each condition studied, referred to the test performed with the standard medium at pH 8.0 as 100%.

Table 2. Influence of Mg^{2+} and Ca^{2+} on XI activity at different pHs (isomerization of fructose 2M, 60 °C). The activity measured in standard medium (pH 8.0, 50 mM $MgSO_4$, 2.5 mM $CoCl_2$) was taken as 100%.

	Medium	pH 8.0		pH 7.0		pH 6.0		pH 5.0	
		IU/mL	%	IU/mL	%	IU/mL	%	IU/mL	%
1	Standard	1896 ± 68	100.0	1590 ± 23	83.9	1466 ± 12	77.4	442 ± 21	23.4
2	$2\times Mg^{2+}$	2041 ± 31	107.7	1890 ± 7	99.7	1820 ± 3	96.9	940 ± 24	49.6
3	$5\times Mg^{2+'}$	2202 ± 67	116.2	2287 ± 43	120.6	2333 ± 12	123.1	1166 ± 35	61.5
4	Standard + Ca^{2+}*	1454 ± 97	76.7	1312 ± 20	69.2	572 ± 85	30.2	380 ± 45	20.1
5	$2\times Ca^{2+}$	1366 ± 72	72.1	944 ± 69	49.8	503 ± 94	26.6	352 ± 45	18.6
6	$5\times Ca^{2+}$	1063 ± 56	56.1	797 ± 56	42.1	356 ± 31	18.8	255 ± 44	13.5
7	$Ca^{2+}* + 2\times Mg^{2+}$	1542 ± 81	81.3	1673 ± 27	88.3	1073 ± 56	56.6	803 ± 20	42.4
8	$Ca^{2+}* + 5\times Mg^{2+}$	1672 ± 88	88.2	2236 ± 39	117.9	1250 ± 16	65.9	966 ± 17	51.0

* $[CaCl_2]$ = 4 g/L (same as used during SHIF supplementation).

The Ca^{2+} ion proved to be an inhibitor of this enzyme, since a significant decrease in the XI activity occurred when the Ca^{2+} concentration in the medium increased. The Mg^{2+} ion, in turn, was able to activate the enzyme, increasing its activity in 16.2% at pH 8.0 and 163% at pH 5.0, both in calcium-free medium. This cofactor was still able to bypass the inhibition caused by calcium, since it reactivated the enzyme in the presence of this ion, increasing its catalytic activity in all studied pHs.

Xylose isomerization catalyzed by XI is initiated by opening the sugar ring, followed by isomerization through the exchange of hydride and finally stabilization of the product by ring closure [32]. Although there is no relationship between the presence of magnesium and the ring opening step, this cation is essential for the isomerization [33]. According to Kasumi et al. [34], the reaction mechanism demands the formation of a binary divalent enzyme–cation complex, since the substrate will bind only to the active site of this complex. Xylose isomerase has two active sites, each containing two divalent cations [33]. Thus, the presence of higher concentrations of magnesium in the reaction medium, improving the probability of the presence of this ion in the active sites of the enzyme, would increase the rates of the isomerization reaction.

Although calcium is a divalent cation and belongs to the same family as magnesium (same configuration in the valence layer) in the periodic table, Ca^{2+} has a larger ionic radius than Mg^{2+}. This fact could be the reason why Mg^{2+} is an activator of the enzyme while Ca^{2+} inhibits XI, that is, the difference in their atomic radii would cause a different interaction with the active site of the enzyme.

In addition to the significant influence of Ca^{2+} and Mg^{2+}, XI showed great sensitivity to pH, losing activity significantly at pH 5.0. The sensitivity of XI to pH was previously observed by Milessi et al. [15], who emphasized the importance of pH control during the simultaneous isomerization and fermentation (SIF) of xylose. However, the activation provided by Mg^{2+} is potentiated at lower pHs. When Ca^{2+} (4 g/L) was added, the magnesium ion was able to recover XI activity more effectively at pH 5.0 than at pH 8.0 (Table 2, media 4 and 7).

Data presented in Table 2 also prove that XI activity was greatly reduced at the SHIF pH range (5.0 to 6.0). The quaternary structure of this enzyme is composed of four subunits that are delicately folded and associated with noncovalent links and without interchain disulfide bonds [31,32]. At pH 8.0, the enzyme structure is composed of all four subunits combined, resulting in the maximum catalytic activity. However, as the pH of the reaction medium lowers, the enzyme is more likely to suffer structure distortions, unfolding and dissociating its tetrameric structure. The presence of Ca^{2+} ions at the low pH of the reaction medium results in a combined effect, acting both on the 3D structure and on the active site. Therefore, the performance of the SHIF process will certainly benefit from a pH control system (pH 5.6). Unfortunately, the substitution of calcium chloride by magnesium chloride in the coagulation solution during sodium alginate gelation was not possible since the resulting beads were not stable. For this reason, SHIF experiments with pH control and excess of magnesium were run, in order to minimize the inhibition of XI caused by the hydrolysis reactions, which release acids from the structures of the xylo-oligomers [35].

2.3. Xylanase Selection

Xylan hydrolysis has to occur efficiently to enable the cascade SHIF process. Hydrolysis cannot be the rate-determining step of these reactions in series: due to the unfavorable equilibrium of xylose to xylulose isomerization, the supply of xylose must not control the reaction [7]. Considering that the composition of each xylanase complex influences the hydrolysis efficiency, different xylanases were evaluated with the purpose of selecting the most efficient for the SHIF process.

The xylanase family is strongly related to the profile of products generated in the process [36]. A xylanase capable of depolymerizing xylan into xylose efficiently is required to ensure that the SHIF process will proceed as expected. Thus, besides Accellerase XY A03304, used in previous SHIF tests, two additional xylanases were evaluated: recombinant *B. subtilis* endoxylanase (XynA) and Multifect CX XL A03139.Hydrolysis profiles and xylooligosaccharide (XOS) composition are reported in Table 3 and Figure 3, respectively.

Table 3. Composition of xylooligosaccharides (XOS) after enzymatic hydrolysis ofbirchwood xylan (25.4 g/L, 3.8 IU/mL, 24 h of reaction at 50 °C, pH 5.6).

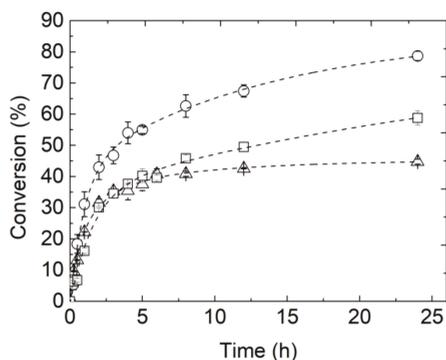| Xylanase	X6 (g/L)	X5 (g/L)	X4 (g/L)	X3 (g/L)	X2 (g/L)	Xylose (g/L)	Conversion (%)
XynA	0	0.20	1.22	4.43	3.49	0.00	44.8
Multifect	0	0.18	0.31	1.85	6.48	13.24	78.7
Accellerase	0	0.00	0.34	1.51	6.45	8.08	58.8

Figure 3. Hydrolysis of birchwood xylan soluble fraction (25,4 g/L) by the studied xylanases (150 IU/g_{xylan} = 3810 IU/L) at 50 °C, 24 h, pH 5.6. XynA (Δ); Multifect (○), and Accellerase (□). Bars are standard errors of triplicates.

Figure 3 shows that Multifect stands out, with a xylan conversion of 78.7%. Moreover, the higher xylose concentration achieved with this enzyme at the end of the experiments indicates that this xylanase complex has a more stable β-xylosidase enzyme, responsible for catalyzing the hydrolysis of xylobiose, the essential final step for the complete xylan hydrolysis. Hence, this xylanase seems to be the most suitable for the production of xylose in the SHIF process, among the studied enzymes.

Accellerase has the highest enzymatic activity under standard conditions. However, in long-term reaction it was able to convert only 58.8% of the available xylan. Several factors may have contributed to Accellerase's inferior performance, such as the amount of each enzyme in the complex, thermal inactivation, substrate affinity, and inhibitory effects.

XynA was already known to have a strictly endoxylanase action, lacking β-xylosidase activity and consequently not producing xylose when hydrolyzing xylan [28]. Accordingly, it presented the lowest conversion (44.8%), probably due to the absence of debranching enzymes.

None of the tested xylanases reached 100% of xylan conversion. Indeed, the incapacity of xylanases to completely hydrolyze xylan has been previously reported. Akpinar et al. [37] observed a yield of 13.8% for tobacco xylan using *Aspergillus niger* xylanase (200 IU/g) at 50 °C after 24 h. Aragon et al. [38] achieved 13% of conversion in the hydrolysis of birch xylan (18 g/L) using *Aspergillus versicolor* endoxylanase immobilized on agarose-glyoxyl at 25 °C and pH 5.0. In fact, since xylan is not a linear polymer of pure xylose, its complete depolymerization requires the use of a varied pool of enzymes [21,23,39]. In this context, the xylanase Multifect CX XL A03139 was selected to be co-immobilized with XI and *S. cerevisia*e in the SHIF process.

2.4. SHIF Assay with pH Control and Excess of Mg^{2+}

In order to overcome the possible inhibition of Ca^{2+} in XI activity, a SHIF assay was performed with pH control and excess of magnesium. Beads without $CaCO_3$ in its composition were prepared, since this salt is only necessary to sustain the pH at the desired range. Moreover, an isomerization free of $CaCO_3$ would contribute to reduce the undesired Ca^{2+} effects. After all these modifications, the obtained derivative of xylanase Multifect presented 330.2 ± 8.1 IU/g (immobilization yield of 96% and recovered activity of 77%). For the SHIF experiment, the medium supplemented with 100 mM $MgSO_4$ (24.6 g/L) and 4 g/L of $CaCl_2$ (to maintain the integrity of the beads) was added together with the beads to the pH-stat stirred reactor. It is important to note that the moderate agitation used during the process did not affect the integrity of the beads. According to Rahim et al. [40], damages to Ca-alginate immobilized biocatalysts due to stirring are usually observed above 200 rpm. In fact, Carvalho et al. [41], in experiments carried out at 300 rpm, noticed a 30% reduction in the size of Ca-alginate beads during experiments with immobilized *Candida guilliermondii*. Accordingly, in the

present work, an agitation of 150 rpm was employed and the structural characteristics of the biocatalyst beads were preserved.

The obtained results, showed in Figure 4, indicated a higher ethanol production in SHIF using pH control and excess of Mg^{2+} (3.1 g/L of ethanol) compared to the value of 2.2 g/L, which was achieved under the original SHIF conditions (Figure 2). Ethanol yield (0.39 g/g, 76% of the theoretical), selectivity (3.12), and productivity (0.26 g/L/h) were also improved.
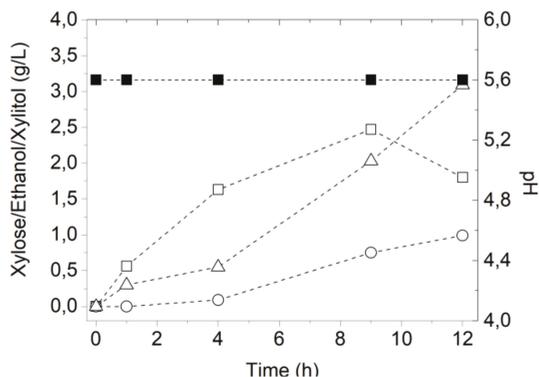

Figure 4. SHIF of previously hydrolyzed birchwood xylan with pH control and excess of Mg^{2+} using immobilized biocatalyst containing 5 w % Multifect xylanase derivative, 15 w % XI derivative, and 10 w % yeast (1.7×10^4 IU/L, 3.7×10^4 IU/L, 50 g/L), initial pH 5.6, 150 rpm stirring and 32 °C; pH (■), xylose (□), ethanol (Δ), and xylitol (○).

To the best of our knowledge, these are the highest yield and productivity reported in the literature for ethanol production from xylan through the simultaneous hydrolysis, isomerization, and fermentation (SHIF) process. Only a few works have studied ethanol production from pentoses using ex vivo isomerization and native *S. cerevisiae*, due to the differences in optimal pH and temperature ranges for each step (and none of them with co-immobilized enzymes/cells). The inclusion of the xylan hydrolysis step should not be a problem in relation to the temperature and pH of the process. Xylanases have the highest catalytic pH and temperature at approximately 5.5 and 50 °C, respectively, whereas XI optimal conditions are pH 7.0–8.0 and 70 °C [7]. Alcoholic fermentation, on the other hand, operates at pH 5.0 and 30 °C. Due to these facts, the process integration for ethanol production is still a challenge. Nakata et al. [42] studied ethanol production of hot-compressed water pretreated Japanese beech using soluble β-xylosidase, XI, and *S. cerevisiae*. It should be stressed that in this work the absence of exo- and endoxylanases would be a restraint to the saccharification step. The best results reported (0.62 g/L of ethanol, corresponding to 13% of theoretical yield) were achieved at pH 5.0, 30 °C after 100 h. Hence, the immobilized biocatalyst containing enzymes and yeast co-encapsulated reported in the present work was significantly more efficient than using enzymes and microorganisms in their soluble form, leading to a better yield and productivity.

Although there are only a few studies addressing the SHIF of xylan, the simultaneous isomerization and fermentation (SIF) of xylose has been more frequently reported. Rao et al. [27] studied the xylose SIF in the presence of 0.05 M borax to shift equilibrium concentration of xylulose/xylose and improve the isomerization step. However, although the isomerization was enhanced, only half of the available xylose was consumed. Lastick et al. [43] observed an ethanol titer of 2.1% (w/v) from the SIF of 6% xylose using XI and *Schizosaccharomyces pombe* (Y-164). Silva et al. [7] studied the SIF of 65 g/L of xylose at 30 °C, using a biocatalyst containing 32.5×10^3 IU/L of xylose isomerase and 20 g/L of yeast co-immobilized in Ca-alginate gel, and reported an ethanol productivity of 0.25 g/L/h. However, the isomerization step became a limiting factor, due to the decrease of the pH from 5.3 (initial) to 4.8 (final). Milessi et al. [15] incorporated $CaCO_3$ into SIF beads to control the pH

of the process. The biocatalysts were prepared with 20% chitosan-immobilized XI and 10% fresh yeast. An ethanol yield of 0.35 g/g (70% of the theoretical yield) and 2 g/L/h productivity was observed. However, the long time needed by *S. cerevisiae* to ferment xylulose makes the SIF process susceptible to contamination by bacteria capable of metabolizing the xylose.

In this context, the proposed SHIF process appears as a promising approach for 2G ethanol production from hemicellulose. Process conditions and enzyme loads in the biocatalyst can be optimized to achieve higher yields and productivity as well as to overcome the difference between optimal pH ranges for each step of the process. Despite the improvement achieved after pH control and supplementation with excess of Mg^{2+}, a small accumulation of xylose was still observed, which suggests that the isomerization step may be still limiting the process. There are other factors that might be affecting the isomerization and/or the fermentation steps, such as the presence of xylooligosaccharides (2–8 xylose units) or other intermediate products released during the xylan hydrolysis, which are not present in the SIF process. The understanding of the influence of these compounds as well as the optimization of the biocatalyst composition regarding the balance of the enzyme pool are important issues to be addressed in order to improve ethanol production rates.

In general, the SHIF process using co-immobilized enzymes and cells stands out for 2G ethanol production. Besides presenting the advantages of multi-enzymatic cascade reactions, it also enables an easy recovery of the biocatalyst, which could be applied in continuous or repeated batch ethanol production runs using a medium that inhibits contamination. In addition, it builds on the advantage of using the same native yeast, already employed in 1G ethanol industry, which simplifies the operation of the fermentation unit. The low genetic stability of recombinant microorganisms together with the strict Brazilian biosafety regulations for genetically modified organisms (GMO) application in the industrial environment make the 2G ethanol production process based on a native yeast an attractive alternative.

3. Materials and Methods

3.1. Materials

GENSWEET® SGI (3400 IU/mL, 127 $mg_{protein}$/mL, DuPont™ Genencor®, Palo Alto, CA, USA), an enzymatic extract of commercial xylose isomerase (EC 5.3.1.5) from *Streptomyces rubiginosus*, and the commercial enzyme preparations Accellerase XY (3670 IU/mL, 9.8 $mg_{protein}$/mL) and Multifect CX XL A03139 (785 IU/mL, 35 $mg_{protein}$/mL) were kindly donated by DuPont™ Genencor® (Palo Alto, CA, USA). The *Bacillus subtilis* recombinant endoxylanase (502 IU/mL, 9.2 $mg_{protein}$/mL) was donated by Verdartis (Ribeirão Preto, SP, Brazil). Powdered chitosan (85.2% deacetylation degree) was supplied by Polymar Ind. Ltda (Fortaleza, CE, Brazil) and 25% glutaraldehyde solution was purchased from Vetec (Duque de Caxias, RJ, Brazil). Birchwood and beechwood xylans were from Sigma-Aldrich (St. Louis, MO, USA). A *Saccharomyces cerevisiae* industrial strain (purchased from Itaiquara®, Tapiratiba, SP, Brazil) was used in all SHIF experiments. All other reagents were of analytical grade.

3.2. Biocatalyst Production

3.2.1. Preparation of Chitosan-Glutaraldehyde Beads

Chitosan gel (2% or 4%, w/v) was prepared as described by Budriene et al. [44], using the coagulation of the chitosan-acetic acid solution in 0.5 M KOH. Activation of the support was carried out by the addition of glutaraldehyde (5%, v/v) in a suspension of chitosan at pH 7.0 (100 mM phosphate buffer, 1:10 $m_{support}$/$v_{suspension}$). After 60 min stirring at 25 °C, the support was filtered under vacuum, washed first with distilled water until neutrality, and then with ultrapure water.

3.2.2. Enzyme Immobilization

Xylose isomerase immobilization was carried out onto 2% (w/v) chitosan-glutaraldahyde according to Silva et al. [7]. The enzyme solution was prepared in 50 mM Tris-maleate buffer (pH 7.0) containing 5 mM $MgSO_4.7H_2O$ and 2.5 mM $CoCl_2.6H_2O$, in order to provide 50 $mg_{protein}/g_{support}$. The support was added to the enzyme solution at a ratio of 1:10 (w/v). After 20 h of immobilization at 25 °C under 150 rpm stirring, sodium borohydride was added (1 mg/mL) and the suspension was kept under gentle agitation for 30 min in an ice bath [45]. The derivatives were filtered and washed under vacuum, first with 200 mM Tris-maleate buffer (pH 7.0), then with ultrapure water, and finally with 50 mM Tris-maleate buffer (pH 7.0), in order to remove borohydride and adsorbed enzyme.

