



*applied sciences*

# Nucleic Acids Conjugates for Biotechnological Applications

---

Edited by

Tamaki Endoh, Eriks Rozners and Takashi Ohtsuki

Printed Edition of the Special Issue Published in *Applied Sciences*

# **Nucleic Acids Conjugates for Biotechnological Applications**



# **Nucleic Acids Conjugates for Biotechnological Applications**

Editors

**Tamaki Endoh**

**Eriks Rozners**

**Takashi Ohtsuki**

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



*Editors*

Tamaki Endoh

Frontier Institute for  
Biomolecular Engineering  
Research (FIBER)

Konan University

Kobe

Japan

Eriks Rozners

Department of Chemistry

Binghamton University

Binghamton

United States

Takashi Ohtsuki

Department of Interdisciplinary

Science and Engineering in

Health Systems

Okayama University

Okayama

Japan

*Editorial Office*

MDPI

St. Alban-Anlage 66

4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Applied Sciences* (ISSN 2076-3417) (available at: [www.mdpi.com/journal/applsci/special\\_issues/Heterogeneous\\_Conjugates\\_Nucleic\\_Acids](http://www.mdpi.com/journal/applsci/special_issues/Heterogeneous_Conjugates_Nucleic_Acids)).

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. <i>Journal Name</i> <b>Year</b> , <i>Volume Number</i> , Page Range.
--

**ISBN 978-3-0365-1514-4 (Hbk)**

**ISBN 978-3-0365-1513-7 (PDF)**

© 2021 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license, which allows users to download, copy and build upon published articles, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons license CC BY-NC-ND.

# Contents

<b>Preface to "Nucleic Acids Conjugates for Biotechnological Applications"</b> . . . . .	vii
<b>Tamaki Endoh, Eriks Rozners and Takashi Ohtsuki</b> Nucleic Acids Chemistry and Engineering: Special Issue on Nucleic Acid Conjugates for Biotechnological Applications Reprinted from: <i>Applied Sciences</i> <b>2021</b> , <i>11</i> , 3594, doi:10.3390/app11083594 . . . . .	1
<b>Alina Derzhalova, Oleg Markov, Alesya Fokina, Yasuo Shiohama, Timofei Zatsepin, Masayuki Fujii, Marina Zenkova and Dmitry Stetsenko</b> Novel Lipid-Oligonucleotide Conjugates Containing Long-Chain Sulfonyl Phosphoramidate Groups: Synthesis and Biological Properties Reprinted from: <i>Applied Sciences</i> <b>2021</b> , <i>11</i> , 1174, doi:10.3390/app11031174 . . . . .	5
<b>Yuichiro Aiba, Gerardo Urbina, Masanari Shibata and Osami Shoji</b> Investigation of the Characteristics of NLS-PNA: Influence of NLS Location on Invasion Efficiency Reprinted from: <i>Applied Sciences</i> <b>2020</b> , <i>10</i> , 8663, doi:10.3390/app10238663 . . . . .	21
<b>Masayasu Mie, Rena Hirashima, Yasumasa Mashimo and Eiry Kobatake</b> Construction of an Enzymatically-Conjugated DNA Aptamer-Protein Hybrid Molecule for Use as a BRET-Based Biosensor Reprinted from: <i>Applied Sciences</i> <b>2020</b> , <i>10</i> , 7646, doi:10.3390/app10217646 . . . . .	31
<b>Shun Nakano, Taiki Seko, Zhengxiao Zhang and Takashi Morii</b> RNA-Peptide Conjugation through an Efficient Covalent Bond Formation Reprinted from: <i>Applied Sciences</i> <b>2020</b> , <i>10</i> , 8920, doi:10.3390/app10248920 . . . . .	41
<b>Kenta Odaira, Ken Yamada, Shogo Ishiyama, Hidenori Okamura and Fumi Nagatsugi</b> Design of the Crosslinking Reactions for Nucleic Acids-Binding Protein and Evaluation of the Reactivity Reprinted from: <i>Applied Sciences</i> <b>2020</b> , <i>10</i> , 7709, doi:10.3390/app10217709 . . . . .	51
<b>Yoko Nakamura, Yuki Sato, Hiroshi Inaba, Takashi Iwasaki and Kazunori Matsuura</b> Encapsulation of mRNA into Artificial Viral Capsids via Hybridization of a -Annulus-dT <sub>20</sub> Conjugate and the Poly(A) Tail of mRNA Reprinted from: <i>Applied Sciences</i> <b>2020</b> , <i>10</i> , 8004, doi:10.3390/app10228004 . . . . .	65
<b>Hiroto Fujita, Yuka Kataoka and Masayasu Kuwahara</b> Bifunctional Aptamer Drug Carrier Enabling Selective and Efficient Incorporation of an Approved Anticancer Drug Irinotecan to Fibrin Gels Reprinted from: <i>Applied Sciences</i> <b>2020</b> , <i>10</i> , 8755, doi:10.3390/app10238755 . . . . .	75
<b>Misaki Sugimoto, Akiyo Suda, Shiroh Futaki and Miki Imanishi</b> Effective RNA Regulation by Combination of Multiple Programmable RNA-Binding Proteins Reprinted from: <i>Applied Sciences</i> <b>2020</b> , <i>10</i> , 6803, doi:10.3390/app10196803 . . . . .	85
<b>Junya Akagi, Takahiro Yamada, Kumi Hidaka, Yoshihiko Fujita, Hirohide Saito, Hiroshi Sugiyama, Masayuki Endo, Shigeyoshi Matsumura and Yoshiya Ikawa</b> An RNA Triangle with Six Ribozyme Units Can Promote a <i>Trans</i> -Splicing Reaction through Trimerization of Unit Ribozyme Dimers Reprinted from: <i>Applied Sciences</i> <b>2021</b> , <i>11</i> , 2583, doi:10.3390/app11062583 . . . . .	95

