biology and life sciences forum

MDPI

Proceeding Paper

Comparative Analysis of RuBisCO Evolution and Intrinsic Differences: Insights from In Silico Assessment in Cyanobacteria, Monocot, and Dicot Plants †

Ishita Biswas [1,2] and Debanjan Mitra [1,*]

[1] Department of Microbiology, Raiganj University, Raiganj 733134, West Bengal, India
[2] Department of Botany, S.B.S. Government College, Hili 733126, West Bengal, India
* Correspondence: debanjanmitra267@gmail.com; Tel.: +91-7479199988

† Presented at the 3rd International Electronic Conference on Agronomy, 15–30 October 2023; Available online: https://iecag2023.sciforum.net/.

Abstract: RuBisCO is the main photosynthetic enzyme of carbon assimilatory pathways in nature. Despite being the most abundant protein on earth, RuBisCO is still relatively underutilised in the food chain. Although there are sequence and structure details in the database, there are few instances of studies on evolutionary relationships. A bioinformatics and in silico study was conducted to check sequence and structural differences of RuBisCO among different photosynthetic organisms. RuBisCO from *Oryza sativa* showed an abundance of charged amino acids, salt-bridges, and intra-protein interactions and was more hydrophilic in nature compared to *Nostoc* sp., *Chlamydomonas reinhardtii*, and *Nicotiana tabacum*. From molecular dynamics simulations, lower root mean square deviation and root mean square fluctuation indicate that RuBisCO from *Oryza sativa* was more stable, followed by *Nicotiana tabacum*, and a lower radius of gyrations indicates their tight packing. From this study, it was clear that some specific evolutions in charged amino acids of RuBisCO of monocot, i.e., *Oryza sativa*, make it more stable and stronger than other plant groups. The study concludes that a more stable nature of RuBisCO is gained from monocot *Oryza sativa*.

Keywords: RuBisCO; evolution; salt-bridge; intra-protein interactions; molecular dynamic simulations

Citation: Biswas, I.; Mitra, D. Comparative Analysis of RuBisCO Evolution and Intrinsic Differences: Insights from In Silico Assessment in Cyanobacteria, Monocot, and Dicot Plants. *Biol. Life Sci. Forum* **2023**, *27*, 44. https://doi.org/10.3390/IECAG2023-15820

Academic Editor: Gianni Bellocchi

Published: 3 November 2023

1. Introduction

The most prevalent protein on earth is undoubtedly RubisCO (Ribulose-1,5-bisphosphate carboxylase oxygenase) [1]. Most autotrophic organisms, starting from prokaryotes and including cyanobacteria, photosynthetic bacteria, chemoautotrophic bacteria, and archaea to eukaryotes such as algae and higher plants, possess this enzyme. According to estimates, Rubisco can make up to 50% of all soluble proteins found in plant leaves or inside microbes [2]. Its presence in marine phytoplankton, which is projected to contribute more than 45% of yearly world net primary production, is perhaps less evident but nonetheless pervasive [3]. The enzyme RuBisCO (EC 4.1.1.39, Ribulose-1,5-bisphosphate carboxylase oxygenase) catalyses the primary photosynthetic CO_2 reduction process, which involves the binding of CO_2 to the acceptor molecule Ribulose-1,5-bisphosphate (RuBP) to produce 3-phosphoglycerate [4]. In addition, it also possesses oxygenase activity through which it binds O_2 with RuBP to form 2-phosphoglycolate. The discovery of the salt bridge's microenvironment is a novel concept in structural biology. An intragenic protein sequence and structural study sheds light on species variety, functions, and evolutionary relationships [5]. Although there are numerous sequences and structures of Rubisco in the database, however, there is no such report on the sequence and structure analysis of this protein in terms of the salt bridge and other interactions of proteins inside. This study was conducted to check the evolutionary pattern in RuBisCO protein, starting from lower group of

photosynthetic organism to higher group of plants through intra-protein interactions and molecular dynamics simulation studies.

2. Methods

2.1. Dataset

Sequences of RuBisCO enzyme from *Nostoc* sp., *Chlamydomonas reinhardtii*, *Oryza sativa*, and *Nicotiana tabacum* were retrieved from the UniProt database [6]. Diverse organisms were taken into consideration to check evolutionary significance between them. The structures of those proteins were extracted from the RCSB PDB database [7].