Xylanase complexes were immobilized onto 4% (w/v) chitosan-glutaraldehyde according to Milessi et al. [28]. The immobilizations were performed at pH 7.0 (100 mM phosphate buffer), 25 °C, and under constant stirring. A load of 20 $mg_{protein}/g_{support}$ was offered, maintaining 1:10 (w/v) ratio of mass of support to volume of enzymatic solution. After the completion of immobilization, sodium borohydride was added (1 mg/mL) and the reduction reaction proceeded for 30 min at 4 °C. The derivatives were extensively washed with 50 mM citrate buffer pH 4.8 and stored until use.

3.2.3. Biocatalyst Co-immobilization

The biocatalyst preparation was carried out through Ca-alginate gel entrapment according to the methodology described in Silva et al. [7]. The industrial strain, supplied as freshly compressed yeast cells, was used as purchased, without previous propagation or activation [46]. A solution of sodium alginate (1% w/v), immobilized XI (15% w/v), immobilized xylanase (5% w/v), and fresh yeast (10% w/v) was gently dropped into a 0.25 M $CaCl_2$/0.25 M $MgCl_2$ solution. Spherical particles (Ø = 1–1.5 mm) were produced using a pneumatic extruder [47]. The procedure was carried out in a laminar flow chamber (Airstream, ESCO, Horsham, PA, USA) and the sodium alginate and coagulation solutions were previously sterilized at 121 °C for 20 min. After immobilization, the beads were cured in a refrigerator for 12–16 h in cure solution (4 g/L of $MgSO_4$, 10 g/L of KH_2PO_4, 3 g/L of urea, 0.2 g/L of $CoCl_2$, and 4 g/L of $CaCl_2.2H_2O$).

3.3. Simultaneous Hydrolysis, Isomerization, and Fermentation (SHIF) of Birchwood Xylan

First, the birchwood xylan substrate was hydrolyzed in smaller xylo-oligomers by the action of recombinant endoxylanase of *B. subtilis* (XynA) in order to make xylan more similar to lignocellulosic hydrolysates, since after a pretreatment step the xylo-oligomers present in the medium are smaller than those in commercial xylan [28]. Xylan pre-hydrolysis was carried out by XynA (150 IU/g_{xylan}, 3.8 IU/mL) immobilized on chitosan-glutaraldehyde (35.0 ± 0.8 IU/g) for 24 h and 50 °C under 150 rpm stirring. At the end, the immobilized enzyme was recovered by filtration and the dried xylan mass retained on the filter was quantified (xylan insoluble fraction). The pH of the xylan solubilized fraction (SHIF substrate) was adjusted to 5.6 with HCl or NaOH 1M and the medium was sterilized by filtration through a 0.22 μm membrane. SHIF experiments were carried out in a shaker incubator (32 °C and 150 rpm), using sealed tubes with a total reaction volume of 2.4 mL (bead ratio of 1:1, 1.2 g of beads and 1.2 mL of medium, bead density was 1 g/cm^3). The composition of SHIF medium was 108 g/L of birchwood xylan supplemented with $MgSO_4$ (4 g/L), KH_2PO_4 (10 g/L), urea (3 g/L), $CoCl_2.6H_2O$ (0.2 g/L), and $CaCl_2.2H_2O$ (4 g/L). Samples were collected at regular intervals for determination of pH, substrate consumption, and product formation.

3.4. Influence of pH, Ca^{2+}, and Mg^{2+} on XI Activity

In order to study the influence of Ca^{2+} and Mg^{2+} ions on the catalytic activity of XI, the enzyme activity was measured at different pHs (5.0 to 8.0) and different ion concentrations. Nine medium compositions were tested for each studied pH. The standard medium was constituted of 2 M fructose at pH 8.0 (50 mM Tris-maleate buffer), 50 mM $MgSO_4$, and 2.5 mM $CoCl_2$.

3.5. Enzymatic Xylan Hydrolysis for Xylanase Selection

Xylan hydrolysis was carried out using the soluble fraction of the birchwood xylan obtained by adding 8 g of commercial birchwood xylan in 100 mL of 50 mM citrate buffer, pH 5.5 at 50 °C. After 1 h at 150 rpm stirring, the solution was centrifuged for 20 min at $9500 \times g$ and 5 °C. The supernatant was then recovered for further use at the concentration of 25 g/L of xylan. It was offered 150 IU/g_{xylan} (3.8 IU/mL$_{reactor}$). The reaction was conducted at 50 °C under mechanical stirring for 24 h.

3.6. Analytical Methods

3.6.1. Xylanase Activity

Xylanase activity was determined according to International Union of Pure and Applied Chemistry (IUPAC) [48] by calculating the initial velocity of xylan hydrolysis catalyzed by a known amount of enzyme. The standard substrate was birchwood xylan (1% w/v) in 50 mM citrate buffer pH 5.5. Enzyme was added to the reaction medium and incubated at 50 °C for 10 min under 250 rpm stirring. Aliquots were withdrawn at 2 min intervals, and the released reducing sugars were quantified by the dinitrosalicylic (DNS) acid method [49]. One unit of activity (IU) was defined as the amount of enzyme required to release 1 μmol of xylose per minute under the assayed conditions.

3.6.2. Xylose Isomerase Activity

Xylose isomerase activity was determined according to Giordano et al. [13], by measuring the initial velocity of fructose isomerization to glucose, under the following conditions: 2 M fructose solution prepared in 50 mM Tris-maleate buffer containing 50 mM $MgSO_4.7H_2O$ and 2.5 mM $CoC1_2.6H_2O$, at pH 7.0 and 60 °C. The glucose concentration was determined colorimetrically using the commercial enzyme kit containing glucose oxidase and peroxidase (GOD-PAP®, Bioclin, Belo Horizonte, Mg, Brazil). One international unit (IU) of xylose isomerase was defined as the amount of enzyme that released 1 μmol of glucose per minute under the assayed conditions.

3.6.3. Substrate and Product Quantification

The concentrations of XOs, xylose, xylulose, xylitol, and ethanol were determined by high performance liquid chromatography (HPLC), equipped with a Waters Sugar-Pak I column (Milford, MA, USA) (300 × 6.5 mm) coupled to a refractive index detector (W410 Waters) (Milford, MA, USA). Ultrapure water was used as eluent at a flow rate of 0.5 mL/min. The column temperature was 80 °C, the detector was set at 40 °C, and the injected volume was 20 μL. Before the analysis, the samples were filtered using a 0.22 μm filter.

4. Conclusions

A new biocatalyst composed of co-immobilized xylanases, xylose isomerase, and unmodified *S. cerevisiae* was able to produce 2G ethanol from birchwood xylan in a one-spot multi-enzymatic simultaneous hydrolysis, isomerization, and fermentation (SHIF) process. Although more studies are required to increase ethanol productivity, the SHIF process showed to be promising from the point of view of its technical viability. The SHIF process brings advantages to industrial applications for easing the integration of 1G and 2G ethanol production processes (because both are based on the same native yeast strains), while reducing the contamination risk due to the use of xylo-oligomers as substrate instead of xylose.

References

1. Antunes, F.A.F.; Chandel, A.K.; Milessi, T.S.S.; Santos, J.C.; Rosa, C.A.; da Silva, S.S. Bioethanol Production from Sugarcane Bagasse by a Novel Brazilian Pentose Fermenting Yeast Scheffersomyces shehatae UFMG-HM 52.2: Evaluation of Fermentation Medium. *Int. J. Chem. Eng.* **2014**, *2014*, 1–8. [CrossRef]
2. UNFCC Conference of the Parties. Adoption of the Paris Agreement. *Report No. FCC/CP/2015/L.9/Rev1. United Nations Framew. Conv. Clim. Chang.* 2015. Available online: http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf (accessed on January 2019).
3. Nigam, P.S.; Singh, A. Production of liquid biofuels from renewable resources. *Prog. Energy Combust. Sci.* **2011**, *37*, 52–68. [CrossRef]
4. Canilha, L.; Chandel, A.K.; Suzane dos Santos Milessi, T.; Antunes, F.A.F.; Luiz da Costa Freitas, W.; das Graças Almeida Felipe, M.; da Silva, S.S. Bioconversion of Sugarcane Biomass into Ethanol: An Overview about Composition, Pretreatment Methods, Detoxification of Hydrolysates, Enzymatic Saccharification, and Ethanol Fermentation. *J. Biomed. Biotechnol.* **2012**, *2012*, 1–15. [CrossRef] [PubMed]
5. Dos Santos, L.V.; de Barros Grassi, M.C.; Gallardo, J.C.M.; Pirolla, R.A.S.; Calderón, L.L.; de Carvalho-Netto, O.V.; Parreiras, L.S.; Camargo, E.L.O.; Drezza, A.L.; Missawa, S.K.; et al. Second-Generation Ethanol: The Need is Becoming a Reality. *Ind. Biotechnol.* **2016**, *12*, 40–57. [CrossRef]
6. Hahn-Hägerdal, B.; Jeppsson, H.; Skoog, K.; Prior, B.A. Biochemistry and physiology of xylose fermentation by yeasts. *Enzyme Microb. Technol.* **1994**, *16*, 933–943. [CrossRef]
7. Silva, C.R.; Zangirolami, T.C.; Rodrigues, J.P.; Matugi, K.; Giordano, R.C.; Giordano, R.L.C. An innovative biocatalyst for production of ethanol from xylose in a continuous bioreactor. *Enzyme Microb. Technol.* **2012**, *50*, 35–42. [CrossRef] [PubMed]
8. Demeke, M.M.; Dietz, H.; Li, Y.; Foulquié-Moreno, M.R.; Mutturi, S.; Deprez, S.; Den Abt, T.; Bonini, B.M.; Liden, G.; Dumortier, F.; et al. Development of a D-xylose fermenting and inhibitor tolerant industrial Saccharomyces cerevisiae strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnol. Biofuels* **2013**, *6*, 89. [CrossRef] [PubMed]
9. Zhang, T.; Chi, Z.; Zhao, C.H.; Chi, Z.M.; Gong, F. Bioethanol production from hydrolysates of inulin and the tuber meal of Jerusalem artichoke by Saccharomyces sp. W0. *Bioresour. Technol.* **2010**, *101*, 8166–8170. [CrossRef] [PubMed]
10. Hector, R.E.; Mertens, J.A.; Bowman, M.J.; Nichols, N.N.; Cotta, M.A.; Hughes, S.R. Saccharomyces cerevisiae engineered for xylose metabolism requires gluconeogenesis and the oxidative branch of the pentose phosphate pathway for aerobic xylose assimilation. *Yeast* **2011**, *28*, 645–660. [CrossRef] [PubMed]
11. Kuyper, M.; Winkler, A.A.; van Dijken, J.P.; Pronk, J.T. Minimal metabolic engineering of Saccharomyces cerevisiae for efficient anaerobic xylose fermentation: A proof of principle. *FEMS Yeast Res.* **2004**, *4*, 655–664. [CrossRef] [PubMed]
12. Ohgren, K.; Bengtsson, O.; Gorwa-Grauslund, M.F.; Galbe, M.; Hahn-Hägerdal, B.; Zacchi, G. Simultaneous saccharification and co-fermentation of glucose and xylose in steam-pretreated corn stover at high fiber content with Saccharomyces cerevisiae TMB3400. *J. Biotechnol.* **2006**, *126*, 488–498. [CrossRef] [PubMed]
13. Giordano, R.L.C.; Giordano, R.C.; Cooney, C.L. A study on intra-particle diffusion effects in enzymatic reactions: Glucose-fructose isomerization. *Bioprocess Eng.* **2000**, *23*, 0159–0166. [CrossRef]
14. Gong, C.S. Recent Advances in D-Xylose Conversion by Yeasts. In *Annual Reports on Fermentation Processes*; Tsao, G.T., Ed.; Academic Press: New York, NY, USA, 1983; Volume 6, pp. 253–297.
15. Milessi, T.S.; Aquino, P.M.; Silva, C.R.; Moraes, G.S.; Zangirolami, T.C.; Giordano, R.C.; Giordano, R.L.C. Influence of key variables on the simultaneous isomerization and fermentation (SIF) of xylose by a native Saccharomyces cerevisiae strain co-encapsulated with xylose isomerase for 2G ethanol production. *Biomass Bioenergy* **2018**, *119*, 277–283. [CrossRef]

16. Silva, S.S.; Mussatto, S.I.; Santos, J.C.; Santos, D.T.; Polizel, J. Cell immobilization and xylitol production using sugarcane bagasse as raw material. *Appl. Biochem. Biotechnol.* **2007**, *141*, 215–227. [CrossRef] [PubMed]

17. Guisan, J.M. New Opportunities for Immobilization of Enzymes. In *Methods in Molecular Biology*; Guisan, J.M., Ed.; Springer: Clifton, NJ, USA, 2013; pp. 1–13.

18. DiCosimo, R.; McAuliffe, J.; Poulose, A.J.; Bohlmann, G. Industrial use of immobilized enzymes. *Chem. Soc. Rev.* **2013**, *42*, 6437–6474. [CrossRef] [PubMed]

19. Choi, J.-M.; Han, S.-S.; Kim, H.-S. Industrial applications of enzyme biocatalysis: Current status and future aspects. *Biotechnol. Adv.* **2015**, *33*, 1443–1454. [CrossRef] [PubMed]

20. Hernández, R.M.A.; Orive, G.; Murua, A.; Pedraz, J.L. Microcapsules and microcarriers for in situ cell delivery. *Adv. Drug Deliv. Rev.* **2010**, *62*, 711–730. [CrossRef] [PubMed]

21. Carvalho, A.F.A.; de Oliva Neto, P.; da Silva, D.F.; Pastore, G.M. Xylo-oligosaccharides from lignocellulosic materials: Chemical structure, health benefits and production by chemical and enzymatic hydrolysis. *Food Res. Int.* **2013**, *51*, 75–85. [CrossRef]

22. Gray, M.C.; Converse, A.O.; Wyman, C.E. Solubilities of Oligomer Mixtures Produced by the Hydrolysis of Xylans and Corn Stover in Water at 180 °C. *Ind. Eng. Chem. Res.* **2007**, *46*, 2383–2391. [CrossRef]

23. Biely, P.; Singh, S.; Puchart, V. Towards enzymatic breakdown of complex plant xylan structures: State of the art. *Biotechnol. Adv.* **2016**, *34*, 1260–1274. [CrossRef] [PubMed]

24. Giordano, R.L.C.; Giordano, R.C.; Zangirolami, T.C.; Tardioli, P.W.; Kopp, W.; Milessi, T.; Silva, C.R.; Cavalcanti-Montano, I.D.; Galeano-Suarez, C.A.; Rojas, M.J.; et al. Sistema catalítico e processo de obtenção de bioetanol 2g a partir de xilana/oligômeros de xilose. Brazilian Patent BR102014023394, 19 September 2014.

25. Macrelli, S.; Galbe, M.; Wallberg, O. Effects of production and market factors on ethanol profitability for an integrated first and second generation ethanol plant using the whole sugarcane as feedstock. *Biotechnol. Biofuels* **2014**, *7*, 26. [CrossRef] [PubMed]

26. Ricca, E.; Brucher, B.; Schrittwieser, J.H. Multi-Enzymatic Cascade Reactions: Overview and Perspectives. *Adv. Synth. Catal.* **2011**, *353*, 2239–2262. [CrossRef]

27. Rao, K.; Chelikani, S.; Relue, P.; Varanasi, S. A Novel Technique that Enables Efficient Conduct of Simultaneous Isomerization and Fermentation (SIF) of Xylose. *Appl. Biochem. Biotechnol.* **2008**, *146*, 101–117. [CrossRef] [PubMed]

28. Milessi, T.S.S.; Kopp, W.; Rojas, M.J.; Manrich, A.; Baptista-Neto, A.; Tardioli, P.W.; Giordano, R.C.; Fernandez-Lafuente, R.; Guisan, J.M.; Giordano, R.L.C. Immobilization and stabilization of an endoxylanase from Bacillus subtilis (XynA) for xylooligosaccharides (XOs) production. *Catal. Today* **2016**, *259*, 130–139. [CrossRef]

29. Fuxreiter, M.; Böcskei, Z.; Szeibert, A.; Szabó, E.; Dallmann, G.; Naray-Szabo, G.; Asboth, B. Role of electrostatics at the catalytic metal binding site in xylose isomerase action: Ca(2+)-inhibition and metal competence in the double mutant D254E/D256E. *Proteins* **1997**, *28*, 183–193. [CrossRef]

30. Kovalevsky, A.Y.; Hanson, L.; Fisher, S.Z.; Mustyakimov, M.; Mason, S.A.; Trevor Forsyth, V.; Blakeley, M.P.; Keen, D.A.; Wagner, T.; Carrell, H.L.; et al. Metal Ion Roles and the Movement of Hydrogen during Reaction Catalyzed by D-Xylose Isomerase: A Joint X-Ray and Neutron Diffraction Study. *Structure* **2010**, *18*, 688–699. [CrossRef] [PubMed]

31. Bhosale, S.H.; Rao, M.B.; Deshpande, V. V Molecular and industrial aspects of glucose isomerase. *Microbiol. Rev.* **1996**, *60*, 280–300. [PubMed]

32. Asboth, B.; Naray-Szabo, G. Mechanism of Action of D-Xylose Isomerase. *Curr. Protein Pept. Sci.* **2000**, *1*, 237–254. [CrossRef] [PubMed]

33. Allen, K.N.; Lavie, A.; Glasfeld, A.; Tanada, T.N.; Gerrity, D.P.; Carlson, S.C.; Farber, G.K.; Petsko, G.A.; Ringe, D. Role of the divalent metal ion in sugar binding, ring opening, and isomerization by D-xylose isomerase: Replacement of a catalytic metal by an amino acid. *Biochemistry* **1994**, *33*, 1488–1494. [CrossRef] [PubMed]

34. Kasumi, T.; Hayashi, K.; Tsumura, N. Roles of Magnesium and Cobalt in the Reaction of Glucose Isomerase from Streptomyces griseofuscus S-41. *Agric. Biol. Chem.* **1982**, *46*, 21–30. [CrossRef]