**Chun-xia Wang, Yusuke Sato, Takashi Sugimoto, Norio Teramae and Seiichi Nishizawa**  
Chloro-Substituted Naphthyridine Derivative and Its Conjugate with Thiazole Orange for  
Highly Selective Fluorescence Sensing of an Orphan Cytosine in the AP Site-Containing  
Duplexes  
Reprinted from: *Applied Sciences* **2020**, *10*, 4133, doi:10.3390/app10124133 . . . . . **111**

# Preface to "Nucleic Acids Conjugates for Biotechnological Applications"

Biological activities are maintained by the cooperative actions of biomolecules. The details of the functionalization mechanisms of these cooperative actions have been elucidated. With the development of biological and chemical technologies, artificial modifications of such systems consisting of various biomolecules has been attempted.

Nucleic acids not only store genetic information in their primary sequence but also exhibit biological functions through the formation of their unique structures. Based on such functions and specific and designable complementary recognition through base pairing, nucleic acids are considered attractive molecules as functional units for biotechnological applications.

This Special Issue focuses on the development of nucleic acid-based conjugates (including chemical conjugates with covalent bonds and complexed conjugate thorough intermolecular interaction) with different types of biomolecules such as proteins, lipids, and organic molecules for cellular engineering and biotechnological applications.

**Tamaki Endoh, Eriks Rozners, Takashi Ohtsuki**

*Editors*



Editorial

# Nucleic Acids Chemistry and Engineering: Special Issue on Nucleic Acid Conjugates for Biotechnological Applications

Tamaki Endoh <sup>1,\*</sup>, Eriks Rozners <sup>2</sup> and Takashi Ohtsuki <sup>3</sup><sup>1</sup> Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, Kobe 650-0047, Japan<sup>2</sup> Department of Chemistry, Binghamton University, New York, NY 13902, USA; erozners@binghamton.edu<sup>3</sup> Department of Interdisciplinary Science and Engineering in Health Systems, Okayama University, Okayama 700-8530, Japan; ohtsuk@okayama-u.ac.jp