2.2. Analysis of Protein Sequences

All protein sequences were subjected to MSA for block preparation through Clustal omega [8]. The non-block format was used for the calculation of amino acid abundance, pI, grand average hydropathy (GRAVY), and aliphatic index through the ProtParam server [9]. The block format of sequences was used in calculations of hydropathy and polarity by the ProtScale server [9].

2.3. Analysis of Crystal Structures

All those protein structures were minimised through the Chimera 1.15rc with an amber forcefield [10,11]. Identification and calculations of intra-protein interactions were made through the PIC server [12]. PDBSum were used to check their type of secondary structures [13].

2.4. Molecular Dynamics Simulations

GROMACS [14] and the GROMOS96 43a1 forcefield were used for molecular dynamic simulations. After equilibration, the energy on the solvated systems was decreased using the steepest descent approach with 5000 steps. The last manufacturing run's molecular dynamic simulations lasted 50 ns at 300 K temperatures. Molecular dynamics simulations were used to estimate the radius of gyration (Rg), solvent accessible surface area (SASA), root mean square deviation (RMSD), root mean square fluctuation (RMSF), and hydrogen bonding.

3. Results and Discussion

3.1. Preferable Amino Acid Abundance

A higher abundance of charged polar residues was observed in *Oryza sativa*, followed by *Nicotiana tabacum* (Figure 1). However, the uncharged polar residues showed a higher presence in the enzyme of *Nostoc* sp. Hydrophobic amino acid showed the highest abundance in *Chlamydomonas reinhardtii*. *Oryza sativa* had a higher presence of amino acid Pro. Due to its capacity to reduce the structural entropy of the denatured state, Pro may be able to improve protein stability. Additionally, Pro is often conserved in proteins and frequently contributes significantly to the structure and function of proteins [15].

A common method for identifying hydrophobic areas in proteins is the Kyte–Doolittle scale. Positive values indicate hydrophobic regions. Transmembrane helices are predicted using hydropathy plots. The non-polar portion of the lipid membrane contains the transmembrane helices, whereas the loops are in a more polar solution [5]. *Nicotiana tabacum* showed the lowest plot, followed by *Oryza sativa*, which means they were more hydrophilic than others.

The specific polarity pattern is crucial to the molecule's structure and functionality. Higher polarity has been shown by protein sequencing of *Nicotiana tabacum* followed by *Oryza sativa*. However, it was observed that in some specific positions in *Chlamydomonas reinhardtii*, there are some pick-points for high polarity. Higher polarity drastically increased the thermal stability of a protein [16,17].