35. Jacobsen, S.E.; Wyman, C.E. Xylose Monomer and Oligomer Yields for Uncatalyzed Hydrolysis of Sugarcane Bagasse Hemicellulose at Varying Solids Concentration. *Ind. Eng. Chem. Res.* **2002**, *41*, 1454–1461. [CrossRef]

36. Dodd, D.; Cann, I.C.K.O. Enzymatic deconstruction of xylan for biofuel production. *GCB Bioenergy* **2009**, *1*, 2–17. [CrossRef] [PubMed]

37. Akpinar, O.; Erdogan, K.; Bostanci, S. Production of xylooligosaccharides by controlled acid hydrolysis of lignocellulosic materials. *Carbohydr. Res.* **2009**, *344*, 660–666. [CrossRef] [PubMed]
38. Aragon, C.C.; Santos, A.F.; Ruiz-Matute, A.I.; Corzo, N.; Guisan, J.M.; Monti, R.; Mateo, C. Continuous production of xylooligosaccharides in a packed bed reactor with immobilized–stabilized biocatalysts of xylanase from Aspergillus versicolor. *J. Mol. Catal. B Enzym.* **2013**, *98*, 8–14. [CrossRef]
39. Biely, P. Xylanolytic Enzymes. In *Handbook of Food Enzymology*; Whitaker, J., Voragen, A., Wong, D., Eds.; CRC Press: Boca Raton, FL, USA, 2003; pp. 879–916.
40. Rahim, S.N.A.; Alawi, S.; Hamid, K.H.K.; Edama, N.A.; Baharuddin, A.S. Effect of Agitation Speed for Enzymatic Hydrolysis of Tapioca Slurry Using Encapsulated Enzymes in an Enzyme Bioreactor. *Int. J. Chem. Eng. Appl.* **2015**, *6*, 38–41.
41. Carvalho, W.; Canilha, L.; Silva, S.S. Semi-continuous xylose-to-xylitol bioconversion by Ca-alginate entrapped yeast cells in a stirred tank reactor. *Bioprocess Biosyst. Eng.* **2008**, *31*, 493–498. [CrossRef] [PubMed]
42. Nakata, T.; Miyafuji, H.; Saka, S. Ethanol production with β-xylosidase, xylose isomerase, and *Saccharomyces cerevisiae* from the hydrolysate of Japanese beech after hot-compressed water treatment. *J. Wood Sci.* **2009**, *55*, 289–294. [CrossRef]
43. Lastick, S.M.; Tucker, M.Y.; Beyette, J.R.; Noll, G.B.; Grohmann, K. Simultaneous fermentation and isomerization of xylose. *Appl. Microbiol. Biotechnol.* **1989**, *30*, 574–579. [CrossRef]
44. Budriene, S.; Gorochovceva, N.; Romaskevic, T.; Yugova, L.V.; Miezeliene, A.; Dienys, G.; Zubriene, A. β-Galactosidase from Penicillium canescens. Properties and immobilization. *Cent. Eur. J. Chem.* **2005**, *3*, 95–105. [CrossRef]
45. Guisán, J. Aldehyde-agarose gels as activated supports for immobilization-stabilization of enzymes. *Enzyme Microb. Technol.* **1988**, *10*, 375–382. [CrossRef]
46. Basso, L.C.; de Amorim, H.V.; de Oliveira, A.J.; Lopes, M.L. Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res.* **2008**, *8*, 1155–1163. [CrossRef] [PubMed]
47. Trovati, J.; Giordano, R.C.; Giordano, R.L.C. Improving the Performance of a Continuous Process for the Production of Ethanol from Starch. *Appl. Biochem. Biotechnol.* **2009**, *156*, 76–90. [CrossRef] [PubMed]
48. Ghose, T.K.; Bisaria, V.S. Measurement of hemicellulase activities: Part I Xylanases. *Pure Appl. Chem.* **1987**, *59*, 1739–1751. [CrossRef]
49. Miller, G.L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **1959**, *31*, 426–428. [CrossRef]

Highly Selective Oxidation of 5-Hydroxymethylfurfural to 5-Hydroxymethyl-2-Furancarboxylic Acid by a Robust Whole-Cell Biocatalyst

Ran Cang [1,†], Li-Qun Shen [1,†], Guang Yang [1], Zhi-Dong Zhang [3], He Huang [1,2,*] and Zhi-Gang Zhang [1,*]

1 School of Pharmaceutical Sciences, Nanjing Tech University, 30 Puzhu Road(S), Nanjing 211816, China; crnj123@126.com (R.C.); shenliqun163@126.com (L.-Q.S.); qwe88227518@njtech.edu.cn (G.Y.)
2 State Key Laboratory of Materials-Oriented Chemical Engineering, Nanjing Tech University, 30 Puzhu Road(S), Nanjing 211816, China
3 Institute of Microbiology, Xinjiang Academy of Agricultural Sciences, 403 Nanchang Rd, Wulumuqi 830091, China; zhangzheedong@sohu.com
* Correspondence: huangh@njtech.edu.cn (H.H.); zhangzg@njtech.edu.cn (Z.-G.Z.)
† These authors contributed equally to this work.

Abstract: Value-added utilization of biomass-derived 5-hydroxymethylfurfural (HMF) to produce useful derivatives is of great interest. In this work, extremely radiation resistant *Deinococcus wulumuqiensis* R12 was explored for the first time as a new robust biocatalyst for selective oxidation of HMF to 5-hydroxymethylfuroic acid (HMFCA). Its resting cells exhibited excellent catalytic performance in a broad range ofpH and temperature values, and extremely high tolerance to HMF and the HMFCA product. An excellent yield of HMFCA (up to 90%) was achieved when the substrate concentration was set to 300 mM under the optimized reaction conditions. In addition, 511 mM of product was obtained within 20 h by employing a fed-batch strategy, affording a productivity of 44 g/L per day. Of significant synthetic interest was the finding that the *D. wulumuqiensis* R12 cells were able to catalyze the selective oxidation of other structurally diverse aldehydes to their corresponding acids with good yield and high selectivity, indicating broad substrate scope and potential widespread applications in biotechnology and organic chemistry.

Keywords: biocatalysis; extremophile; 5-hydroxymethylfurfural; 5-hydroxymethylfuroic acid; platform chemicals; whole cells

1. Introduction

The production of bio-fuels and chemicals from carbon-neutral and renewable biomass is attracting increasing interest [1–5]. Biomass is regarded as a sustainable resource from which some platform chemicals can be manufactured [6,7]. 5-hydroxymethylfurfural (HMF), derived from lignocellulosic materials via dehydration of carbohydrates, is one of the most important platform chemicals [8–10]. It has been listed as one of "Top 10+4" bio-based chemicals by the U.S. Department of Energy (DOE) [11], being applied in the synthesis of a variety of value-added pharmaceutical and biomaterial intermediates [12]. Due to its high reactivity, HMF is a versatile molecule that can be converted into various useful furan derivatives [12–14]. Its structure comprises a furan ring, an aldehyde group and a hydroxymethyl group which can be subjected to upgrading processes by selective redox reactions, leading to 5-hydroxymethylfuroic acid (HMFCA), 2,5-diformylfuran (DFF), 5-formylfuroic acid (FFCA), 2,5-furandicarboxylic acid (FDCA), maleic anhydride (MA) and 2,5-bis

(hydroxymethyl) furan (BHMF) (Scheme 1). Among these HMF derivatives, the completely oxidized product FDCA displays very promising application potential and may serve as a "greener" substitute for terephthalate in the manufacture of polyester and polyamide materials [15,16]. HMFCA is the oxidation product of the aldehyde group in HMF and a promising starting material for the synthesis of various polyesters [17]. It was reported that HMFCA can also be used as an antitumor agent and interleukin inhibitor [18,19].
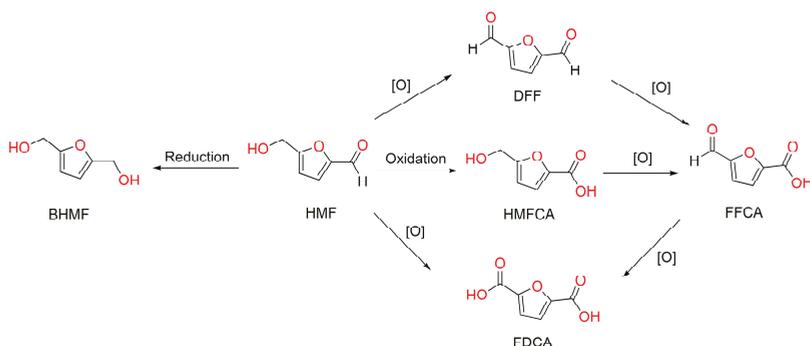
Scheme 1. Catalytic biotransformation of 5-hydroxymethylfurfural (HMF) into high value derivatives.

In order to form HMFCA, selective oxidation of the aldehyde group in HMF is required, while the alcohol group is left intact. Chemoselective oxidation methods are mainly used in the synthesis of HMFCA from HMF, in which noble metal catalysts are generally used [12,20–22]. Recently, HMF was selectively oxidized to HMFCA by an immobilized molybdenum complex in toluene within 3 h, with a yield of approximately 87% [23]. Han et al. reported a selective and mild photocatalytic method for HMFCA synthesis from HMF under ultraviolet and visible light conditions with a yield of 90–95% [24]. In addition, the conversion of HMF to HMFCA via the Cannizzaro reaction is of great value [25,26]. However, the maximal selectivity of HMFCA was 50% due to the formation of an equimolar by-product.

Biocatalytic oxidation of HMF to HMFCA represents a promising alternative to chemical methods [14,27]. Biocatalysis offers many advantages, such as mild, environmentally friendly reaction conditions and often excellent selectivity, as well as high efficiency. However, compared to chemical methods, there are only a few reports on biotransformation of HMF to selectively form HMFCA in the literature [28–32]. In seminal work, Sheldon et al. reported the chloroperoxidase-catalyzed oxidation of HMF affording HMFCA with a selectivity of 25–40% [31]. Krystof et al. reported lipase-mediated and peracid-assisted oxidation of the HMF process to produce HMFCA [32]. Recently, Li and co-workers made use of a molybdenum-dependent enzyme—xanthine oxidase from *Escherichia coli*—for the biocatalytic oxidation of HMF to form HMFCA, with 94% yield and 99% selectivity [29].

Relative to the use of isolated enzymes, we believe that, in HMF oxidation, whole-cell biocatalysts have advantages. They are not only inexpensive and relatively stable, but they also do not require cofactor regeneration [27,33]. Biocatalysis is more efficient when recombinant whole cells that overexpress the enzyme(s) important for catalysis are used [34]. However, employing whole-cell biocatalysts for HMF oxidation is still challenging due to the well-known toxicity of HMF to microbial cells [30]. In addition, due to the variety of enzymes in microbial cells many side reactions are likely to occur during the process of HMF oxidation with formation of HMFCA [28]. Hence, exploring highly tolerant and selective microbial strains is crucial for the biotransformation of HMF into value-added derivative. To our knowledge, there are only a few studies on whole-cell-catalyzed selective oxidation of HMF to form HMFCA in the literature [28,30]; processes that are accompanied by a certain amount of HMF derivatives as byproducts. It was reported that some *Pseudomonas* strains have an HMF degradation pathway, in which HMF is converted to HMFCA as an intermediate [35–37]. A careful

literature search did not reveal any studies describing the use of this system for the production of HMFCA. In 2010, Koopman et al. reported the production of 2,5-furandicarboxylic acid (FDCA) from HMF by using recombinant *P. putida* S12_hmfH. As part of this biotransformation, HMFCA hardly accumulated, leading to a mixture of other metabolites [38,39]. Therefore, in the challenging quest to obtain large amounts of pure HMFCA, the use of the *Pseudomonas* strain metabolic pathway is not feasible. Moreover, long standing issues still exist, such as low substrate loading, substrate toxicity and insufficient selectivity, etc. Therefore, searching for new and robust biocatalytic systems with high selectivity is a demanding task.

Extremophiles are organisms that have evolved to thrive under one or more extreme adverse environmental conditions where other organisms cannot survive [40,41]. They are regarded as an ideal and valuable source of biocatalysts, allowing biotransformation under relatively harsh industrial conditions [42–44]. Nevertheless, employing whole-cells or isolated enzymes derived from extremophiles for biocatalysis in a general manner is just beginning to be implemented experimentally. Recently, a *Deinococcus sp*, designated as *Deinococcus wulumuqiensis* R12, was isolated from radiation-polluted soil [45,46]. Previous studies showed that it is phylogenetically more closely related to a prototype strain of the *Deinococcus* genus, namely *Deinococcus radiodurans* R1 [47]. It was found that this strain was capable of producing carotenoids with good yield, and related biosynthesis genes were subsequently cloned and heterogeneously expressed in *E. coli.* by Xu et al. [48]. Furthermore, its whole genome was sequenced by Huang et al. [49]. Recently, genes encoding heat shock proteins from *D. wulumuqiensis* R12 were introduced into *Clostridium acetobutylicum* ATCC824 in order to improve the robustness and butanol titers of host cells [50]. Considering the robustness of *D. wulumuqiensis* R12, it would be of great interest to explore the catalytic properties of its whole cells in biotransformation or bioconversion.

In this study, we report that the radiation resistant strain *D. wulumuqiensis* R12 that can indeed be used as a whole-cell biocatalyst in HMFCA synthesis by selective oxidation of HMF (Scheme 2). The catalytic properties of this strain were evaluated in the transformation of HMF, and the reaction conditions were optimized. In addition, the substrate scope of this new whole-cell biocatalyst was also investigated.
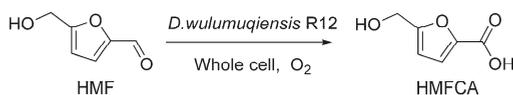
Scheme 2. Chemoselective oxidation of MF to 5-hydroxymethylfuroic acid (HMFCA) with whole *D. wulumuqiensis* R12 cells.

2. Results and Discussion

2.1. Growing and Resting Deinococcus Cells as Catalysts in HMF Oxidation with Selective Formation of HMFCA

Similar to the prototype strain of the *Deinococcus* genus, *D. radiodurans* R1, *D. wulumuqiensis* R12 is also well known for its excellent ability to resist extremely high doses of gamma and UV radiation [45]. In order to explore its potential applications in biocatalysis, growing and resting cells of this strain were applied as biocatalysts in the conversion of HMF to form HMFCA. As shown in Figure 1a, 100 mM of the HMF substrate were converted almost completely within 12 h using resting cells, whereas growing cells gave only a 32% yield at a prolonged reaction time of 36 h. Resting cells enabled a much higher yield with more than 98% of HMFCA and a trace amount of 2,5-bis (hydroxymethyl) furan (BHMF) as sole byproduct, indicating excellent chemoselectivity in this biocatalytic process. Increasing substrate concentration further did not affect the selectivity of the resting cells (Figure S2). Considering the reported degradation mechanism of HMF in microbial cells [51], it is reasonable to speculate that the intermediate HMF alcohol (from HMF reduction) was almost completely oxidized in a very

short time to form the final HMFCA, or the HMF substrate was oxidized directly—which constitutes a different mechanistic hypothesis. However, to validate this inference, more efforts need to be invested.
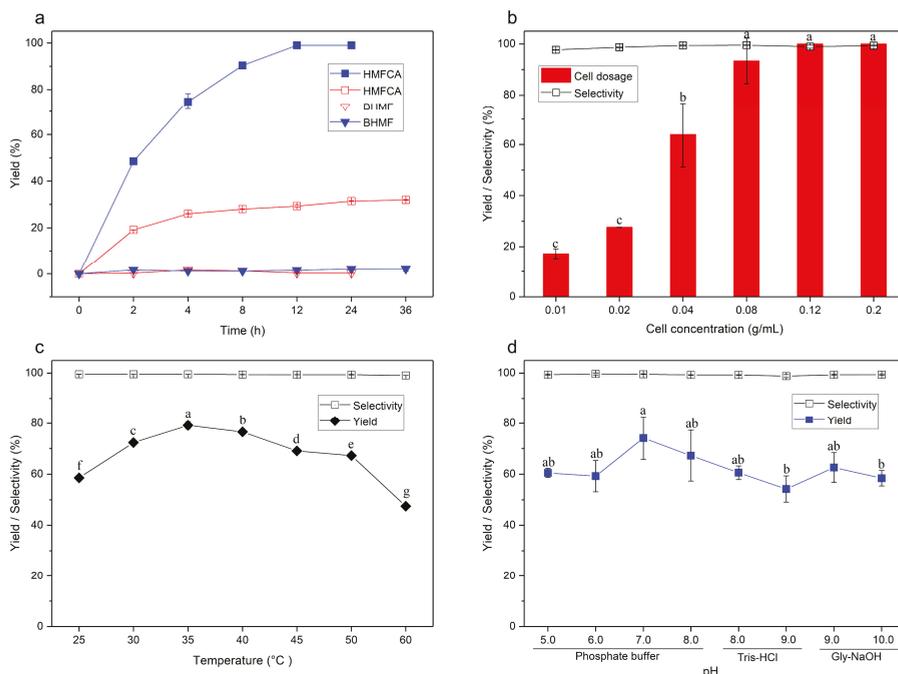

Figure 1. The influence of various factors on HMFCA synthesis by HMF. General conditions unless otherwise stated: 100 mM HMF, 0.12 g/mL microbial cells, 100 mM,pH 7.4, phosphate buffer, 850 rpm, 30 °C, 4h reaction time. (**a**) Resting (solid symbols) and growing cells (open symbols), wherein tryptone glucose yeast extract (TGY) culture was used when growing cells were employed under 200 rpm, 30 °C; (**b**) cell concentration; (**c**) temperature; and (**d**)pH values. Time courses of HMF biotransformation using resting (solid symbols) and growing cells (open symbols).

In addition, we were pleased to discover the performance of two other *Deinococcus* strains stored in our lab, *D. radiodurans* R1 and *Deinococcus xibeiensis* R13, which were also used in the biocatalytic oxidation of HMF to HMFCA. It was found that both radiation resistant strains selectively oxidized HMF with formation of HMFCA. The conversions achieved by *D. radiodurans* R1 and *D. xibeiensis* R13 were slightly lower than that of *D. wulumuqiensis* R12 under the same reaction conditions (Figure S3). These results suggest that the *D. wulumuqiensis* R12 cells act as a catalytic system with high activity and excellent chemoselectivity in the oxidation of HMF to HMFCA. Its catalytic properties were subsequently investigated in greater detail (Figure 1).