\* Correspondence: t-endoh@konan-u.ac.jp; Tel.: +81-78-303-1314

## 1. Introduction

Nucleic acids not only store genetic information in their primary sequence but also exhibit biological functions through the formation of their unique structures. Based on such functions and specific and designable complementary recognition through base pairing, nucleic acids are considered attractive molecules as functional units for biotechnological applications. Life in biological systems is maintained by the cooperative actions of various biomolecules, such as nucleic acids, proteins, lipids, and small organic molecules. With the development of chemical and biological technologies related to nucleic acids, the details of the mechanisms of such cooperative actions between nucleic acids and other biomolecules have been elucidated. This Special Issue focuses on the biotechnological applications of nucleic acid conjugates that transcend the natural cooperative systems.

## 2. Addition of Superior Functions to Natural Nucleic Acid through Covalent Conjugations

Nucleic acids show complementary recognition by base pairing, which cannot be achieved by other biomolecules. A specific region on messenger RNA (mRNA) can be targeted by rationally designing oligonucleotides, such as antisense and small interfering RNA (siRNA), that enable the suppression of the protein expression from the target mRNA. An artificial nucleic acid mimic typified by peptide nucleic acid (PNA) further enables targeting a specific region on genomic double-stranded DNA through the formation of a double-duplex invasion complex. When these oligonucleotides are applied for gene regulation, the delivery of the oligonucleotides into cells and localization in appropriate cellular compartments are issues that need to be addressed. The conjugation of a functional unit for providing cell permeability to oligonucleotides is a promising strategy for the intracellular delivery. D. Stetsenko et al. synthesized a novel type of lipid-oligonucleotide conjugates (LONs) based on their original chemistry, performing the conjugation on solid support. They demonstrated that siRNAs built of LONs downregulated gene expression through self-internalization into cells [1]. Since the constructed LONs spontaneously formed particles, the additional functionality of resistance to degradation for in vivo application is expected. Localization in a specific cellular compartment can be achieved by the conjugation of a short peptide tag, such as nuclear localization signal (NLS). Y. Aiba and O. Shoji et al. have constructed an NLS-PNA conjugate, which was superior to natural PNA in its invasion efficiency into the target DNA duplex [2]. The demonstration motivates the applicability for in cell studies because the NLS-PNA could form the double-duplex invasion complex even in the presence of physiological salt concentrations, whereas the invasion efficiency of PNA decreased as the salt concentration increased.

In addition to the complementary sequence, nucleic acids form complex tertiary structures that recognize other biomolecules such as proteins and small molecule metabolites. Aptamers, nucleic acids with potential to specifically recognize another molecule, have



**Citation:** Endoh, T.; Rozners, E.; Ohtsuki, T. Nucleic Acids Chemistry and Engineering: Special Issue on Nucleic Acid Conjugates for Biotechnological Applications. *Appl. Sci.* **2021**, *11*, 3594. <https://doi.org/10.3390/app11083594>