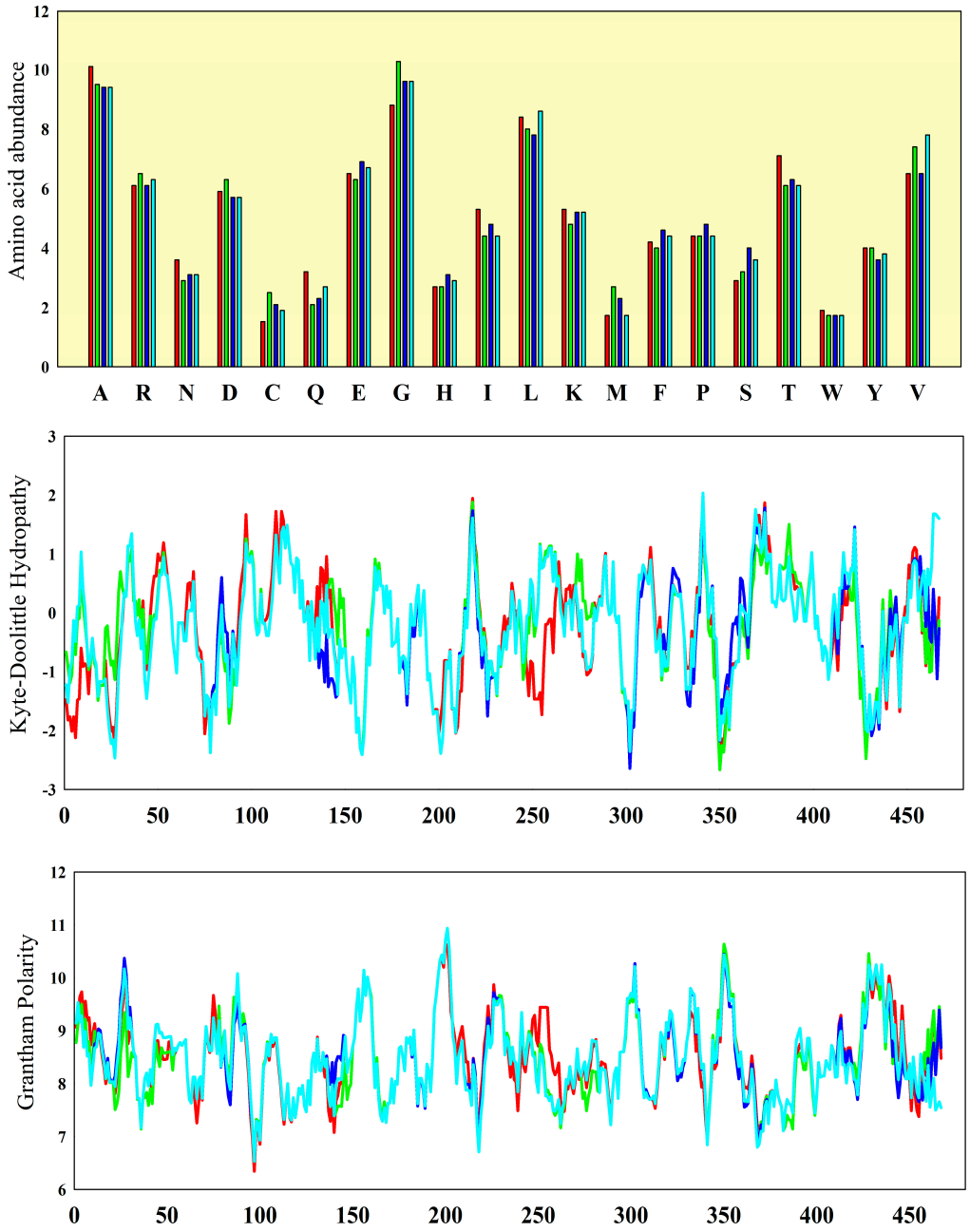

Figure 1. Amino acid abundance, Kyte-Dolittle hydropathy, Grantham polarity for RuBisCO enzyme from *Nostoc* sp. (red), *Chlamydomonas reinhardtii* (green), *Oryza sativa* (blue), and *Nicotiana tabacum* (cyan).

3.2. Secondary Structure Assessment

Nostoc sp. and *Chlamydomonas reinhardtii* both showed four sheets, four beta-alpha-beta units, two beta hairpins, three beta bulges, and sixteen strands. Difference has been shown in *Nostoc* sp., where there are 22 helices, 23 helix–helix interactions, 22 beta turns, 5 gamma turns and 1 disulphide structure. However, in those places, in *Chlamydomonas reinhardtii*, there are 21 helices, 21 helix–helix interactions, 31 beta turns, and 4 gamma turns. The RuBisCO of *Oryza sativa* possesses five sheets, four beta-alpha-beta units, two beta hairpins, four beta bulges, sixteen strands, twenty-three helices, twenty-eight helix–helix interactions, twenty-three beta turns, and three gamma turns, whereas *Nicotiana tabacum* had three sheets, five beta-alpha-beta units, two beta hairpins, four beta bulges, fifteen strands, twenty-two helices, twenty-five helix–helix interactions, twenty-three beta turns,

one gamma turn, and one disulphide bond. Increasing the amount of helix in the protein secondary structure ultimately increases the protein stability [18].

3.3. Intra-Protein Interactions

Intra-protein interactions are crucial interactions to enhance protein stability. Salt-bridges, aromatic–aromatic interactions, aromatic–sulphur interactions, and cation–pi interactions have significant contributions in this field. Generally, they act as a single pair called isolated; however, sometimes multiple isolated are connected to each other to make a network formation.