2.2. Effect of Cell Dosage in the Reaction System for HMFCA Synthesis

Figure 1b shows the influence of microbial cell dosage on HMFCA synthesis based on selective oxidation of HMF. The yield of HMFCA increased steadily from 18% to 99% with increasing cell dosage in the presence of 100 mM of HMF substrate. The maximal yield of 99% was achieved when the cell dosage reached 0.12 g/mL, and further increasing did not improve the HMFCA yield, indicating that the biocatalyst was potentially saturated by substrate under the given reaction conditions. Our results imply that the conversion of HMF to HMFCA correlates with the cell dosage employed in the biocatalytic system. A small amount of cell dosage was required to reach maximal conversion

when the substrate concentration decreased to 40 mM under the same reaction conditions (Figure S4). In addition, cell dosage had no significant effect on the selectivity of the reactions (>98%). A higher cell dosage may result in higher viscosity, however, which could impact mass transfer of the reaction mixture. Thus, the optimal cell dosage of 0.12 g/mL wet cells was used in subsequent experiments.

2.3. Effect of Temperature and pH on HMFCA Synthesis

The influence of temperature and pH on HMFCA synthesis in the whole-cell catalyzed oxidation of HMF was also studied. As shown in Figure 1c, the effect of reaction temperature on HMF selective oxidation was determined by performing the transformation at different temperatures. Remarkably, the microbial cell biocatalyst showed considerable activity at a broad temperature range, from 25 to 60 °C. The maximal substrate conversion of 79% was obtained at 35 °C after 4 h in the presence of 100 mM HMF substrate. In addition, even at 50 °C, 67% of the HMF substrate was converted to HMFCA, which is in accord with an early report that *D. wulumuqiensis* R12 has a broad growth temperature range [45]. Slightly decreased conversion is possibly due to the inactivation of the enzymes in the microbial cells at 60 °C. It should be mentioned that HMFCA was obtained as essentially the only oxidative product—with a yield of 99%—in the reaction within the temperature range of 25 °C to 60 °C, indicating excellent catalytic selectivity of the whole-cell biocatalyst. Considering the thermostability of cells and energy efficiency, a temperature of 35 °C was set for subsequent experiments.

In addition, we further studied the pH profile of the whole-cell catalyst in HMFCA synthesis via selective oxidation of HMF (Figure 1d). It was found that the microbial cells had a broad pH activity profile and exhibited a particularly good catalytic performance in the pH range of 5.0 to 10.0. The best yield of 81% was achieved in 100 mM phosphate buffer at a pH 7.0 after 4 h. Interestingly, a conversion percentage of 60% and 58% was obtained in phosphate buffer of pH 5.0 and Gly-NaOH buffer of pH 10.0, respectively, after a reaction time of 4 h. In addition, it appeared that the buffer types had a moderate influence on the conversion of HMF, as a yield of 54% was obtained in Tris-HCl buffer (pH 9.0), compared to 63% in Gly-NaOH buffer at the same pH. One should not be surprised that *D. wulumuqiensis* R12 cells are able to resist such harsh reaction conditions with extreme pH values. In their studies, Wang et al. reported that the *D. wulumuqiensis* R12 strain is able to grow in a wide pH range from 5.0 to 12.0 [45]. Compared to *Comamonas testosterone* SC1588, which has been applied in HMFCA synthesis from HMF [28], *D. wulumuqiensis* R12 cells showed higher tolerance to extreme pH values. Therefore, the optimal pH value of 7.0 was selected for all subsequent experiments.

2.4. Inhibitory and Toxic Effect of Substrate

HMF is a well-known toxic inhibitor of microbial cells, inhibiting their growth and hindering their upgrading of HMF by whole-cell biocatalysis [52]. The catalytic performance of *D. wulumuqiensis* R12 cells towards HMF under varying concentrations was therefore tested. As shown in Figure 2a, HMFCA was synthesized in 99% yield within 12 h when the substrate concentration was 150 mM. However, the yield decreased slightly to 81% at the substrate concentration of 200 mM. The yield decreased gradually in the substrate concentration range of 250–1000 mM. Remarkably, 41% yield of HMFCA was obtained when the substrate concentration reached 500 mM, and a yield of 23% was observed at the substrate concentration of 1000 mM. The essentially complete selectivity remained almost constant at these varying substrate concentrations.
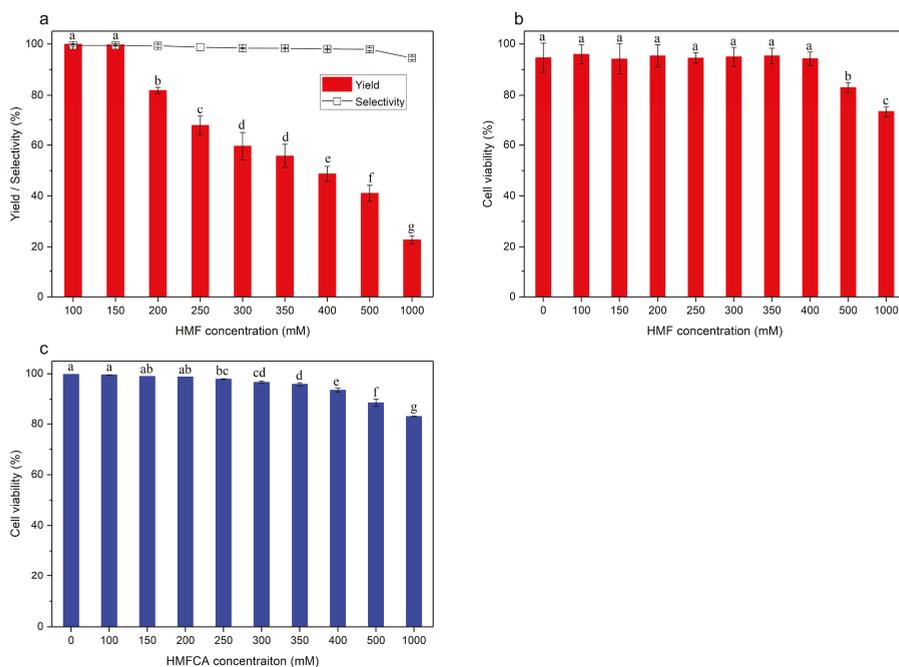
Figure 2. Effects of HMF concentration on (**a**) HMFCA synthesis and (**b**) cell viability. Reaction conditions: 0.12 g/mL microbial cells, in phosphate buffer (100 mM,pH 7.0) under 850 rpm and 35 °C. Fresh harvested cells incubated without HMF/HMFCA under identical conditions were used as a control. Reaction periods: 12 h for 100 mM 150 mM; 24 h for 200 mM, 250 mM; and 36 h for 300 mM, 350 m, 400 mM, 500 mM, 1000 mM. (**c**) Effects of the product concentration on cell viability. HMFCA was incubated for 24 h under the same conditions.

It is well known that the cytotoxicity of HMF to microbial cells is a key parameter in whole-cell biocatalytic conversion of HMF. As shown in Figure 2b, the effect of substrate concentration on cell viability was investigated by using a cell viability assay. The microbial cells were incubated with varying concentrations of HMF under the usual reaction conditions, and the cell viability was subsequently measured using an Annexin V-FITC/PI Apoptosis Detection Kit. Microbial cells incubated in buffer without substrate under the same reaction condition were used as a control. Good cell viability (94%) was unexpectedly obtained in the presence of 400 mM substrate, which was not significantly different than that of the control. This result suggests that the microbial cells can tolerate as much as 400 mM of HMF substrate without losing viability. Further increasing HMF concentration to 500 mM led only to a slight decrease in cell viability to 83%. When 1000 mM of HMF substrate was used, 73% cell viability was still observed, indicating that *D. wulumuqiensis* R12 is extremely tolerant to HMF substrate. However, the conversion of HMF in the oxidation reaction decreased dramatically with increased substrate concentration. Since detailed toxic mechanisms of HMF to microbial cells are not completely understood, nor the reason for the extreme resistance of *D. wulumuqiensis* R12 cells to harsh environmental factors, the present results are not considered surprising.

2.5. Inhibitory and Toxic Effect of the HMFCA Product

In the oxidation of HMF to HMFCA, the product is an acidic compound harboring a carboxylic and hydroxyl group. In our work, the pH of the reaction system decreased over the reaction time due to the accumulation of the HMFCA product. This situation could conceivably become critical when high concentrations of substrate are produced. Thus, it would be of great interest and importance to

investigate the possible inhibition and toxicity of the product towards microbial cells. Microbial cells incubated in phosphate buffer without the product was used as the control for such investigations. Based on the high cell viability (>93%), which was comparable to that of the control, there appeared to be no significant toxicity towards the microbial cells at HMFCA product concentrations less than 400 mM, as shown in Figure 2c. However, a product amount of 500 mM resulted in slight toxicity, as the cell viability value decreased to 87%. A further increased product concentration of 1000 mM led to significant toxicity of the microbial cells, with a viability of 84%. To our surprise, the data showed that the product toxicity to the viability of microbial cells was not as high as expected, even at extremely high concentrations. Considering that *D. wulumuqiensis* R12 is a robust strain, able to grow in a broad range ofpH values, one should not be surprised that the microbial cells are highly resistant to HMFCA production, with excellent cell viability at extremely high product concentrations. Further product toxicity tests with still higher concentrations were not performed, as the conversion of HMF was already very low at the substrate concentration of 1000 mM.

2.6. Manufacture of HMFCA Under Optimized Conditions

Obtaining large amounts of HFMCA is highly desired in biocatalytic oxidation of HMF by whole-cell biocatalysis, with great potential applications in industrial production. Therefore, further enhancement of the catalytic performance of *D. wulumuqiensis* R12 was investigated by optimizing the biocatalytic parameters of the conversion process. It was found that the HMFCA product yields were affected significantly by increasing the substrate concentration. For example, due to the known negative effect of HMF, the yield of HMFCA decreased significantly when HMF concentrations were higher than 300 mM (Figure 2a). Based on the catalytic properties of this strain, increasing the dosage of microbial cells in the reaction may further enhance the HMFCA yield. Recently, Zhang et al. reported that improved synthesis of HFMCA from HMF was obtained by tuning the pH of the reaction mixture using NaOH solution during the catalytic process [28]. Thus, we decided to employ the same strategy. In addition, it has been reported that adding furfural and furfural alcohol as inducers during cultivation of microbial cells for biocatalysis can trigger the expression of the enzymes responsible for HMF oxidation, which can facilitate HMFCA production. Therefore, three strategies (increasing microbial cell dosage, using inducing cells and tuningpH of reaction mixture) were applied in subsequent studies.

Increasing the dosage of microbial cells proved to be effective for enhancing the yield of HMFCA (Figure S4a). For example, in the presence of 300 mM HMF, the yield of HMFCA increased considerably from 59% to 71% when the concentration of cells increased from 0.12 g/mL to 0.2 g/mL. A further increase in cells dosage was not performed considering the negative effect of higher viscosity in the reaction mixture. On the other hand, influencing the expression of cells by the use of furfural and furfural alcohol for enhancing the yield of HMFCA proved not to be effective (Figure S4b). The reason for this might be that enzymes in *D. wulumuqiensis* R12 responsible for HMF oxidation are expressed constitutively. Finally,pH tuning was found to be an effective method for improving the yield of HMFCA (Figure S4c). Thus, the pH of the reaction mixture was tuned to approximately 7.0 using a NaOH solution. Compared with the control withoutpH tuning, the HMFCA yield improved from 66% to 83% in the presence of 300 mM HMF substrate, and from 48% to 65% at an HMF concentration of 500 mM.

Therefore, both increasing the dosage of microbial cells and pH tuning was applied together for enhancing the production of HMFCA. As shown in Figure 3, at a high HMFCA concentration, a yield of 90% was achieved after 36 h when the substrate concentration was 300 mM. This demonstrates the considerable effectiveness of the combined strategy. In addition, the oxidative conversion of HMF to HMFCA reached 80% in the presence of 350 mM HMF after 48 h. When the concentration of HMF was set to 500 mM, 66% of the substrate was still converted after 48 h, but further prolonged reaction times did not lead to an increase in HMFCA yield.
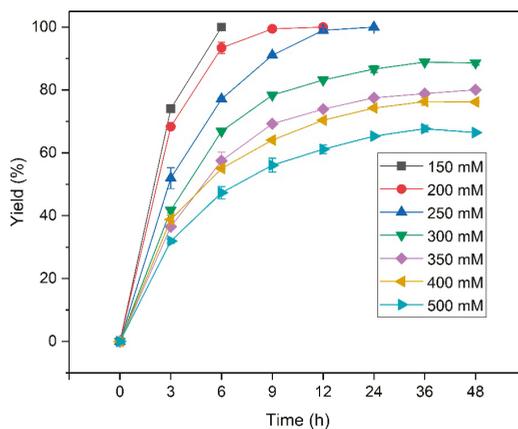
Figure 3. Synthesis of HMFCA under optimized conditions. Reaction conditions: HMF of the designated concentration, 0.2 g/mL microbial cells, 5 mL phosphate buffer (100 mM,pH 7.0), 35 °C, and 850 rpm. Tuning the pH of the reaction system to approximately 7.0 occurred every 3 h in the first 12 h, and then every 12 h in the 36 h that followed.

Compared with previous reported biocatalysis results, the data obtained in this work proved to be more efficient and selective because of a higher substrate concentration and simpler catalytic process. As shown in Table 1, the substrate concentrations used in the reported biocatalytic routes were still very low and co-enzymes were usually required when isolated enzymes were applied. Although *C. testosterone* SC1588 cells also display good selectivity and high HMF tolerance, its catalytic performance is highly sensitive topH [28]. Thus, a considerable amount of histidine co-substrate is required for efficient selective oxidation of HMF. The extreme environment-derived *D. wulumuqiensis* R12 strain used in this work exhibited excellent resistance to highpH and temperatures, and proved to be a robust biocatalyst for HMFCA synthesis by way of selective oxidation of HMF.

Table 1. HMFCA synthesis via HMF oxidation by various biocatalytic systems.

Biocatalysts	Reaction Conditions	t (h)	Yield (%)	Ref.
Chloroperoxidase	50 mM HMF, 1 equiv H_2O_2 per 2 h	2.5	25–40	[31]
Serratia liquefaciens LF14	10 mM HMF, 18.2 mg/mL dry cells, in phosphate buffer	1	97	[30]
Immobilized lipase B	50 mM HMF, 10 mg/mL catalase, addition of aqueous H_2O_2 (30% v/v) hourly, reaction media: acyl butyrate/tBuOH (1:1 v/v), 40 °C	24	76	[32]
Xanthine oxidase (XO)	26 mM HMF, 5.6 U *E. coli* XO, 1.1 mg catalase, phosphate buffer, 37 °C, 150 rpm, air bubbling	7	94	[29]
Comamonas testosterone SC1588	160 mM HMF, 30 mg/mL induced microbial cells, phosphate buffer, 20 mM histidine, 150 rpm, 30 °C, tuningpH of the reaction mixture to approximately 7.0 every 24 h.	36	98	[28]
Aldehyde dehydrogenases	20 mM HMF, 10 µM catalase, 5 µM [NOx], 100 µg/mL [DTT], 0.5 mM [NAD^+] , phosphate buffer, 40 °C, 180 rpm	24	91	[53]
D. wulumuqiensis R12	300 mM HMF, 200 mg/mL microbial cells, phosphate buffer, 850 rpm, 35 °C, tuningpH of the reaction mixture to approximately 7.0 every 3 h.	36	90	This work

2.7. Efficient Synthesis of HMFCA by a Fed-Batch Strategy

As mentioned above, excellent yields of HMFCA from selective oxidation of HMF were obtained under optimized conditions. It is highly desirable to manufacture HMFCA on a large scale in an effort to create a practical biocatalytic process. Thus, by applying a fed-batch strategy, in which HMF substrate was added continuously, the accumulation of high concentrations of product was achieved.

Figure 4 shows the results of biocatalytic synthesis of high concentrations of HMFCA. It was found that 511 mM of product was produced within 20 h after three-batch feeding of HMF, affording a total yield of 85% and a productivity of approximately 44 g/L per day. Only 4 mM of BHMF was observed as the sole byproduct (<1%). Chemoselectivity towards the target product reached more than 99%. In addition, a decrease in yield of HMFCA in each batch feeding was observed, indicating possible substrate and/or product inhibition in the whole-cell biocatalyst. An attempt to improve the yield of HMFCA further in this fed-batch process was performed by prolonging the reaction time, but no significant improvement was observed (Data not shown).
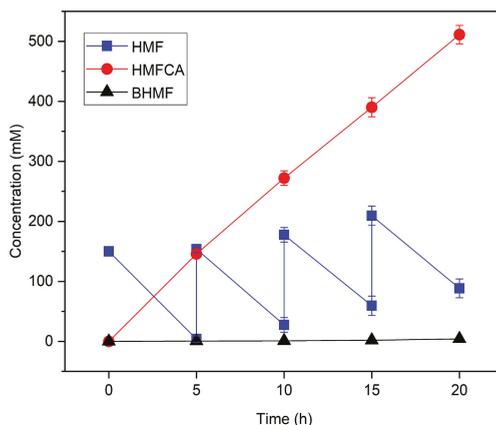
Figure 4. Synthesis of HMFCA by a fed-batch method. Reaction conditions: 150 mM HMF, 0.2 g/mL microbial cells, 5 mL phosphate buffer (100 mM, pH 7.0), 35 °C, and 850 rpm. In each cycle of 5 h, 0.75 mmoL of HMF was added.