Received: 13 April 2021

Accepted: 14 April 2021

Published: 16 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

been applied as functional units in biosensors. Biosensors require a unit for recognizing the target molecule and another unit for outputting signal, which can be easily recorded and quantified such as fluorescence and luminescence. Since aptamers only recognize the target molecule, the construction of aptamer-based biosensors has been demonstrated by conjugations with other functional units. For example, the conjugation of an aptamer with a protein is one of the useful strategies because some proteins exhibit unique photochemical output signals. M. Mie et al. have demonstrated the construction of nucleic acid–protein hybrid molecules using a protein involved in viral replication initiation that acts as a conjugation unit. They constructed an aptamer–protein hybrid that detects the target molecule by a bioluminescence resonance energy transfer (BRET) through designing a fusion protein of the conjugation unit between NanoLuc luciferase and Venus, which output luminescence and fluorescence as a donor and acceptor of BRET, respectively [3]. The construct of the aptamer–protein hybrid would be applicable for sensing various aptamer targets because change in the BRET signal occurs upon steric rearrangement of the protein units upon the interaction of the aptamer and its target. When the target of the aptamer is a small molecule, it may be difficult to cause a large steric rearrangement of the protein units. In such cases, aptamer-based biosensors can also be constructed by optimally locating a small fluorophore, in which fluorescence is sensitive to the microenvironment, close to the aptamer unit. T. Morii et al. have constructed fluorometric sensors based on a ribonucleopeptide (RNP), which acts as a scaffold to locate RNA aptamer and fluorophore units in proximity. Although the prototype RNP-based sensors were just formed by a noncovalent interaction between RNA and peptide, the present study covalently linked the complex to improve chemical and thermal stability through optimization of the conjugation conditions [4]. The covalently linked RNP sensor would be applicable for the simultaneous detection of multi targets by preventing the interexchange of the functional units between RNP sensors.

The formation of covalent linkage between nucleic acid and protein can also be used for entrapping proteins that interact with a nucleic acid of interest. Although crosslinking chemical reagents for biomolecules have been developed, in order to capture a protein that binds to a specific site of the target nucleic acid, it is necessary to conjugate the reactive unit at the specific site of nucleic acid. F. Nagatsugi et al. have constructed oligonucleotides bearing an unnatural nucleotide consisting of 4-amino-6-oxo-2-vinyltriazine as the reactive unit and acyclic linker in place of deoxyribose. The vinyl group in the reactive unit enables covalent linkage with amino acid residues with nucleophilic species at a proximity that captures the proteins interacting with the specific region on the duplex [5]. The formation of covalent linkage is useful for the basic study of nucleic acid binding proteins. In addition, it is also expected to be possible to use the reactive oligonucleotide as an irreversible inhibitor of the proteins like a suicide inhibitor of catalytic proteins.

### 3. Molecular Assemblies Including Nucleic Acids for Functional Materials and Systems

The original and fundamental function of nucleic acids is to store the genetic information required to synthesize proteins. Since mRNA encodes the amino acid sequence of the protein to be expressed, it is attracting attention as biomedical material exemplified as a vaccine against SARS-CoV-2. When mRNA is used as a nucleic acid-based drug, it should be delivered into cells to express the protein of interest. Due to the considerably large size of mRNA compared to oligonucleotide therapeutics, chemical conjugation through the technology introduced above becomes difficult. In this case, the encapsulation of the mRNA into an assembled complex, which shows cellular internalization function, would provide an efficient strategy to deliver the mRNAs. K. Matsuura et al. have constructed an artificial capsule by using a short peptide unit derived from tomato bushy stunt virus. They have chemically conjugated the peptide with a short poly-T oligonucleotide and succeeded to encapsulate mRNA, which hybridizes to the poly-T strand through its poly-A tail. The capsule was further modified to internalize into cells and express proteins by the modification of the outside of the capsule with a cell penetrating peptide [6]. The

modification functions like a viral envelope and is expected to be applicable for targeting specific cells. In contrast to the use of nucleic acid as a drug, a strategy for internalizing a small drug molecule into an assembled complex via nucleic acid is also promising for the therapeutic application. M. Kuwahara et al. have constructed a bifunctional aptamer (bApt) consisting of two aptamer units for thrombin and a drug molecule, respectively. During the fibrin gel formation by the catalytic activity of thrombin, the thrombin aptamer unit enables the internalization of bApt into fibrin gel. Since another aptamer unit, which is interacting with the drug molecule, can be easily exchanged depending on the therapeutic target, the assembled fibrin gel can be a biocompatible material that stores and releases various drug molecules [7].