Strong hydrogen bonds that are formed by the interaction of two charged residues are known as salt bridges or ion pairs. In contrast to surface salt bridges, a subsurface salt bridge destabilises by 3–4 kcal/mol when one partner is removed. Nowadays, beside the isolated and network salt bridge, a special salt-bridge, i.e., a cyclic salt bridge, has been discovered [19]. The highest number of isolated salt bridges was found in RuBisCO of *Oryza sativa* (23), whereas *Nicotiana tabacum* showed the highest number of network salt bridges (10) (Table 1). *Nostoc* sp. showed 15 isolated and 7 network salt bridges, whereas *Chlamydomonas reinhardtii* had 19 isolated and 9 network salt bridges. The higher formation of salt bridges in monocot and dicot makes them more stable than the other two species. Important non-covalent interactions in proteins involve aromatic–aromatic interactions between several aromatic amino acids (Phe, Tyr, and Trp). *Nostoc* sp. showed the highest number of isolated aromatic–aromatic interactions, i.e., seven, with three network aromatic–aromatic interactions. *Chlamydomonas reinhardtii* showed five isolated and three network aromatic–aromatic interactions. RuBisCO of monocot *Oryza sativa* had five isolated and six network formation, which was highest formation of network aromatic–aromatic interactions. *Nicotiana tabacum* showed an equal number of isolated and network aromatic–aromatic interactions. Aromatic–sulphur interactions were higher in *Chlamydomonas reinhardtii* in the form of isolated bonds. However, the only formation of network aromatic–sulphur interactions was observed here. The formation of isolated cation–pi interactions was almost the same in every species, except *Chlamydomonas reinhardtii.* However, the network formation was higher in *Nicotiana tabacum*, whereas others showed an equal number, i.e., two network formations. The formation of a higher number of intra-protein interactions in RuBisCO of dicot and monocot gives an advantage to them to gain more stability over the algae and cyanobacteria.

Table 1. Intra-protein interactions in RuBisCO of *Nostoc* sp., *Chlamydomonas reinhardtii*, *Oryza sativa*, and *Nicotiana tabacum*.

Protein	Salt Bridge		Aromatic–Aromatic		Aromatic–Sulphur		Cation–pi	
	Isolated	Network	Isolated	Network	Isolated	Network	Isolated	Network
Nostoc sp.	15	7	7	3	3	0	8	2
Chlamydomonas reinhardtii	19	9	5	3	8	1	7	2
Oryza sativa	23	6	5	6	7	0	8	2
Nicotiana tabacum	11	10	5	5	5	0	8	4

3.4. Stability through Simulation Study

The 50 ns molecular dynamics simulations provide the details of RMSD, RMSF, Rg, and SASA. From the RMSD, it was observed that *Nostoc* sp. had a higher RMSD than the others. It started to deviate from 0.2 nm and ended at 0.6 nm (Figure 2). On other hand, *Oryza sativa* and *Nicotiana tabacum* showed lower and almost equal trajectories throughout the 50 ns RMSD analysis, from the start at 0.2 nm, becoming stabilised and ending at almost the same range. RMSF analysis revealed that all proteins showed an almost similar trajectory throughout the path. The Rg plot also showed similarity with the plot to RMSD. *Oryza sativa* and *Nicotiana tabacum* showed a lower Rg than the others, which means they

had the tightest packing RuBisCO [20]. SASA was also high in *Oryza sativa* and *Nicotiana tabacum*. Increasing the value of SASA enhanced the stability and protein folding [21].