2.8. Exploring the Substrate Scope of D. wulumuqiensis R12

In order to examine the substrate spectrum of this novel whole-cell biocatalyst, a set of structurally unique aldehyde compounds was applied in the oxidation reactions catalyzed by *D. wulumuqiensis* R12 cells (Scheme 3). Considering possible solubility and toxicity effects, proper concentrations of these aldehyde compounds were applied. The results showed that the whole cells of the strain readily accept furfural as a substrate, thereby enabling an efficient synthesis of furoic acid (>99% yield). Furoic acid can be used in the pharmaceutical, agrochemical, and flavor industries [9]. In the case of aldehydes containing an additional hydroxyl group, the microbial cells proved to be strictly chemoselective for the aldehyde group, affording the corresponding carboxylic acids with good to excellent yields (Table 2). For example, oxidation of DFF and FFCA to FDCA was achieved with 100% and 63% yields, respectively. Due to solubility issues, higher DFF substrate concentrations were not applied. We also discovered that the aldehyde group of vanillin could be selectively oxidized by the microbial cells to vanillic acid. Vanillic acid has important applications in the pharmaceutical industry, and also as a monomer in polyester synthesis [54]. In addition, terephthalic acid and p-hydroxybenzoic acid were also prepared by selective oxidation of the corresponding aldehydes. In short, the results showed that *D. wulumuqiensis* R12 as a biocatalyst has an amazingly broad substrate spectrum.
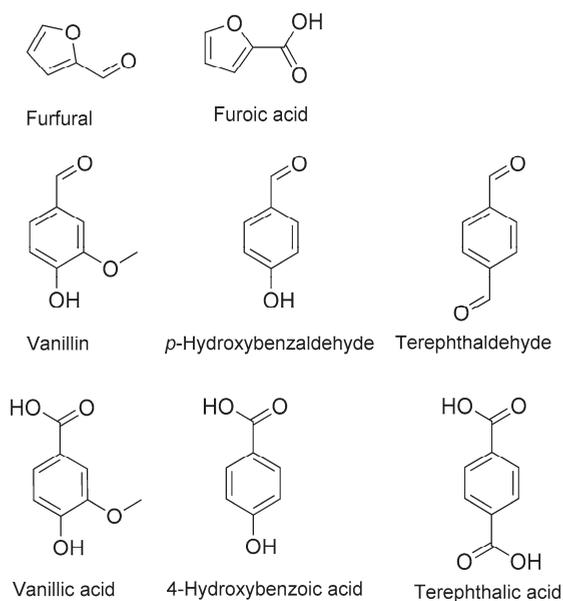
Scheme 3. Aldehyde substrates used and products obtained in biocatalytic oxidation by *D. wulumuqiensis* R12 whole cells.

Table 2. Whole-cell biocatalytic oxidation of aldehydes.

Entry	Substrate	Substrate Concentration (mM)	Product	Yield (%)
1	Furfural	100	Furoic acid	>99
2	DFF	30	FDCA	>99
3	FFCA	100	FDCA	63
4	Vanillin	5	Vanillic acid	>99
5	*p*-Hydroxybenzaldehyde	5	4-Hydroxybenzoic acid	70
6	Terephthaldehyde	2	Terephthalic acid	>99

Reaction conditions: 0.12 g/mL microbial cells, 5 mL of phosphate buffer (100 mM, pH 7.0), 850 rpm, 35 °C for 12 h.

3. Materials and Methods

3.1. Chemicals and Strains

Extremely radiation resistant strains *D. wulumuqiensis* R12 (DSM 28115T), *D. radiodurans* R1 (ATCC NO.13939), and *D. xibeiensis* R13 (NRBC 105666T) were acquired from Zhi-Dong Zhang at the Institute of Microbiology, Xinjiang Academy of Agricultural Sciences in China and stored in our laboratory. The GenBank accession number for the 16S rDNA sequence was KJ784486, while the whole genome sequence was APCS01000000.

HMF (97%) and HMFCA (97%) were purchased from Macklin Biochemical Co., Ltd (Shanghai, China). FDCA (98%), DFF (98%) and furfural (98%) were obtained from Aladdin Biochemical Technology Co., Ltd (Shanghai, China). BHMF (98%) was purchased from Ark Pharm, Inc (Arlington Heights, IL, USA). HMFCA (97%) was obtained from J&K Scientific Ltd (Beijing, China). Both furfuryl alcohol (99.5%) and furoic acid (98%) were obtained from TCI (Shanghai, China). Annexin V-FITC/PI Apoptosis Detection Kit was purchased from Nanjing KeyGen Biotech. Co. Ltd. (Jiangsu, China) for the cell viability assay.

3.2. Cultivation of D. wulumuqiensis R12 Cells

The glycerol stock of *D. wulumuqiensis* R12 was pre-cultivated at 30 °C, 200 rpm for 24 h in tryptone glucose yeast extract (TGY) medium containing 0.5% tryptone, 0.1% glucose, and 0.3% yeast extract. Then, 1% of the overnight preculture was transferred to fresh TGY medium. The culture was incubated at 30 °C, 200 rpm for 48 h and was centrifuged under 5000 rpm for 10 min to harvest cells. The cell pellet was washed twice with 100 mM phosphate buffer (pH 7.4) and resuspended in phosphate buffer with a final cell concentration of 0.12 g/mL (cell wet weight).

3.3. General Procedure for the Biocatalytic Oxidation of Aldehyde Substrates

Five milliliters of phosphate buffer (0.1 M, pH 7.0) containing the designated amounts of microbial cells (cell wet weight) and substrates was incubated at 35 °C and 850 rpm for a given reaction time. Aliquots of the reaction mixture were withdrawn at specified reaction times and diluted with the phosphate buffer prior to high-performance liquid chromatography (HPLC) assays. The conversion of HMF and other aldehydes by biocatalytic oxidation was defined as the percentage of the consumed substrate amount in the initial amount. The selectivity of the reaction was defined as the ratio of HMFCA product amount to the sum of all the products. The yield was defined as the percentage of the measured product amount in the theoretical product amount based on the initial amount of HMF.

$$\% \; yield \; = \; \frac{\text{Actual yield}}{\text{Theoretical yield}} \times 100\%$$

3.4. Analytical Method

The reaction products were analyzed by HPLC following a previously reported method with slight modifications [28]. Briefly, a reverse-phase HPLC (Thermo Fisher ultimate 3000), equipped with Sepax GP-C18 column (4.6 mm × 250 mm, 5 μm), was used at 25 °C. The mobile phase was the gradient of acetonitrile in 20 mM KH_2PO_4 (pH 6.0) at a flow of 1.0 mL min^{-1}, increasing from 10% to 24% within 7 min and from 24% to 10% within 3 min. The HPLC retention time of the HMFCA product and HMF were 2.90 min and 6.20 min, respectively. All experiments were performed in triplicate and mean values are presented. Data are expressed as the mean ± standard deviation. Duncan's multiple range test (using SPSS software 16.0, Chicago, IL, USA) was used to analyze the statistical significance of differences between the groups. A significance difference was judged to exist at a level of $p < 0.05$. HPLC runs are shown in Figure S1.

3.5. Cell Viability Assay

Cell viability assay experiments were performed using an Annexin V-FITC/PI Apoptosis Detection Kit and flow cytometry following the manufacturers' instructions. Cell viability was determined using ACEA NovoCyte Flow Cytometer (ACEA Biosciences, Inc., San Diego, CA, USA) with the excitation light and emission light wavelengths set at 488 nm and 530 nm, respectively. Data were collected and analyzed using NovoExpress software. The cell viability of *D. wulumuqiensis* R12 when using HMF as the substrate is presented as the percentage of living cells to the total amount of cells.

3.6. Synthesis of HMFCA by the Substrate Fed-Batch Feeding Process

Five milliliters of 100 mM phosphate buffer, pH 7.0, which contained 150 mM HMF substrate and 0.2 g/mL of microbial cells, was incubated at 35 °C and 850 rpm. After 5 h, 0.75 mmol of HMF was repeatedly added to the reaction mixture. During the whole biocatalytic process, the pH of the reaction mixture was adjusted to the range of 7.0–8.0 with NaOH solution and the concentration of substrate and products was analyzed by HPLC.

4. Conclusions

Herein, we successfully explored for the first time the use of a radiation resistant *D. wulumuqiensis* R12 strain as a whole-cell biocatalyst for the efficient synthesis of HMFCA from HMF. The whole cells of this strain proved to be highly tolerant to HMF and the product, HMFCA. The whole-cell system is an excellent biocatalyst for the selective oxidation of HMF. An excellent yield of HMFCA of up to 90% was achieved within 36 h in the presence of 300 mM HMF substrate under optimized conditions. A yield of 80% to 66% was obtained when the substrate concentration increased from 350 mM to 500 mM, while the selectivity towards HMFCA remained at approximately 98%. In addition, up to 511 mM of HMFCA was synthesized in 20 h via a fed-batch method, resulting in a productivity of 44 g/L per day. Thus, *D. wulumuqiensis* R12 cells are a promising catalyst in the biocatalytic process of HMF upgrading. Moreover, the cells were able to transform a set of structurally different aldehyde compounds into their corresponding carboxylic acids with good to excellent selectivity. Since the genome sequence of this strain has been sequenced, exploring the genes that encode the enzymes responsible for HMFCA synthesis from HMF has become feasible in future work. The catalytic properties of these microbial cells can also be further engineered by introduction of other oxidases to form a cell factory for HMF biotransformation. Furthermore, this strain may also have potential applications for the biodetoxification of lignocellulosic hydrolysates in the process of biofuel production. Discovery of *D. wulumuqiensis* R12 as an efficient biocatalyst broadens the toolbox of biocatalysts for the biotransformation of HMF into value-added derivatives and will further facilitate the utilization of biomass for the production of useful chemicals and biofuels.

References

1. Sheldon, R.A. Green and sustainable manufacture of chemicals from biomass: State of the art. *Green Chem.* **2014**, *16*, 950–963. [CrossRef]
2. Wu, L.; Moteki, T.; Gokhale, A.A.; Flaherty, D.W.; Toste, F.D. Production of Fuels and Chemicals from Biomass: Condensation Reactions and Beyond. *Chem* **2016**, *1*, 32–58. [CrossRef]
3. Tuck, C.O.; Pérez, E.; Horváth, I.T.; Sheldon, R.A.; Poliakoff, M. Valorization of Biomass: Deriving More Value from Waste. *Science* **2012**, *337*, 695–699. [CrossRef] [PubMed]
4. Zhang, Z.; Deng, K. Recent Advances in the Catalytic Synthesis of 2,5-Furandicarboxylic Acid and Its Derivatives. *ACS Catal.* **2015**, *5*, 6529–6544. [CrossRef]
5. Christensen, C.H.; Rass-Hansen, J.; Marsden, C.C.; Taarning, E.; Egeblad, K. The Renewable Chemicals Industry. *ChemSusChem* **2008**, *1*, 283–289. [CrossRef] [PubMed]
6. Gallezot, P. Conversion of biomass to selected chemical products. *Chem. Soc. Rev.* **2012**, *41*, 1538–1558. [CrossRef] [PubMed]
7. Isikgor, F.H.; Becer, C.R. Lignocellulosic biomass: A sustainable platform for the production of bio-based chemicals and polymers. *Polym. Chem.* **2015**, *6*, 4497–4559. [CrossRef]

8. Chheda, J.N.; Román-Leshkov, Y.; Dumesic, J.A. Production of 5-hydroxymethylfurfural and furfural by dehydration of biomass-derived mono- and poly-saccharides. *Green Chem.* **2007**, *9*, 342–350. [CrossRef]
9. Van Putten, R.-J.; van der Waal, J.C.; de Jong, E.; Rasrendra, C.B.; Heeres, H.J.; de Vries, J.G. Hydroxymethylfurfural, A Versatile Platform Chemical Made from Renewable Resources. *Chem. Rev.* **2013**, *113*, 1499–1597. [CrossRef] [PubMed]
10. Teong, S.P.; Yi, G.; Zhang, Y. Hydroxymethylfurfural production from bioresources: Past, present and future. *Green Chem.* **2014**, *16*, 2015–2026. [CrossRef]
11. Bozell, J.J.; Petersen, G.R. Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited. *Green Chem.* **2010**, *12*, 539–554. [CrossRef]
12. Rosatella, A.A.; Simeonov, S.P.; Frade, R.F.M.; Afonso, C.A.M. 5-Hydroxymethylfurfural (HMF) as a building block platform: Biological properties, synthesis and synthetic applications. *Green Chem.* **2011**, *13*, 754–793. [CrossRef]
13. Koopman, F.; Wierckx, N.; de Winde, J.H.; Ruijssenaars, H.J. Efficient whole-cell biotransformation of 5-(hydroxymethyl)furfural into FDCA, 2,5-furandicarboxylic acid. *Bioresour. Technol.* **2010**, *101*, 6291–6296. [CrossRef] [PubMed]
14. Hu, L.; He, A.; Liu, X.; Xia, J.; Xu, J.; Zhou, S.; Xu, J. Biocatalytic Transformation of 5-Hydroxymethylfurfural into High-Value Derivatives: Recent Advances and Future Aspects. *ACS Sustain. Chem. Eng.* **2018**, *6*, 15915–15935. [CrossRef]
15. Gandini, A.; Silvestre, A.J.D.; Neto, C.P.; Sousa, A.F.; Gomes, M. The furan counterpart of poly(ethylene terephthalate): An alternative material based on renewable resources. *J. Polym. Sci. Part A Polym. Chem.* **2009**, *47*, 295–298. [CrossRef]
16. Sousa, A.F.; Vilela, C.; Fonseca, A.C.; Matos, M.; Freire, C.S.R.; Gruter, G.-J.M.; Coelho, J.F.J.; Silvestre, A.J.D. Biobased polyesters and other polymers from 2,5-furandicarboxylic acid: A tribute to furan excellency. *Polym. Chem.* **2015**, *6*, 5961–5983. [CrossRef]
17. Hirai, H. Oligomers from Hydroxymethylfurancarboxylic Acid. *J. Macromol. Sci. Part A Chem.* **1984**, *21*, 1165–1179. [CrossRef]
18. Tamura, G. Antitumor Activity of 5-Hydroxy-methyl-2-furoic Acid AU—Munekata, Masanobu. *Agric. Biol. Chem.* **1981**, *45*, 2149–2150. [CrossRef]
19. Braisted, A.C.; Oslob, J.D.; Delano, W.L.; Hyde, J.; McDowell, R.S.; Waal, N.; Yu, C.; Arkin, M.R.; Raimundo, B.C. Discovery of a Potent Small Molecule IL-2 Inhibitor through Fragment Assembly. *J. Am. Chem. Soc.* **2003**, *125*, 3714–3715. [CrossRef]
20. Casanova, O.; Iborra, S.; Corma, A. Biomass into Chemicals: Aerobic Oxidation of 5-Hydroxymethyl-2-furfural into 2,5-Furandicarboxylic Acid with Gold Nanoparticle Catalysts. *ChemSusChem* **2009**, *2*, 1138–1144. [CrossRef]
21. Gorbanev, Y.Y.; Klitgaard, S.K.; Woodley, J.M.; Christensen, C.H.; Riisager, A. Gold-Catalyzed Aerobic Oxidation of 5-Hydroxymethylfurfural in Water at Ambient Temperature. *ChemSusChem* **2009**, *2*, 672–675. [CrossRef]
22. Davis, S.E.; Houk, L.R.; Tamargo, E.C.; Datye, A.K.; Davis, R.J. Oxidation of 5-hydroxymethylfurfural over supported Pt, Pd and Au catalysts. *Catal. Today* **2011**, *160*, 55–60. [CrossRef]
23. Zhang, Z.; Liu, B.; Lv, K.; Sun, J.; Deng, K. Aerobic oxidation of biomass derived 5-hydroxymethylfurfural into 5-hydroxymethyl-2-furancarboxylic acid catalyzed by a montmorillonite K-10 clay immobilized molybdenum acetylacetonate complex. *Green Chem.* **2014**, *16*, 2762–2770. [CrossRef]
24. Zhou, B.; Song, J.; Zhang, Z.; Jiang, Z.; Zhang, P.; Han, B. Highly selective photocatalytic oxidation of biomass-derived chemicals to carboxyl compounds over Au/TiO$_2$. *Green Chem.* **2017**, *19*, 1075–1081. [CrossRef]
25. Subbiah, S.; Simeonov, S.P.; Esperança, J.M.S.S.; Rebelo, L.P.N.; Afonso, C.A.M. Direct transformation of 5-hydroxymethylfurfural to the building blocks 2,5-dihydroxymethylfurfural (DHMF) and 5-hydroxymethyl furanoic acid (HMFA) via Cannizzaro reaction. *Green Chem.* **2013**, *15*, 2849–2853. [CrossRef]
26. Kang, E.-S.; Chae, D.W.; Kim, B.; Kim, Y.G. Efficient preparation of DHMF and HMFA from biomass-derived HMF via a Cannizzaro reaction in ionic liquids. *J. Ind. Eng. Chem.* **2012**, *18*, 174–177. [CrossRef]
27. Domínguez de María, P.; Guajardo, N. Biocatalytic Valorization of Furans: Opportunities for Inherently Unstable Substrates. *ChemSusChem* **2017**, *10*, 4123–4134. [CrossRef]

28. Zhang, X.-Y.; Zong, M.-H.; Li, N. Whole-cell biocatalytic selective oxidation of 5-hydroxymethylfurfural to 5-hydroxymethyl-2-furancarboxylic acid. *Green Chem.* **2017**, *19*, 4544–4551. [CrossRef]

29. Qin, Y.-Z.; Li, Y.-M.; Zong, M.-H.; Wu, H.; Li, N. Enzyme-catalyzed selective oxidation of 5-hydroxymethylfurfural (HMF) and separation of HMF and 2,5-diformylfuran using deep eutectic solvents. *Green Chem.* **2015**, *17*, 3718–3722. [CrossRef]

30. Mitsukura, K.; Sato, Y.; Yoshida, T.; Nagasawa, T. Oxidation of heterocyclic and aromatic aldehydes to the corresponding carboxylic acids by *Acetobacter* and *Serratia* strains. *Biotechnol. Lett* **2004**, *26*, 1643–1648. [CrossRef]

31. Van Rantwijk, F.; Sheldon, R.A. Chloroperoxidase-Catalyzed Oxidation of 5-Hydroxymethylfurfural. *J. Carbohydr. Chem.* **1997**, *16*, 299–309. [CrossRef]

32. Krystof, M.; Pérez-Sánchez, M.; Domínguez de María, P. Lipase-Mediated Selective Oxidation of Furfural and 5-Hydroxymethylfurfural. *ChemSusChem* **2013**, *6*, 826–830. [CrossRef] [PubMed]

33. Ni, Y.; Xu, J.-H. Biocatalytic ketone reduction: A green and efficient access to enantiopure alcohols. *Biotechnol. Adv.* **2012**, *30*, 1279–1288. [CrossRef] [PubMed]