Assembly between nucleic acids and proteins is one of the major mechanisms that modulate gene expression in cells. For example, the assembly of Puf family proteins on the 3' untranslated region of mRNA affects protein expression. Since the target RNA sequence of the Puf family proteins can be altered by rationally designing amino acids in their RNA binding domain, a strategy that assembles multiple Puf proteins on target mRNA would enable the efficient regulation of protein expression. M. Imanishi et al. have fused the engineered RNA binding domain of the Puf family protein with tristetraproline (TTP), which is involved in mRNA decay and degradation. They demonstrated that a tandem assembly of four TTP-PUF fusion proteins effectively repressed protein expression from target mRNA [8]. Since the strong suppression was observed only when the four types of TTP-PUF fusion proteins are present as it is like a logic gate, the technology is expected to be applied as a gene regulation tool in the field of synthetic biology.

Molecular assembly is also applied for in vitro applications. The hybridization and self-organization of nucleic acids are designable based on the sequence complementarity. The reconstruction of assembly-based nucleic acid structures is applicable for controlling the subsequent output of functional units. As examples, Y. Ikawa et al. have demonstrated the regulation of catalytic activity of RNA through self-organization of split RNAs. They constructed an RNA triangle with six catalytic ribozyme units, in which the triangle complex was further supported by protein assembly on the corners. It is an attractive RNA-based nanoarchitecture and would be a unit for a system that exhibits more complex biological functions as the RNA triangle has succeeded to catalyze the trans-splicing of RNA with an arbitrary sequence [9]. S. Nishizawa et al. have synthesized a chemical probe for fluorometric detection of an orphan cytosine, which has no base pairing partner due to an abasic nucleotide on the opposite strand. They also designed an oligonucleotide, which hybridizes to the target strand and locates abasic nucleotide opposite a specific cytosine, and demonstrated the detection of target nucleic acid by fluorescence signal through induction of the orphan cytosine and subsequent assembly of the chemical probe [10]. Since the assembly of the chemical probe is specific for the orphan cytosine sandwiched by matched base pairs, the technique is expected to be applicable for the detection of single nucleotide polymorphisms including viral mutations in relatively wide sequence variations.

#### 4. Summary

In the published papers in this Special Issue, advanced research works involved in nucleic acid conjugates are reported in wide application fields, such as artificial gene regulation, biomolecular sensing, and therapeutics. In general, biotechnologies have been developed through discoveries and utilizations of functional mechanisms in life systems. The field of nucleic acid conjugates for technological applications will continue to develop since the full extent of natural nucleic acid functions has not been clarified.

**Funding:** This research received no external funding.

**Acknowledgments:** This Special Issue would not have been possible without the contributions of authors who responded to the invitation. We sincerely appreciate all authors for their efforts to summarize the research works and the editorial team in Applied Sciences for their support.