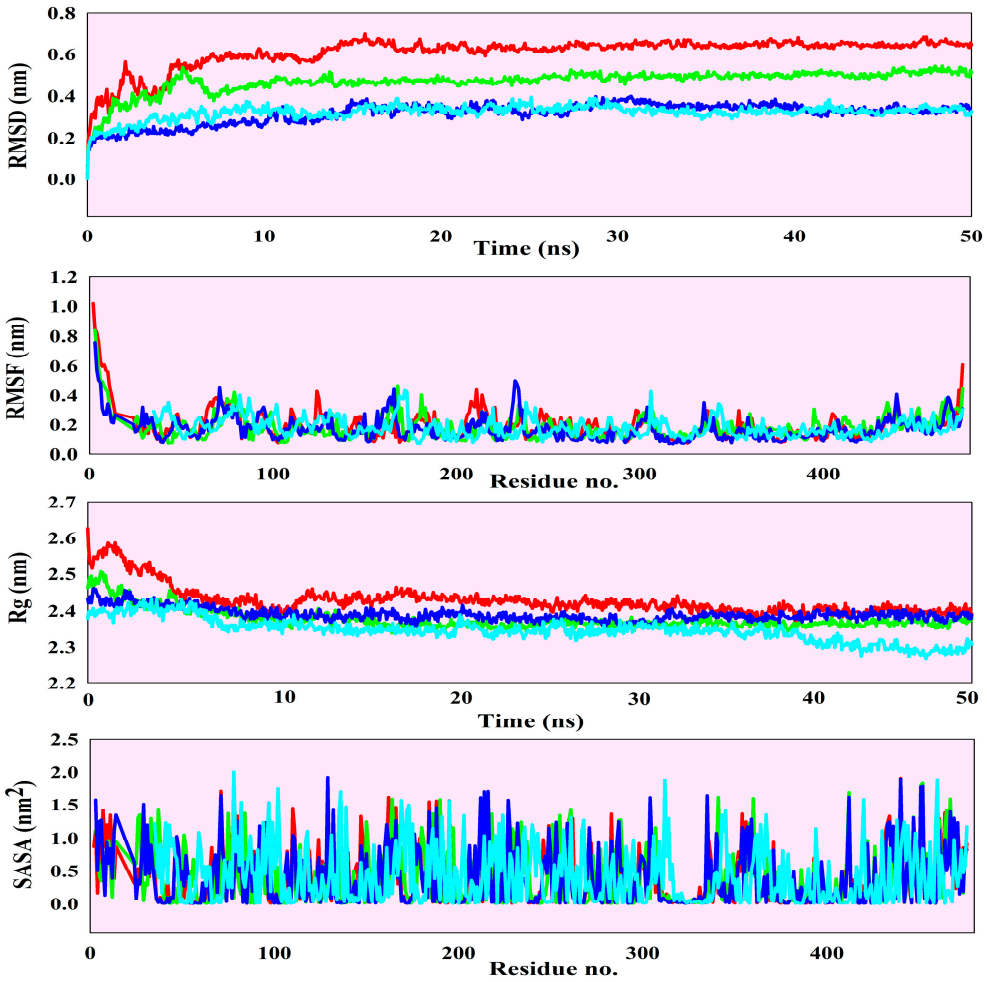

Figure 2. RMSD, RMSF, Rg, and SASA for RuBisCO enzyme from *Nostoc* sp. (red), *Chlamydomonas reinhardtii* (green), *Oryza sativa* (blue), and *Nicotiana tabacum* (cyan).

4. Conclusions

In silico investigation into different RuBisCO revealed how the amino acid evolutions make significant changes to gain more stability and flexibility in higher groups of plants. Charged amino acid residue abundance was found mainly in *Nicotiana tabacum* and *Oryza sativa* plants. Moreover, a higher hydrophilicity and higher polarity enhance the stability, functionality, and flexibility. The increase of helices in secondary structures further boost the stability. Molecular dynamics simulations revealed higher stability, flexibility, and folding patterns of RuBisCO from *Oryza sativa* and *Nicotiana tabacum*. This study will be helpful to understand the protein evolution and role played in protein engineering.

Author Contributions: Conceptualization, D.M. and I.B.; methodology, I.B.; software, D.M.; validation, D.M.; formal analysis, I.B.; investigation, D.M.; resources, D.M.; data curation, I.B.; writing—original draft preparation, I.B.; writing—review and editing, D.M.; visualization, I.B.; supervision, D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