34. Ku, S. Finding and Producing Probiotic Glycosylases for the Biocatalysis of Ginsenosides: A Mini Review. *Molecules* **2016**, *21*, 645. [CrossRef] [PubMed]

35. Trudgill, P.W. The metabolism of 2-furoic acid by *Pseudomanas* F2. *Biochem. J.* **1969**, *113*, 577–587. [CrossRef]

36. Koenig, K.; Andreesen, J.R. Xanthine dehydrogenase and 2-furoyl-coenzyme A dehydrogenase from *Pseudomonas putida* Fu1: Two molybdenum-containing dehydrogenases of novel structural composition. *J. Bacteriol.* **1990**, *172*, 5999–6009. [CrossRef]

37. López, M.J.; Nichols, N.N.; Dien, B.S.; Moreno, J.; Bothast, R.J. Isolation of microorganisms for biological detoxification of lignocellulosic hydrolysates. *Appl. Microbiol. Biotechnol.* **2004**, *64*, 125–131. [CrossRef]

38. Koopman, F.; Wierckx, N.; de Winde, J.H.; Ruijssenaars, H.J. Identification and characterization of the furfural and 5-(hydroxymethyl)furfural degradation pathways of *Cupriavidus basilensis* HMF14. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4919–4924. [CrossRef]

39. Nikel, P.I.; Chavarría, M.; Danchin, A.; de Lorenzo, V. From dirt to industrial applications: *Pseudomonas putida* as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr. Opin. Chem. Biol.* **2016**, *34*, 20–29. [CrossRef]

40. Hoover, R.B.; Tang, J. Microbial Extremophiles at the Limits of Life AU - Pikuta, Elena V. *Crit. Rev. Microbiol.* **2007**, *33*, 183–209. [CrossRef]

41. Madigan, M.T.; Orent, A. Thermophilic and halophilic extremophiles. *Curr. Opin. Microbiol.* **1999**, *2*, 265–269. [CrossRef]

42. Schiraldi, C.; De Rosa, M. The production of biocatalysts and biomolecules from extremophiles. *Trends Biotechnol.* **2002**, *20*, 515–521. [CrossRef]

43. Van den Burg, B. Extremophiles as a source for novel enzymes. *Curr. Opin. Microbiol.* **2003**, *6*, 213–218. [CrossRef]

44. Elleuche, S.; Schröder, C.; Sahm, K.; Antranikian, G. Extremozymes—Biocatalysts with unique properties from extremophilic microorganisms. *Curr. Opin. Biotechnol.* **2014**, *29*, 116–123. [CrossRef] [PubMed]

45. Wang, W.; Mao, J.; Zhang, Z.; Tang, Q.; Xie, Y.; Zhu, J.; Zhang, L.; Liu, Z.; Shi, Y.; Goodfellow, M. *Deinococcus wulumuqiensis sp. nov.*, and *Deinococcus xibeiensis sp. nov.*, isolated from radiation-polluted soil. *Int. J. Syst. Evol. Microbiol.* **2010**, *60*, 2006–2010. [CrossRef] [PubMed]

46. Hong, S.; Farrance, C.E.; Russell, A.; Yi, H. Reclassification of *Deinococcus xibeiensis* Wang et al. 2010 as a heterotypic synonym of *Deinococcus wulumuqiensis* Wang et al. 2010. *Int. J. Syst. Evol. Microbiol.* **2015**, *65*, 1083–1085. [CrossRef]

47. Battista, J.R.; Earl, A.M.; Park, M.-J. Why is *Deinococcus radiodurans* so resistant to ionizing radiation? *Trends Microbiol.* **1999**, *7*, 362–365. [CrossRef]

48. Xu, X.; Tian, L.; Xu, J.; Xie, C.; Jiang, L.; Huang, H. Analysis and expression of the carotenoid biosynthesis genes from *Deinococcus wulumuqiensis* R12 in engineered Escherichia coli. *AMB Express* **2018**, *8*, 94. [CrossRef]

49. Xu, X.; Jiang, L.; Zhang, Z.; Shi, Y.; Huang, H. Genome Sequence of a Gamma- and UV-Ray-Resistant Strain, *Deinococcus wulumuqiensis* R12. *Genome Announc.* **2013**, *1*, e00206–e00213. [CrossRef]

50. Liao, Z.; Zhang, Y.; Luo, S.; Suo, Y.; Zhang, S.; Wang, J. Improving cellular robustness and butanol titers of *Clostridium acetobutylicum* ATCC824 by introducing heat shock proteins from an extremophilic bacterium. *J. Biotechnol.* **2017**, *252*, 1–10. [CrossRef]

51. Wierckx, N.; Koopman, F.; Ruijssenaars, H.J.; de Winde, J.H. Microbial degradation of furanic compounds: Biochemistry, genetics, and impact. *Appl. Microbiol. Biotechnol.* **2011**, *92*, 1095–1105. [CrossRef] [PubMed]
52. Palmqvist, E.; Hahn-Hägerdal, B. Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition. *Bioresour. Technol.* **2000**, *74*, 25–33. [CrossRef]
53. Knaus, T.; Tseliou, V.; Humphreys, L.D.; Scrutton, N.S.; Mutti, F.G. A biocatalytic method for the chemoselective aerobic oxidation of aldehydes to carboxylic acids. *Green Chem.* **2018**, *20*, 3931–3943. [CrossRef]
54. Pang, C.; Zhang, J.; Zhang, Q.; Wu, G.; Wang, Y.; Ma, J. Novel vanillic acid-based poly(ether–ester)s: From synthesis to properties. *Polym. Chem.* **2015**, *6*, 797–804. [CrossRef]

Screening and Comparative Characterization of Microorganisms from Iranian Soil Samples Showing ω-Transaminase Activity toward a Plethora of Substrates

Najme Gord Noshahri [1], Jamshid Fooladi [1,*], Christoph Syldatk [2], Ulrike Engel [2], Majid M. Heravi [3], Mohammad Zare Mehrjerdi [4] and Jens Rudat [2,*]

[1] Department of Biotechnology, Faculty of Biology Science, Alzahra University, 1993893973 Tehran, Iran
[2] BLT_II: Technical Biology, Karlsruhe Institute of Technology (KIT), Fritz-Haber-Weg 4, 76131 Karlsruhe, Germany; christoph.syldatk@kit.edu (C.S.); ulrike.engel@kit.edu (U.E.)
[3] Department of Chemistry, School of Sciences, Alzahra University, 1993891176 Tehran, Iran; mmheravi@alzahra.ac.ir
[4] Higher Education Complex of Shirvan, 9468194477 Shirvan, Iran; mzarem@um.ac.ir
[*] Correspondence: jfooladi@alzahra.ac.ir (J.F.); jens.rudat@kit.edu (J.R.); Tel.: +98-21-880-440-40 (J.F.); +49-721-608-484-28 (J.R.)

Abstract: In this study, soil microorganisms from Iran were screened for ω-transaminase (ω-TA) activity based on growth on minimal media containing (*rac*)-α-methylbenzylamine (rac-α-MBA) as a sole nitrogen source. Then, for the selection of strains with high enzyme activity, a colorimetric o-xylylendiamine assay was conducted. The most promising strains were identified by 16S rDNA sequencing. Five microorganisms showing high ω-TA activity were subjected to determine optimal conditions for ω-TA activity, including pH, temperature, co-solvent, and the specificity of the ω-TA toward different amine donors and acceptors. Among the five screened microorganisms, *Bacillus halotolerans* turned out to be the most promising strain: Its cell-free extract showed a highly versatile amino donor spectrum toward aliphatic, aromatic chiral amines and a broad range of pH activity. Transaminase activity also exhibited excellent solvent tolerance, with maximum turnover in the presence of 30% (*v/v*) DMSO.

Keywords: ω-transaminase; α-methylbenzylamine; biocatalysis; chiral amine; biotransformation

1. Introduction

Chiral amines as building blocks are prevalent in pharmaceuticals and chemical industries [1]. For instance, chiral amines are present in roughly 40% of FDA approved pharmaceuticals [2,3]. Furthermore, chiral β-amino acids (β-aas) have attracted significant attention recently [4]. They are critical building blocks in highly important medicines and natural compounds such as maraviroc (HIV remedy) [5,6], taxol (cytostatic drug against breast and ovarian cancer), jaspamide (insecticidal and antifungal agent), theopalauamide (antifungal compound), and dolastatin (antitumor agent) [7]. β-aas are also employed in the production of peptidomimetics that are stable against degradation by proteolytic enzymes [4,8].

The significant challenge in industry is to use optically pure amines and amino acids from cost-effective production [1,9]. For the mass production of such compounds, nowadays biocatalysts are frequently used to comply with green chemistry as well be economically feasible [1,5]. Among various enzymatic methods for producing chiral amine compounds, ω-transaminases (ω-TAs) are promising enzymes to synthesize chiral amines. They show high substrate spectra,

enantioselectivity, and no requirement to regenerate external cofactors [10–13]. They belong to pyridoxal-5′–phosphate-(PLP)-dependent enzymes that transfer an amino group from an amino donor to amino accepters such as prochiral ketones or aldehydes [14,15]. In addition, ω-TAs can be used for providing optically enriched chiral amino acids via the kinetic resolution of racemic amines or asymmetric synthesis from prochiral ketones [12].

A restricted substrate scope for large molecules [16], as well as rather low enzyme stability under extreme reaction conditions associated with pH, temperature [17], and organic solvents [5,13] have been major hurdles of using ω-TAs. Thus, these limitations should be circumvented by additional screening approaches to fill the toolbox of industrially applicable enzymes.

Basically, there are different approaches to discover new ω-TAs: metagenomic screenings [18,19], site-directed mutagenesis of well-established enzymes [20], and screening new biocatalysts by identifying novel microorganisms from as-yet unexplored screening sites [21–23].

Results of several studies indicated that Iranian landscapes contain highly diverse microbial storages due to considerable variety in geographical features, including mountain range peaks with volcano activity, sandy deserts (Dasht-e Lut is one of the hottest points of the Earth), rivers [24], forest areas, hot springs, glaciers, and frequent snowing. Iran also has two coastlines to its north and south. Besides the extremely versatile environments in Iran, industries such as petroleum (oil well) and petrochemical units releasing chemical substances cause influences on the microorganism populations of the above-mentioned country. With such a variety of environments, Iran's soil can be considered a rich source for screening microorganisms, which are able to produce special extremophile biocatalysts [25–28].

The purpose of this study is to identify novel ω-TA from Iranian soil microorganisms to be examined as a catalyst in the synthesis of enantiopure β-amino acid (β-aa) and other chiral amino compounds. Relevant properties of this novel ω-TA, such as the effect of temperature, pH, and solvent and substrate spectra were also investigated.

2. Results and Discussion

2.1. Microorganisms Screening and Identification

In this study, 42 strains were isolated by enrichment culture against (*rac*)-α-methylbenzylamine (rac-MBA) as a sole nitrogen source. Induction and amount of ω-TA is affected by nitrogen source in media. In complex media such as Luria-Bertani (LB) medium, some of the microorganisms produce enzymes with lower activity [29]. Therefore, minimal media (MIM) containing (*rac*)-α-MBA as a model amine compound [22,23] was used in the experiment.

They were subsequently validated through *o*-xylylenediamine (*o*-XDA) assay, with the color changing from yellow to black after 4 h. *o*-XDA after deamination undergoes spontaneous cyclization, tautomerization, and irreversible polymerization to form a black product [1]. Five strains showing ω-TA activity were found to be promising. Four of these strains were identified by 16S rDNA sequencing to be members of the genus *Bacillus* close to *Bacillus halotolerans* (BaH) (99.93% similarity) and *Bacillus endophyticus* (BaE) (99.76% similarity). Two subspecies belong to *Bacillus subtilis* with 99.92% and 99.93% similarity to *stercoris* (BaS) and *inaquosorum* (BaI), respectively. The fifth bacterium is associated to *Rhizobium radiobacter* (RhZ) with 99.77% similarity (Table 1). All of the strains except BaE were isolated from first method of screening.

Table 1. Results of 16SrDNA sequencing.

Isolate	Closest Relative in GenBank	% Similarity/Sequence Length (bp)
BaH	*Bacillus halotolerans* (LPVF01000003)	99.9/1388
BaE	*Bacillus endophyticus* (AF295302)	99.7/1273
BaS	*Bacillus subtilis* subsp. *stercoris* (JHCA01000027)	99/1385
BaI	*Bacillus subtilis* subsp. *inaquosorum* (AMXN01000021)	99.9/1406
RhZ	*Rhizobium radiobacter* (AJ389904)	99.4/1338

Number in brackets is GenBank accession number.

2.2. Enzymatic Properties and Substrate Specificity

Appropriate temperature, pH, and organic solvents play important roles in ω-TA activity [5,30]. To survey such effects, the optimum reaction condition against these three factors as investigated in the screened bacteria. Several studies have been performed by applying cell and crude extract to convert different substrates [2,18,21,29,31–33], so instead of using high-priced purified protein, we applied cell-free extracts to optimize reaction parameters and to investigate substrate scope.

The relative values of ω-TA activity were calculated based on producing acetophenone as a product of *(S)*-MBA deamination. The highest activity was considered as 100%.

2.2.1. Effect of Temperature on ω-TA Activity

The cell-free extracts were incubated at defined temperatures (25–65 °C) for one hour. The crude extracts of BaH, RhZ, and BaS shared similar expected mesophilic temperature profiles between 30 and 65 °C, with highest activity at 35 °C, which is reported for other ω-TAs [21,34]. They exhibited reasonable activity of around 70% up to 45 °C. Among all of them, 40–50% of activity remained at 50 °C, which decreased to approximately 10–20% at temperatures above 60 °C (Figure 1a).
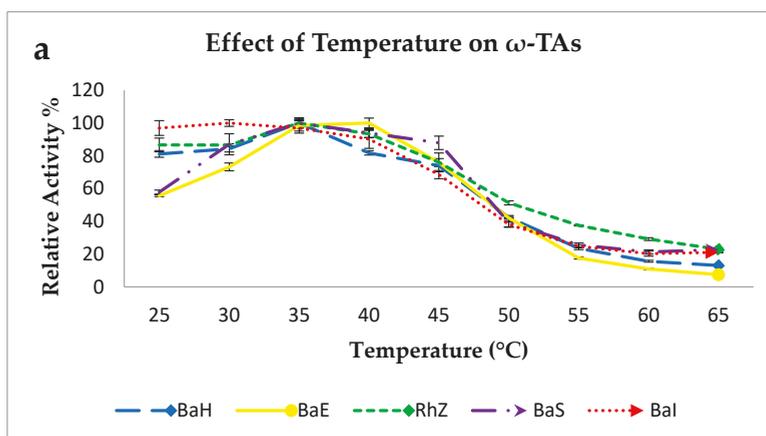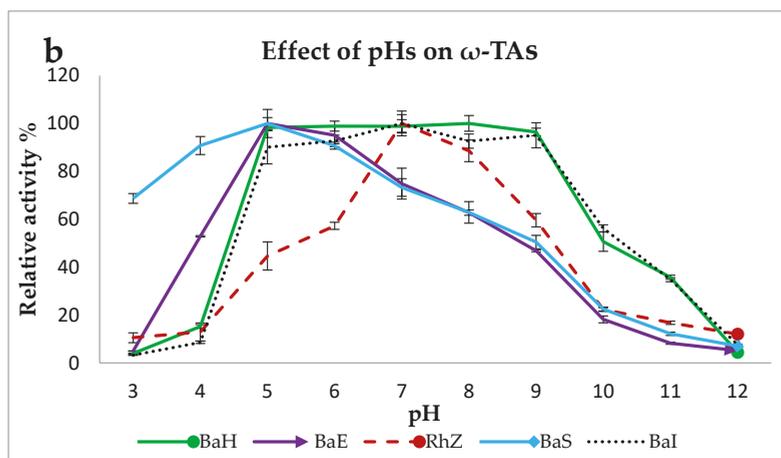

Figure 1. *Cont.*

Figure 1. The effect of temperature (**a**) and pH (**b**) on the amination of pyruvate and (*S*)-MBA as amino donor using 1 mg/mL of cell-free extracts of screened strains: *Bacillus halotolerans* (BaH), *Bacillus endophyticus* (BaE), *Rhizobium radiobacter* (RhZ), *Bacillus subtilis* subsp. s*tercoris* (BaS) and *Bacillus subtilis* subsp. i*naquosorum* (BaI).Reaction conditions are given in the method section. Formation of acetophenone was detected by HPLC at 254 nm. The value of 100% corresponded to the highest activity. Error bars represent ± standard deviation.

For ω-TABaE and ω-TABaI, the optimum reaction temperature was 40 °C and 30 °C, respectively. No significant change of activity was observed at 35 °C. Thus, for surveying the effect of pH and solvent on enzyme activity on all crude extracts, the temperature was maintained at 35 °C. Additionally, ω-TABaI showed approximately the same activity at temperatures between 20 and 30 °C (Table S1).

2.2.2. Effect of pH on ω-TA Activity

In contrast to other ω-TAs that usually show their maximum activity at slightly alkaline pH [12,21,31], ω-TABaI and ω-TABaH exhibited an almost constant activity in the pH range of 5 to 9. Surprisingly, for ω-TABaS and ω-TABaE, the highest enzyme activity was observed under acidic conditions (pH 5). Notably, 70% of ω-TABaS activity remained at pH 3, Furthermore, ω-TARhz preferred neutral conditions (Figure 1b). No enzyme activity was observed at pH 12. In the literature data, there is no report concerning ω-TA activity in acidic conditions. ω-TA activity at different pHs was mostly tested in pH range from 6 to 11 [4,17,18].

2.2.3. Effect of Organic Solvents on ω-TA Activity

Organic solvents are often added to ω-TA reactions to increase substrate solubility. However, this mostly causes negative effects on enzyme activity [10]. Therefore, the examination of the tolerance of the applied enzyme to different solvents is highly required [5]. To analyze whether solvents affect enzyme activity, four different solvents with a 0%–20% with the exception of DMSO which increased up to 40% *v/v* concentration were tested. The ω-TA activity in aqueous buffer without addition of organic solvents was set as 100%. The results were expressed as relative activity in %.