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

## References

1. Derzhalova, A.; Markov, O.; Fokina, A.; Shiohama, Y.; Zatsepin, T.; Fujii, M.; Zenkova, M.; Stetsenko, D. Novel Lipid-Oligonucleotide Conjugates Containing Long-Chain Sulfonyl Phosphoramidate Groups: Synthesis and Biological Properties. *Appl. Sci.* **2021**, *11*, 1174. [[CrossRef](#)]
2. Aiba, Y.; Urbina, G.; Shibata, M.; Shoji, O. Investigation of the Characteristics of NLS-PNA: Influence of NLS Location on Invasion Efficiency. *Appl. Sci.* **2020**, *10*, 8663. [[CrossRef](#)]
3. Mie, M.; Hirashima, R.; Mashimo, Y.; Kobatake, E. Construction of an Enzymatically-Conjugated DNA Aptamer-Protein Hybrid Molecule for Use as a BRET-Based Biosensor. *Appl. Sci.* **2020**, *10*, 7646. [[CrossRef](#)]
4. Nakano, S.; Seko, T.; Zhang, Z.; Morii, T. RNA-Peptide Conjugation through an Efficient Covalent Bond Formation. *Appl. Sci.* **2020**, *10*, 8920. [[CrossRef](#)]
5. Odaira, K.; Yamada, K.; Ishiyama, S.; Okamura, H.; Nagatsugi, F. Design of the Crosslinking Reactions for Nucleic Acids-Binding Protein and Evaluation of the Reactivity. *Appl. Sci.* **2020**, *10*, 7709. [[CrossRef](#)]
6. Nakamura, Y.; Sato, Y.; Inaba, H.; Iwasaki, T.; Matsuura, K. Encapsulation of mRNA into Artificial Viral Capsids via Hybridization of a  $\beta$ -Annulus-dT20 Conjugate and the Poly(A) Tail of mRNA. *Appl. Sci.* **2020**, *10*, 8004. [[CrossRef](#)]
7. Fujita, H.; Kataoka, Y.; Kuwahara, M. Bifunctional Aptamer Drug Carrier Enabling Selective and Efficient Incorporation of an Approved Anticancer Drug Irinotecan to Fibrin Gels. *Appl. Sci.* **2020**, *10*, 8755. [[CrossRef](#)]
8. Sugimoto, M.; Suda, A.; Futaki, S.; Imanishi, M. Effective RNA Regulation by Combination of Multiple Programmable RNA-Binding Proteins. *Appl. Sci.* **2020**, *10*, 6803. [[CrossRef](#)]
9. Akagi, J.; Yamada, T.; Hidaka, K.; Fujita, Y.; Saito, H.; Sugiyama, H.; Endo, M.; Matsumura, S.; Ikawa, Y. An RNA Triangle with Six Ribozyme Units Can Promote a Trans-Splicing Reaction through Trimerization of Unit Ribozyme Dimers. *Appl. Sci.* **2021**, *11*, 2583. [[CrossRef](#)]
10. Wang, C.-X.; Sato, Y.; Sugimoto, T.; Teramae, N.; Nishizawa, S. Chloro-Substituted Naphthyridine Derivative and Its Conjugate with Thiazole Orange for Highly Selective Fluorescence Sensing of an Orphan Cytosine in the AP Site-Containing Duplexes. *Appl. Sci.* **2020**, *10*, 4133. [[CrossRef](#)]

## Article

# Novel Lipid-Oligonucleotide Conjugates Containing Long-Chain Sulfonyl Phosphoramidate Groups: Synthesis and Biological Properties

Alina Derzhalova <sup>1,2,†</sup>, Oleg Markov <sup>2,†</sup>, Alesya Fokina <sup>1,3,†</sup>, Yasuo Shiohama <sup>4</sup>, Timofei Zatsepin <sup>5,6</sup> , Masayuki Fujii <sup>7</sup> , Marina Zenkova <sup>2</sup>  and Dmitry Stetsenko <sup>1,3,\*</sup> 

<sup>1</sup> Department of Physics, Novosibirsk State University, 630090 Novosibirsk, Russia; alina.derzhalova@gmail.com (A.D.); alesya\_fokina@mail.ru (A.F.)

<sup>2</sup> Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia; markov\_oleg@list.ru (O.M.); marzen@niboch.nsc.ru (M.Z.)

<sup>3</sup> Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

<sup>4</sup> Tropical Biosphere Research Center, University of the Ryukyus, Okinawa 903-0213, Japan; h194005@comb.u-ryukyu.ac.jp

<sup>5</sup> Skolkovo Institute of Science and Technology, Innovation Centre Skolkovo, 121205 Moscow, Russia; tsz@yandex.ru

<sup>6</sup> Department of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia

<sup>7</sup> Department of Biological & Environmental Chemistry, Kindai University, Iizuka, Fukuoka 820-8555, Japan; mfujii@fuk.kindai.ac.jp

\* Correspondence: d.stetsenko@nsu.ru; Tel.: +7-383-363-4963

† These authors contributed equally to the manuscript.