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

References

1. Tabita, F.R.; Hanson, T.E.; Li, H.; Satagopan, S.; Singh, J.; Chan, S. Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 576–599. [CrossRef] [PubMed]
2. Ellis, R.J. The most abundant protein in the world. *Trends Biochem. Sci.* **1979**, *4*, 241–244. [CrossRef]
3. Field, C.B.; Behrenfeld, M.J.; Randerson, J.T.; Falkowski, P. Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **1998**, *281*, 237–240. [CrossRef] [PubMed]
4. Andersson, I.; Backlund, A. Structure and function of Rubisco. *Plant Physiol. Biochem.* **2008**, *46*, 275–291. [CrossRef] [PubMed]
5. Biswas, I.; Mitra, D.; Bandyopadhyay, A.K.; Mohapatra, P.K. Contributions of protein microenvironment in tannase industrial applicability: An in-silico comparative study of pathogenic and non-pathogenic bacterial tannase. *Heliyon* **2020**, *6*, e05359. [CrossRef] [PubMed]
6. UniProt Consortium. UniProt: A hub for protein information. *Nucleic Acids Res.* **2015**, *43*, D204–D212. [CrossRef] [PubMed]
7. Kouranov, A.; Xie, L.; de la Cruz, J.; Chen, L.; Westbrook, J.; Bourne, P.E.; Berman, H.M. The RCSB PDB information portal for structural genomics. *Nucleic Acids Res.* **2006**, *34* (Suppl. 1), D302–D305. [CrossRef] [PubMed]
8. Sievers, F.; Higgins, D.G. Clustal omega. *Curr. Protoc. Bioinform.* **2014**, *48*, 3–13. [CrossRef] [PubMed]
9. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.E.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. *Protein Identification and Analysis Tools on the ExPASy Server*; Humana Press: Totowa, NJ, USA, 2005.
10. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef] [PubMed]
11. Mitra, D.; Das Mohapatra, P.K. In silico comparative structural and compositional analysis of glycoproteins of RSV to study the nature of stability and transmissibility of RSV A. *Syst. Microbiol. Biomanuf.* **2023**, *3*, 312–327. [CrossRef] [PubMed]
12. Tina, K.G.; Bhadra, R.; Srinivasan, N. PIC: Protein interactions calculator. *Nucleic Acids Res.* **2007**, *35* (Suppl. 2), W473–W476. [CrossRef] [PubMed]
13. Laskowski, R.A. PDBsum: Summaries and analyses of PDB structures. *Nucleic Acids Res.* **2001**, *29*, 221–222. [CrossRef] [PubMed]
14. Mitra, D.; Paul, M.; Thatoi, H.; Mohapatra, P.K. Study of potentiality of dexamethasone and its derivatives against COVID-19. *J. Biomol. Struct. Dyn.* **2022**, *40*, 10239–10249. [CrossRef] [PubMed]
15. Bajaj, K.; Madhusudhan, M.S.; Adkar, B.V.; Chakrabarti, P.; Ramakrishnan, C.; Sali, A.; Varadarajan, R. Stereochemical criteria for prediction of the effects of proline mutations on protein stability. *PLoS Comput. Biol.* **2007**, *3*, e241. [CrossRef] [PubMed]
16. Mitra, D.; Das Mohapatra, P.K. Cold adaptation strategy of psychrophilic bacteria: An in-silico analysis of isocitrate dehydrogenase. *Syst. Microbiol. Biomanuf.* **2021**, *1*, 483–493.
17. Mitra, D.; Pal, A.K.; Das Mohapatra, P.K. Intra-protein interactions of SARS-CoV-2 and SARS: A bioinformatic analysis for plausible explanation regarding stability, divergency, and severity. *Syst. Microbiol. Biomanuf.* **2022**, *2*, 653–664. [CrossRef]
18. Mitra, D. Deep Dive Into the DNA Polymerase through Insilico Analysis: An Information to Get Better PCR Enzyme from the Ancient One. *Preprint* **2023**. [CrossRef]
19. Mitra, D.; Das Mohapatra, P.K. Discovery of novel cyclic salt bridge in thermophilic bacterial protease and study of its sequence and structure. *Appl. Biochem. Biotechnol.* **2021**, *193*, 1688–1700. [CrossRef] [PubMed]
20. Mitra, D.; Afreen, S.; Das Mohapatra, P.K.; Abdalla, M. Threat of respiratory syncytial virus infection knocking the door: A proposed potential drug candidate through molecular dynamics simulations, a future alternative. *J. Mol. Model.* **2023**, *29*, 91. [CrossRef] [PubMed]
21. Ausaf Ali, S.; Hassan, I.; Islam, A.; Ahmad, F. A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. *Curr. Protein Pept. Sci.* **2014**, *15*, 456–476.