Dimethyl sulfoxide (DMSO) is a common solvent that was used in concentrations ranging from 5%–20% *v/v* in different projects [5]. DMSO as co-solvent led to the highest activity, followed by methanol, N,N-Dimethylformamide (DMF), and isopropanol. In some cases (BaI and Rhz), the reactions without any co-solvent showed better activity than methanol. It is remarkable to mention that ω-TABaH activity was enhanced threefold with a concentration of DMSO up to 30%. The results again suggest that DMSO is the most suitable organic solvent for catalytic activation of ω-TAs. However,

methanol in some cases also increased enzyme activity (Figure 2). Methanol was mentioned as a second co-solvent after DMSO for enhancing ω-TA activity in Pawer et al.'s experiment [2].
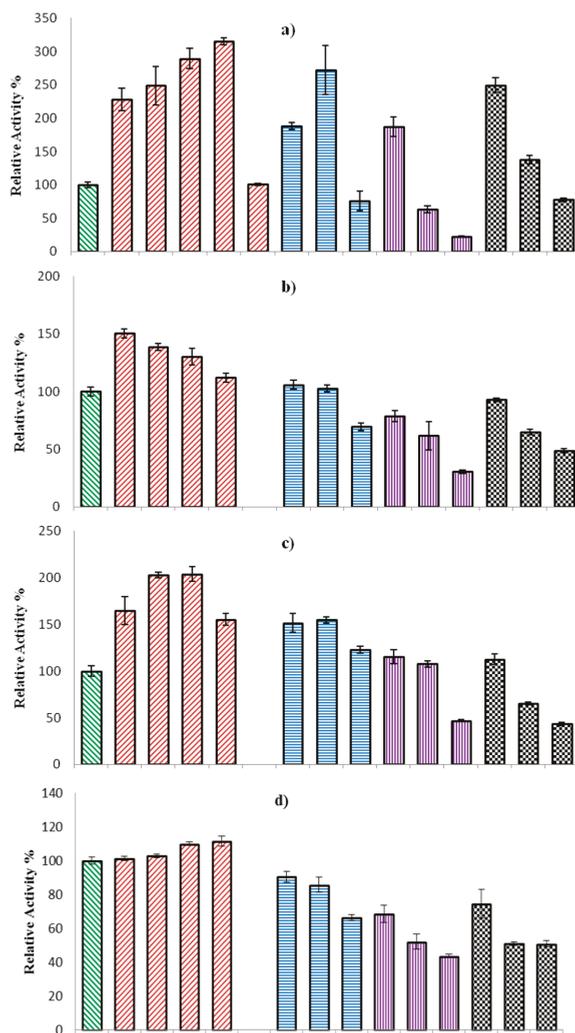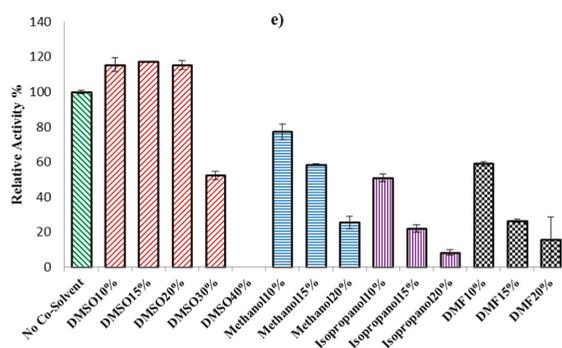

Figure 2. *Cont.*

Figure 2. Effect of co-solvent on ω-TA activity of screened strains. (**a**) BaH; (**b**) BaE; (**c**) BaS; (**d**) RhZ; (**e**) BaI. The value of 100% corresponds to enzyme activity in the no co-solvent condition. 1 mg/mL cell-free extract was added to react under the following conditions: (*S*)-MBA (10 mM), pyruvate (10 mM), PLP (0.1 mM), and co-solvent (0–30% *v/v*) in Britton–Robinson buffer (pH 7) at 35 °C. Formation of acetophenone was detected by HPLC at 254 nm. Error bars represent ± standard deviation.

2.2.4. Substrate Specificity and Enantioselectivity

The substrate spectrum of all strains was investigated in order to find the most suitable amino donor and acceptor.

Regarding amino donors, 28 substrates, including aliphatic and aromatic amino compounds were reacted to pyruvate as an amino acceptor to investigate amino donor specificity, and alanine generation was monitored by the HPLC technique. Racemates as well as the pure, *(S)* and *(R)* enantiomers (if available) of amino donors were chosen to explore the enantioselectivity of ω-TA. The results of the substrate screening are illustrated in Table 2. HPLC analysis proved that all enzymes show predominant formation of an *(S)*-configuration. This is in accordance with most previously reported publications in the literature indicating that *(R)*-stereoselectivity are rather rare in wild-type microorganisms [35]. All cell-free extracts showed high activity against *(S)*- α-methylbenzylamine (*S*-MBA) **1s**.

BaH showed considerable activity toward a wide range of substrates, including *(S)*, *(R)*, and racemates of aromatic and aliphatic amino donors. Although this cell-free extract also showed activity toward some tested *(R)*-enantiomers (**1r, 2r, 4r, 5r, 7r, 8r, 13r**), (S)-selectivity appeared to be not so strict in this strain. Notably, it is difficult to conclude that BaH showed *(R)*-selectivity since in this study, crude extract was applied. There is a possibility that (*R*) converted to (*S*)-enantiomer via enzymatic racemization [36], and then the latter was subjected to ω-TA activity. On the other hand, (*R*)-preference towards **13r** in compare to **13s** leads to another hypothesis, in which both R and S transaminases are present in crude extract.

While the cell-free extract of BaH revealed varying enantiopreference toward all tested amino donors to different degrees, the substrate specificity was found to resemble what has previously been reported by Mathew et al. [5,30], as a preference to react with aromatic β-amino acids is clearly visible [5]. *(S)*-MBA **1s** is converted with the highest activity, which therefore was set as 100%. Although a few ω-TAs are described with activity toward aromatic β-amino acids (56), it is remarkable that BaH uses β-phenylalanine (**13, 13s, 13r**) and (*rac*)-β-homophenylalanine (**14**) as an amino donor; both of these were inert for crude extraction from other bacteria in this experiment, which implies the presence of a broad binding pocket in the active site of ω-TABaH. ω-TA activity against **14** has been reported by Mathew et al. for ω-TABG, with 50% relative activity in comparison with **13** [30]. A recent study by Buß et al. explained the lack of ω-TA activity against **14** (by an ω-TA that was highly active against **13**) by sterice hindrance caused by the additional carbon atom located between the amino group and phenyl group [36]. Several substituted derivatives of β-phenylalanine are converted as well.

Table 2. Amino donor screened assays were carried out in Britton–Robinson buffer, pH 7 with amino donor (10 mM), pyruvate (10 mM), PLP (0.1 mM), and 15% (*v/v*) DMSO. The reaction was initiated by adding crude extract to a final concentration of 1 mg protein per ml. Reactions were performed at 35 °C, 600 rpm for 20 h. Relative activity was measured by HPLC detection of alanine formation at 338 nm. All reactions were conducted in triplicate, and each cell-free extract was separately analyzed for alanine content as a control.

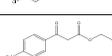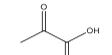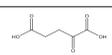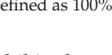| Amino Donor			Relative Activity%				
			BaH	BaE	BaS	BaI	RhZ
1	*(rac)*		53	48	45	60	36
1s	*(S)*	α-methylbenzylamine	100	100	100	100	100
1r	*(R)*		23	0	0	6	0
2	*(rac)*		19	0	0	0	0
2s	*(S)*	3-amino-3-(4-chlorophenyl) propionic acid	20	0	0	0	0
2r	*(R)*		16	0	0	0	0
3	*(rac)*	3-amino-3-(3,4-dichloro-phenyl)propionic acid	20	0	0	0	0
4	*(rac)*	3-amino-3-(4-nitrophenyl)propionic acid	51	0	0	0	0
4r	*(R)*		19	0	0	0	0
5r	*(R)*	(R)3-amino-3-(2-nitro-phenyl)-propionic acid	16	0	0	0	0
6	*(rac)*	3-amino-3-(4-methoxyphenyl) propionic acid	21	0	0	0	0
6s	*(S)*		17	0	0	0	0
7s	*(S)*	3-amino-3-(4-hydroxy-phenyl)propionic acid	18	0	0	0	0
7r	*(R)*		15	0	0	0	0
8	*(rac)*	3-amino-3-(4-bromophenyl)-propionic acid	21	0	0	0	0
8r	*(R)*		20	0	0	0	0
9	(rac)	3-amino-3-(4-fluorophenyl)propionic acid	43	0	0	0	0
9s	*(S)*		16	0	0	0	0
10	*(rac)*	Methyl(-3-amino-3-phenyl propanoate hydrochloride)	11	0	0	0	0
11	*(rac)*	3-amino-3-(4-isopropylphenyl) propionic acid	20	0	0	0	0
12	*(rac)*	1-aminocyclopropane-1-carboxylic acid	17	17	0	30	0
13	*(rac)*		17	0	0	0	0
13s	*(S)*	β-phenylalanine	16	0	0	0	0
13r	*(R)*		22	0	0	0	0
14	*(rac)*	β-Homophenylalanine	16	0	0	0	0
15		Isopropylamine	20	0	0	21	0
16		Sec-butylamine	25	7	0	27	0
17s	*(S)*	3-amino butyric acid	91	59	54	20	135

The activity against *(S)*-MBA was set as 100. An amount under 1% was considered as zero activity.

BaE presented only *(S)*-selectivity in relation to tested substrates. Similar to BaH, the strain's crude extract exhibited activity toward **1s**, almost double that of **1**. BaE and BaH share the same activity toward **12**, while BaI showed almost double the activity in comparison with those. The relative activity of ω-TA of BaI toward *(rac)*-MBA **1** exceeded 60%, which might be related to the presence of racemase in the crude extract of bacteria [37,38].

Crude extracts of RhZ and BaS displayed *(S)*-selectivity toward the tested amino donors with similar preference profiles. RhZ showed more activity toward **17s** as an aliphatic substrate than MBA **1s,** which was not observed for other screened strains.

Amino acceptors play a critical role in ω-TA activity [30]. Eight ketones were reacted with *(S)*-MBA to explore the amino acceptor specificities by HPLC detection of acetophenone production (Table 3). Aliphatic ketones such as sodium pyruvate **7b** and α-ketoglutaric acid **8b** showed high conversion, whereas most aromatic ketones were not accepted by the investigated enzymes.

Table 3. Amino acceptor screened assays were carried out in Britton–Robinson buffer, pH 7 with *(S)*-MBA (10 mM), amino acceptor (10 mM), PLP (0.1 mM), and 15% (*v/v*) DMSO. The reaction was started by the addition of crude extract to a final concentration of 1 mg/mL in the following conditions: 35 °C, 600 rpm for 20 h. Relative activity was measured by HPLC detection of acetophenone formation at 254 nm. All reactions were conducted in triplicate, and each cell-free extract was separately analyzed as a control.

	Amino Acceptors		Relative Activity%				
			BaH	BaE	BaS	BaI	RhZ
1b	Ethyl benzoylacetate		40.95	0	20.84	14.87	0
2b	Ethyl 3-(4-methoxyphenyl)-3-oxopropanoate		0	0	0	60.48	0
3b	Sodium 3-oxo-3-phenylpropanoate		0	0	0	0	0
4b	Ethyl(4-flurobenzoyl)acetate		0	0	0	0	0
5b	Ethyl(4-chlorobenzoyl)acetate		0	0	0	0	0
6b	Ethyl 4-nitrobenzoylacetate		0	0	0	0	0
7b	Pyruvate		100	100	43.73	100	36.39
8b	α-ketoglutarate		17.09	7.31	100	0	100
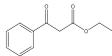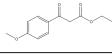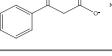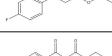
In each column, the highest activity was defined as 100%. Activity less than 1% was considered zero.

Cell-free extract of BaH, BaS, and BaI exhibited some activity toward bulkier substrates, including **1b** or **2b**. It is supposed that β-amino acid esters were produced by using these acceptors, but this needs to be investigated in further experiments. BaH, BaS, and BaI showed activity toward **1b** as amino acceptor. BaI was the only extract to show activity toward acceptor **2b**.

The substrate scope of ω-TA is determined by the active site, which consists of a large and a small binding pocket. The former can accommodate bulky residues, including aryl groups, whereas the latter is loaded by a small group such as a methyl substituent [14]. So most ω-TAs have a limitation in accepting bulky substrates without applying protein engineering. Voss et al. reported that after double mutation on ω-TA 3FCR from *Ruegeria* sp. TM1040, this enzyme gained activity toward bulkier substrates such as phenylpropylamine and phenylbutylamine [16]. Strain BaH showed a good

ability to convert bulky substrates and showed different substrate spectra with known ω-TAs [15,35]. Purification and further investigation of this enzyme is intended for further studies.

3. Materials and Methods

3.1. Chemicals

All solvents and chemicals used in this study were purchased from Sigma-Aldrich (St Louis, MO, USA), Carl Roth GmbH (Karlsruhe, Germany), and Peptech (Burlington, MA, USA).

3.2. Screening and Identification

3.2.1. Enrichment of Microorganisms on (rac)-α-MBA as Sole Nitrogen Source

Due to the role of transaminase in nitrogen metabolism [29], areas enriched with nitrogen compounds were chosen as potential sites for screening novel ω-transaminase. Soil samples (2 g) from a petrochemical site, oil well, and an agriculture field in Iran were separately suspended in 200 mL sterile minimal medium (MIM) in 500 mL Erlenmeyer flasks containing 10 mM (*rac*)-α-MBA, 100 mM glycerol, 1 g/L MgSO$_4$ .7H$_2$O, 0.02 mg/L H$_3$BO$_3$, 0.2 mM CaCl$_2$, 0.1 mg/L MnSO$_4$. 4H$_2$O, 0.1 mg/L CuSO$_4$ .5H$_2$O, 0.1 mg/L NiSO$_4$. 6H$_2$O, 2.0 mg/L NaMoO$_4$, 0.05 mg/L CoCl$_2$, 0.1 mg/L ZnCl$_2$, 4 mg/L FeSO$_4$. 7H$_2$O, and potassium phosphate buffer (50 mM, pH 7.0) [39]. FeSO$_4$ and (*rac*)-α-MBA were separately sterile filtrated and added to the medium. Autoclaving causes precipitation of ferric hydrate [40]. Each sample was prepared in duplicate. Enrichment went through two strategies:

In the first method, Erlenmeyer flasks were incubated for 24 h at 180 rpm at 35 °C (depending on the original temperature of the soil samples). After 24 h of incubation, a 100 µL sample was transferred to 20 mL of fresh minimal medium and incubated for 24 h under the same conditions as described before. This serial dilution was repeated 3 times. After that, the culture was diluted with PBS buffer (10: 90 µL) and poured to MIM agar plates (mineral medium, 10 mM (*rac*)-α-MBA with 1.5% agar) and incubated at 35 °C. As soon as colonies were visible, single colonies were streaked on new MIM agar plates to obtain uniform colonies under the same temperature.

In the second method, Erlenmeyer flasks were transferred to a dark place, without shaking, at 35 °C. During incubation, 1 mL of concentrated MIM (5X), 0.5% *v/v* glycerol, and 10 mM (*rac*)-α-MBA were added to the flasks every week. After 3 months, the culture broth was centrifuged (1000 *g*, 2 min) and 100 µL of supernatant was spread on MIM agar plates. The colonies were isolated after 3 days of incubation at 35 °C. This was continued until the isolation of single colonies was achieved.

3.2.2. Selection of the Most Promising Strains

To identify whether isolated strains use (*rac*)-α-MBA as a sole nitrogen source or atmospheric nitrogen, *o-Xylylenediamine* was applied as a smart amino donor. Each colony was cultivated in a shake flask containing 20 mL MIM medium with 12 mM (*rac*)-α-MBA for three days at 35 °C at 120 rpm. The cells were harvested by centrifugation at 6000× *g* for 10 min at 4 °C and resuspended in HEPES buffer (50 mM, pH 7.5). The cell concentration was adjusted to around 20 mg/mL by measuring dry cell mass based on Buß et al. [36]. The cells were frozen at −80 °C and thawed at room temperature for cell disruption.

Reaction was started by adding 50 µL of whole cells to 150 µL reaction solution containing 7.5 mM *o-Xylylenediamine,* 5 mM pyruvate, 1 mM PLP, and 10% DMSO in HEPES buffer (50 mM, pH 7.5) [41] in a 96-well plate. Each reaction was conducted in triplicate. The plate was incubated overnight at 35 °C, 150 rpm.

3.2.3. Identification of Bacteria

The positive strains were grown separately overnight in LB broth at 35 °C, 150 rpm. DNA was extracted by ZR soil Microbe DNA Kit^{TM}. PCR was done with a Q5 high-fidelity PCR kit (NEB,

Germany). Universal Primers (27F and 1492R) were applied to amplify 1.5 kb 16S rDNA fragments. PCR was conducted following the manufacturer's instructions for Q5 polymerase. Gene sequencing was performed by Eurofins Company (Ebersberg, Germany). The resulting 16S rRNA gene sequences were compared with available gene sequences in the EZ bioCloud database [42].

3.3. Enzymatic Properties and Substrate Specificity

3.3.1. Preparation of Crude Extract

The inoculum (3 mL) was prepared by picking a single colony and cultivating it for three days in MIM with 12 mM (*rac*)-α-MBA. The medium (400 mL) was inoculated by the addition of 1% (*v/v*) of the latter culture into 1 L shaking flasks containing MIM with 0.5 g/L yeast extract and 12 mM (*rac*)-α-MBA overnight at 150 rpm and 35 °C. Cells were harvested before reaching stationary phase by centrifugation at 6000× *g* for 10 min at 4 °C. After washing the pellet with sodium phosphate buffer (50 mM, pH 7), cells were resuspended in 10 mL lysis buffer (sodium phosphate (50 mM, pH 7), 0.1 μM PLP, 100μg/mL lysozyme]. Afterwards, they were incubated for 1 h at room temperature, followed by sonification (3 cycles: 30 s pulse, 20 s pause; 60% amplitude) on ice. Subsequently, the cell debris was removed by centrifugation (20,000 rpm, 4 °C, 15 min) in a JA-30–50 rotor (Coulter–Beckman centrifuge). The protein concentration of crude extract was determined by using the Roti®-Quant universal kit (Carl Roth, Karlsruhe, Germany) following the manufacturer's instructions. Eventually, they were mixed with 15% glycerol and preserved in −80 °C for further analysis.