**Citation:** Derzhalova, A.; Markov, O.; Fokina, A.; Shiohama, Y.; Zatsepin, T.; Fujii, M.; Zenkova, M.; Stetsenko, D. Novel Lipid-Oligonucleotide Conjugates Containing Long-Chain Sulfonyl Phosphoramidate Groups: Synthesis and Biological Properties. *Appl. Sci.* **2021**, *11*, 1174. <https://doi.org/10.3390/app11031174>

Academic Editor: Tamaki Endoh, Eriks Rozners and Takashi Ohtsuki

Received: 8 December 2020

Accepted: 24 January 2021

Published: 27 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** New lipid conjugates of DNA and RNA incorporating one to four [(4-dodecylphenyl)sulfonyl] phosphoramidate or (hexadecylsulfonyl)phosphoramidate groups at internucleotidic positions near the 3' or 5'-end were synthesized and characterized. Low cytotoxicity of the conjugates and their ability to be taken up into cells without transfection agents were demonstrated. Lipid-conjugated siRNAs targeting repulsive guidance molecules a (RGMa) have shown a comparable gene silencing activity in PK-59 cells to unmodified control siRNA when delivered into the cells via Lipofectamine mediated transfection.

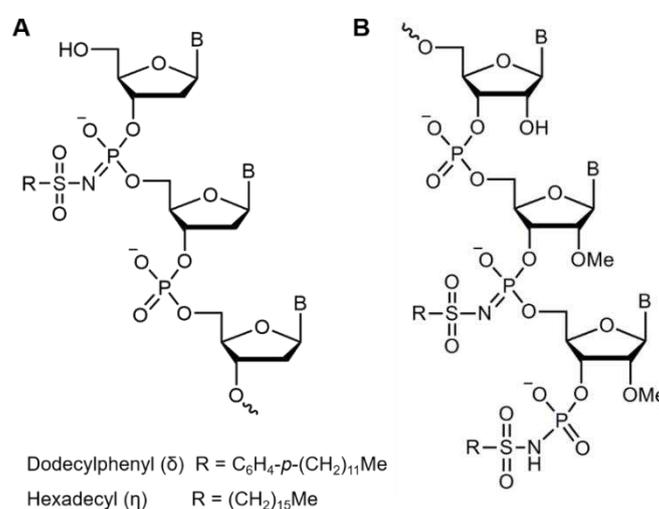
**Keywords:** therapeutic nucleic acid; drug delivery; nanoparticles; cytotoxicity; macrophages; cellular uptake; small interfering RNA; multiple sclerosis; repulsive guidance molecule a

## 1. Introduction

Nucleic acid derivatives such as antisense oligonucleotides or small interfering RNAs (siRNAs) are extensively studied as therapeutic agents that target biologically important RNAs; either mRNAs or noncoding RNAs such as micro-RNAs via either enzyme-mediated RNA cleavage [1–3] or steric blocking of RNA function [4,5]. While the former may be applicable to a wide range of diseases including some genetic disorders [6], cancer [7], viral [8] and bacterial [9] infections, where the target RNA can be eliminated, the latter is often employed where a fine tuning of the RNA activity is required such as in splice switching [10]. The progress in the development of oligonucleotide-based therapies in the past few years has resulted in the recent approval of several antisense and siRNA drugs by the FDA and EMA [11]. However, widespread application of therapeutic oligonucleotides still suffers from their relatively inefficient cellular uptake and often poor bioavailability [12,13]. The ways of promoting cellular delivery of oligonucleotides and their analogues include their structural modification [14,15], noncovalent binding [16] or covalent conjugation [17,18] to the molecules that could improve translocation through the outer membrane and direct intracellular trafficking of nucleic acids. Among these, lipids in a wider sense

including biomolecules such as cholesterol, tocopherol (vitamin E) or fatty acids, and their synthetic analogues such as, e.g., DOTMA [19] have been shown to improve cellular uptake of oligonucleotides, most relevantly, in the form of lipid-oligonucleotide conjugates (LONs) [20–22].