3.3.2. Effect of Temperature, pH, and Co-Solvent

The effects of temperature and pH on enzyme activity were examined at different temperatures (25–65 °C at pH7) and various pHs (3–12 at 35 °C) by using Britton–Robinson buffer (0.04 M H_3BO_3, 0.04 M H_3PO_4, 0.04 M CH_3COOH) containing an amino donor (10 mM), amino acceptor (10 mM), PLP (0.1 mM), and DMSO (10%). The total reaction volume was set to 0.25 mL. The reactions were initiated by adding crude extract of enzyme (~1 mg/mL).

To study the effect of solvent on ω-TA activity, reactions were carried out applying various organic solvents with 0–20% (*v/v*) (for DMSO 0–40%). (rac)-α-MBA (10 mM) was used as amino donor and pyruvate (10 mM) as amino acceptor in the presence of PLP (0.1 mM) and cell-free extract (~1 mg/mL). The reactions were carried out in Britton–Robinson buffer (pH 7) at 35 °C as described above.

Each reaction was conducted in triplicate and incubated (ThermoMixer, Eppendorf) at 600 rpm for 1 h. As a control, the reaction was conducted without adding enzyme. Furthermore, the amount of acetophenone in crude extract of enzyme was also evaluated. Reactions were stopped by heating to 95 °C for 5 min. After centrifugation, the supernatant was analyzed by HPLC to detect produced acetophenone according to Section 3.4. Results are shown as relative activity.

3.3.3. Substrate Specificity and Enantioselectivity

The reactions were carried out according to Section 3.3.2. Various amino donors (10 mM) listed in Table 1 were tested using sodium pyruvate (10 mM) as amino acceptor in the Britton–Robinson buffer (pH 7) and 15% (*v/v*) DMSO. The reaction was carried out at 35 °C, 600 rpm, for 20 h (ThermoMixer, Eppendorf). The produced alanine was analyzed by HPLC according to Section 3.4.

Different amino acceptors (10 mM) were tested with (S)-α-MBA as amino donor (10 mM) in Britton–Robinson buffer, pH 7, at 35 °C, 600 rpm, for 20 h. The production of acetophenone was analyzed by HPLC according to Section 3.4.

Each reaction was conducted in triplicate, and the average of three independent reactions was used for evaluation. Enzyme inactivation was applied by heating to 95 °C for 5 min. The supernatant was used for analysis after centrifugation. In addition, the amounts of alanine and acetophenone were measured in enzyme crude extracts as blanks.

3.4. HPLC Analytics

All samples were analyzed by Agilent 1100 series HPLC system (Santa Clara, CA, USA). For the analysis of alanine, derivatization by using ortho-phthalaldehyde was carried out according to Brucher et al. [43] and Buß et al. [44], with an automated precolumn derivatization. A reversed-phase C_{18} column (150 × 4.6 mm HyperClone 5 μm ODS, Phenomenex, Germany) with isocratic elution with 35% (*v/v*) methanol and 65% (*v/v*) sodium phosphate buffer (40 mM, pH 6.5) at a flow rate of 0.8 mL min^{-1} and detection at 338 nm at 25 °C was used.

The acetophenone concentration in the samples was determined chromatographically by isocratic elution with acetonitrile/water (50/50, *v/v*) at flow rate of 0.6 mL min^{-1} with UV detection at 254 nm [45], at 25 °C applying a C_{18} Hypersil-keystone column (250 × 4.6 mm 5 μ Hypersil). The injection volume was adjusted to 1 μL.

4. Conclusions

Our motivation for screening novel microorganisms exhibiting transaminase activity resulted from several challenges for the industrial application of ω-TAs, e.g., lacking activity at acidic pH, high temperature, unnatural substrates, (*R*)-configured molecules, and elevated concentrations of organic solvents (needed due to low substrate solubility in aqueous systems). We succeeded in finding promising new strains from a high variety of Iranian soil samples by enrichment culture using (*rac*)-α-methylbenzylamine (α-MBA) as a sole nitrogen source.

In particular, a *Bacillus halotolerans* (BaH) strain was isolated from a petroleum refinery, exhibiting ω-TA activity using a broad spectrum of amino donors over a pH range of 5–9 at elevated concentrations of DMSO and other organic solvents up to 30% (*v/v*). Two other strains isolated from an agricultural field (*Bacillus endophyticus*, BaE) and an oilfield (*Bacillus subtilis*, BaS) showed the highest ω-TA activity against α-MBA at pH 5 with 70% remaining activity at pH 3 (BaS), whereas ω-TAs are usually described to prefer slightly alkaline conditions. The enzymes dedicated to these extraordinary activities will be purified and subjected to in-depth studies for application-technical characterization.

References

1. Gomm, A.; Lewis, W.; Green, A.P.; O'Reilly, E. A New Generation of Smart Amine Donors for Transaminase-Mediated Biotransformations. *Chem. A Eur. J.* **2016**, *22*, 12692–12695. [CrossRef]
2. Pawar, S.V.; Hallam, S.J.; Yadav, V.G. Metagenomic Discovery of a Novel Transaminase for Valorization of Monoaromatic Compounds. *RSC Adv.* **2018**, *8*, 22490–22497. [CrossRef]
3. Aleku, G.A.; France, S.P.; Man, H.; Mangas-Sanchez, J.; Montgomery, S.L.; Sharma, M.; Leipold, F.; Hussain, S.; Grogan, G.; Turner, N.J. A Reductive Aminase from Aspergillus Oryzae. *Nat. Chem.* **2017**, *9*, 961. [CrossRef]
4. Kim, G.H.; Jeon, H.; Khobragade, T.P.; Patil, M.D.; Sung, S.; Yoon, S.; Won, Y.; Choi, I.S.; Yun, H. Enzymatic Synthesis of Sitagliptin Intermediate Using a Novel ω-Transaminase. *Enzym. Microb. Technol.* **2019**, *120*, 52–60. [CrossRef]

5. Mathew, S.; Nadarajan, S.P.; Chung, T.; Park, H.H.; Yun, H. Biochemical Characterization of Thermostable ω-Transaminase from Sphaerobacter Thermophilus and its Application for Producing Aromatic β-and γ-Amino Acids. *Enzym. Microb. Technol.* **2016**, *87*, 52–60. [CrossRef]

6. Haycock-Lewandowski, S.J.; Wilder, A.; Åhman, J. Development of a Bulk Enabling Route to Maraviroc (UK-427,857), a CCR-5 Receptor Antagonist. *Org. Process. Res. Dev.* **2008**, *12*, 1094–1103. [CrossRef]

7. Slomka, C.; Zhong, S.; Fellinger, A.; Engel, U.; Syldatk, C.; Bräse, S.; Rudat, J. Chemical Synthesis and Enzymatic, Stereoselective Hydrolysis of a Functionalized Dihydropyrimidine for the Synthesis of β-Amino Acids. *AMB Express* **2015**, *5*, 85. [CrossRef]

8. Weiner, B.; Szymański, W.; Janssen, D.B.; Minnaard, A.J.; Feringa, B.L. Recent Advances in the Catalytic Asymmetric Synthesis of β-Amino Acids. *Chem. Soc. Rev.* **2010**, *39*, 1656–1691. [CrossRef]

9. Turner, N.J.; Truppo, M.D. Biocatalytic Routes to Nonracemic Chiral Amines. In *Chiral Amine Synthesis: Methods, Developments and Applications*; Nugent, T.C., Ed.; Wiley-VCH: Weinheim, Germany, 2010; pp. 431–459.

10. Mathew, S.; Jeong, S.S.; Chung, T.; Lee, S.H.; Yun, H. Asymmetric Synthesis of Aaromatic β-Amino Acids Using ω-Transaminase: Optimizing the Lipase Concentration to Obtain Thermodynamically Unstable β-Keto Acids. *Biotechnol. J.* **2016**, *11*, 185–190. [CrossRef]

11. Malik, M.S.; Park, E.S.; Shin, J.S. Features and Technical Applications of ω-Transaminases. *Appl. Microbiol. Biotechnol.* **2012**, *94*, 1163–1171. [CrossRef]

12. Jiang, J.; Chen, X.; Feng, J.; Wu, Q.; Zhu, D. Substrate Profile of an ω-Transaminase from Burkholderia Vietnamiensis and its Potential for the Production of Optically Pure Amines and Unnatural Amino Acids. *J. Mol. Catal. B Enzym.* **2014**, *100*, 32–39. [CrossRef]

13. Cerioli, L.; Planchestainer, M.; Cassidy, J.; Tessaro, D.; Paradisi, F. Characterization of a Novel Amine Transaminase from Halomonas Elongata. *J. Mol. Catal. B Enzym.* **2015**, *120*, 141–150. [CrossRef]

14. Voss, M.; Das, D.; Genz, M.; Kumar, A.; Kulkarni, N.; Kustosz, J.; Kumar, P.; Bornscheuer, U.T.; Höhne, M. In Silico Based Engineering Approach to Improve Transaminases for the Conversion of Bulky Substrates. *ACS Catal.* **2018**, *8*, 11524–11533. [CrossRef]

15. Rudat, J.; Brucher, B.R.; Syldatk, C. Transaminases for the Synthesis of Enantiopure Beta-Amino Acids. *AMB Express* **2012**, *2*, 11. [CrossRef]

16. Pavlidis, I.V.; Weiß, M.S.; Genz, M.; Spurr, P.; Hanlon, S.P.; Wirz, B.; Iding, H.; Bornscheuer, U.T. Identification of (S)-Selective Transaminases for the Asymmetric Synthesis of Bulky Chiral Amines. *Nat. Chem.* **2016**, *8*, 1076–1082. [CrossRef]

17. Kelly, S.A.; Magill, D.J.; Megaw, J.; Skvortsov, T.; Allers, T.; McGrath, J.W.; Allen, C.C.; Moody, T.S.; Gilmore, B.F. Characterisation of a Solvent-Tolerant Haloarchaeal (R)-Selective Transaminase Isolated from a Triassic Period Salt Mine. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 5727–5737. [CrossRef]

18. Leipold, L.; Dobrijevic, D.; Jeffries, J.W.; Bawn, M.; Moody, T.S.; Ward, J.M.; Hailes, H.C. The Identification and Use of Robust Transaminases from a Domestic Drain Metagenome. *Green Chem.* **2019**, *21*, 75–86. [CrossRef]

19. Ferrandi, E.E.; Previdi, A.; Bassanini, I.; Riva, S.; Peng, X.; Monti, D. Novel Thermostable Amine Transferases from Hot Spring Metagenomes. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 4963–4979. [CrossRef]

20. Genz, M.; Vickers, C.; van den Bergh, T.; Joosten, H.-J.; Dörr, M.; Höhne, M.; Bornscheuer, U. Alteration of the Donor/Acceptor Spectrum of the (S)-Amine Transaminase from Vibrio Fluvialis. *Int. J. Mol. Sci.* **2015**, *16*, 26953–26963. [CrossRef]

21. Kelly, S.A.; Megaw, J.; Caswell, J.; Scott, C.J.; Allen, C.C.; Moody, T.S.; Gilmore, B.F. Isolation and Characterisation of a Halotolerant ω-Transaminase from a Triassic Period Salt Mine and Its Application to Biocatalysis. *ChemistrySelect* **2017**, *2*, 9783–9791. [CrossRef]

22. Shin, J.S.; Kim, B.G. Kinetic Resolution of α-Methylbenzylamine with o-Transaminase Screened from Soil Microorganisms: Application of a Biphasic System to Overcome Product Inhibition. *Biotechnol. Bioeng.* **1997**, *55*, 348–358. [CrossRef]

23. Pavkov-Keller, T.; Strohmeier, G.A.; Diepold, M.; Peeters, W.; Smeets, N.; Schürmann, M.; Gruber, K.; Schwab, H.; Steiner, K. Discovery and Structural Characterisation of New Fold Type IV-Transaminases Exemplify the Diversity of This Eenzyme Fold. *Sci. Rep.* **2016**, *6*, 38183. [CrossRef] [PubMed]

24. Yazdi, A.; Emami, M.H.; Shafiee, S.M. Dasht-E Lut in Iran, the Most Complete Collection of Beautiful Geomorphological Phenomena of Desert. *Open J. Geol.* **2014**, *4*, 249–261. [CrossRef]

25. Ataee, N.; Fooladi, J.; Namaei, M.H.; Rezadoost, H.; Mirzajani, F. Biocatalysts Screening of Papaver Bracteatum Flora for Thebaine Transformation to Codeine and Morphine. *Biocatal. Agric. Biotechnol.* **2017**, *9*, 127–133. [CrossRef]

26. Ghasemi, Y.; Rasoul-Amini, S.; Ebrahiminezhad, A.; Kazemi, A.; Shahbazi, M.; Talebnia, N. Screening and Isolation of Extracellular Protease Producing Bacteria from the Maharloo Salt Lake. *Iran. J. Pharm. Sci.* **2011**, *7*, 175–180.

27. Alghabpoor, S.S.; Panosyan, H.; Trchounian, A.; Popov, Y. Purification and Characterization of a Novel Thermostable and Acid Stable α-Amylase from Bacillus Sp. Iranian S1. *Int. J. Eng. Trans. B Appl.* **2013**, *26*, 815–820. [CrossRef]

28. Shirsalimian, M.; Amoozegar, M.; Sepahy, A.A.; Kalantar, S.; Dabbagh, R. Isolation of Extremely Halophilic Archaea from a Saline River in the Lut Desert of Iran, Moderately Resistant to Desiccation and Gamma Radiation. *Microbiology* **2017**, *86*, 403–411. [CrossRef]

29. Shin, J.S.; Kim, B.G. Comparison of the ω-Transaminases from Different Microorganisms and Application to Production of Chiral Amines. *Biosci. Biotechnol. Biochem.* **2001**, *65*, 1782–1788. [CrossRef]

30. Mathew, S.; Bea, H.; Nadarajan, S.P.; Chung, T.; Yun, H. Production of Chiral β-Amino Acids Using ω-Transaminase from Burkholderia Graminis. *J. Biotechnol.* **2015**, *196*, 1–8. [CrossRef]

31. Guo, F.; Berglund, P. Transaminase Biocatalysis: Optimization and Application. *Green Chem.* **2017**, *19*, 333–360. [CrossRef]

32. Dreßen, A.; Hilberath, T.; Mackfeld, U.; Billmeier, A.; Rudat, J.; Pohl, M. Phenylalanine Ammonia Lyase from Arabidopsis Thaliana (AtPAL2): A Potent MIO-Enzyme for the Synthesis of Non-Canonical Aromatic Alpha-Amino Acids: Part I: Comparative Characterization to the Enzymes from Petroselinum Crispum (PcPAL1) and Rhodosporidium Toruloides (RtPAL). *J. Biotechnol.* **2017**, *258*, 148–157.

33. Schätzle, S.; Höhne, M.; Robins, K.; Bornscheuer, U.T. Conductometric Method for the Rapid Characterization of the Substrate Specificity of Amine-Transaminases. *Anal. Chem.* **2010**, *82*, 2082–2086. [CrossRef]

34. Schätzle, S.; Höhne, M.; Redestad, E.; Robins, K.; Bornscheuer, U.T. Rapid and Sensitive Kinetic Assay for Characterization of ω-Transaminases. *Anal. Chem.* **2009**, *81*, 8244–8248. [CrossRef] [PubMed]

35. Koszelewski, D.; Tauber, K.; Faber, K.; Kroutil, W. ω-Transaminases for the Synthesis of Non-Racemic α-Chiral Primary Amines. *Trends Biotechnol.* **2010**, *28*, 324–332. [CrossRef] [PubMed]

36. Buß, O.; Dold, S.M.; Obermeier, P.; Litty, D.; Muller, D.; Grüninger, J.; Rudat, J. Enantiomer Discrimination in β-Phenylalanine Degradation by a Newly Isolated Paraburkholderia Strain BS115 and Type Strain PsJN. *AMB Express* **2018**, *8*, 149. [CrossRef] [PubMed]

37. Conti, E.; Stachelhaus, T.; Marahiel, M.A.; Brick, P. Structural Basis for the Activation of Phenylalanine in the Non-Ribosomal Biosynthesis of Gramicidin S. *EMBO J.* **1997**, *16*, 4174–4183. [CrossRef] [PubMed]

38. Yamada, M.; Kurahashi, K. Further Purification and Properties of Adenosine Triphosphate-Dependent Phenylalanine Racemase of Bacillus Brevis Nagano. *J. Biochem.* **1969**, *66*, 529–540. [CrossRef]

39. Yun, H.; Lim, S.; Cho, B.K.; Kim, B.G. ω-Amino Acid: Pyruvate Transaminase from Alcaligenes Denitrificans Y2k-2: A New Catalyst for Kinetic Resolution of β-Amino Acids and Amines. *Appl. Environ. Microbiol.* **2004**, *70*, 2529–2534. [CrossRef]

40. Temple, K.L.; Colmer, A.R. The Autotrophic Oxidation of Iron by a New Bacterium: Thiobacillus Ferrooxidans. *J. Bacteriol.* **1951**, *62*, 605.

41. Buß, O.; Voss, M.; Delavault, A.; Gorenflo, P.; Syldatk, C.; Bornscheuer, U.; Rudat, J. β-Phenylalanine Ester Synthesis from Stable β-Keto Ester Substrate Using Engineered ω-Transaminases. *Molecules* **2018**, *23*, 1211. [CrossRef]

42. Yoon, S.H.; Ha, S.M.; Kwon, S.; Lim, J.; Kim, Y.; Seo, H.; Chun, J. Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 1613. [PubMed]

43. Brucher, B.; Rudat, J.; Syldatk, C.; Vielhauer, O. Enantioseparation of Aromatic β³-Amino Acid by Precolumn Derivatization with o-Phthaldialdehyde and N-Isobutyryl-l-Cysteine. *Chromatographia* **2010**, *71*, 1063–1067. [CrossRef]

44. Buß, O.; Muller, D.; Jager, S.; Rudat, J.; Rabe, K.S. Improvement in the Thermostability of a β-Amino Acid Converting ω-Transaminase by Using FoldX. *ChemBioChem* **2018**, *19*, 379–387. [CrossRef] [PubMed]

45. Gao, S.; Su, Y.; Zhao, L.; Li, G.; Zheng, G. Characterization of a (R)-Selective Amine Transaminase from Fusarium Oxysporum. *Process Biochem.* **2017**, *63*, 130–136. [CrossRef]

MDPI