In particular, lipid conjugates of antisense oligonucleotides and siRNAs have improved transfection ability and extended half-life in the bloodstream with beneficial impact on therapeutic activity *in vivo* [23–25]. We describe herein a series of novel DNA and RNA lipid conjugates, which have been chemically modified by one to four long-chain *N*-(sulfonyl)phosphoramidate groups, namely, [(4-dodecylphenyl)sulfonyl]-phosphoramidate group ( $\delta$ ) or (hexadecylsulfonyl)phosphoramidate group ( $\eta$ ) at the internucleotidic position(s) near the 3'- or 5'-end of the oligonucleotide chain (Figure 1). The oligonucleotides were obtained by an automated phosphoramidite solid-phase synthesis protocol that employs the Staudinger reaction between the support-bound phosphite triester and the respective sulfonyl azide instead of usual aqueous iodine oxidation to introduce either of the lipophilic modifications following a previously developed procedure [26–28].



**Figure 1.** Structures and positions of hydrophobic groups in lipid-oligonucleotide conjugates used in this study: oligodeoxynucleotide (**A**); small interfering RNA (siRNA) (**B**).

The obtained lipid conjugates of oligodeoxyribonucleotides with one or two of the respective hydrophobic groups were nearly non-cytotoxic for either murine or human macrophages up to 20  $\mu\text{M}$  concentration. Both types of lipid conjugates were able to enter the cells in a carrier free manner. Both features are beneficial for potential therapeutic application. The corresponding lipid conjugates of siRNAs with either two or four hydrophobic groups were designed to downregulate a validated target in multiple sclerosis, namely, repulsive guidance molecule a (RGMa) [29]. The conjugates with two groups were shown to possess silencing ability comparable to the unmodified siRNA when delivered by lipofectamine. This type of lipid conjugate may represent a useful addition to the inventory of therapeutic oligonucleotide derivatives.

## 2. Materials and Methods

### 2.1. Oligonucleotide Synthesis

Oligodeoxynucleotides and siRNAs containing ((4-dodecylphenyl)sulfonyl)phosphoramidate ( $\delta$ ) or (hexadecylsulfonyl)phosphoramidate ( $\eta$ ) groups were assembled by automated solid-phase synthesis on a Mermaid MM-12 DNA/RNA synthesizer according to a modified  $\beta$ -cyanoethyl phosphoramidite protocol replacing aqueous iodine oxidation by a Staudinger reaction with either 4-dodecylbenzenesulfonyl azide or 1-hexadecanesulfonyl azide (0.5 M in acetonitrile for ( $\delta$ ) or in acetonitrile—THF (2:1 *v/v*) for ( $\eta$ ), 120 min at

ambient temperature) in an appropriate cycle of chain elongation [27,28]. Standard *N*-protected 5'-DMTr-3'-β-cyanoethyl-*N,N'*-phosphoramidites of deoxy, 2'-*O*-TBDMS-ribo and 2'-*O*-methylribonucleotides (Sigma-Aldrich Inc., St Louis, MO, USA) and 1000 Å CPG polymer supports (Glen Research Corp, Sterling, VA, USA) were used. After the completion of the synthesis, oligonucleotides were cleaved from support and deprotected by standard procedures followed by consecutive PAGE and RP-HPLC purification as described previously [30]. Molecular masses of the oligonucleotides were verified by ESI LC-MS (Table 1).

**Table 1.** Structures and molecular masses of oligonucleotides and their lipid conjugates (LONs).

Type	Designation	Oligonucleotide Sequence, 5'-3' <sup>1</sup>	Molecular Mass, Da <sup>2</sup>	
			Calc. [M]	Exp. [M]
Model antisense oligonucleotide	ODN	agtctcgacttgctacc <sup>3</sup>	5121.25	5120.98
	ODN- <i>F</i> <sup>4</sup>	agtctcgacttgctacc-Flu	5688.74	5687.93
	1δ	a <sup>δ</sup> gtctcgacttgctacc	5248.86	5426.08
	2δ	a <sup>δ</sup> g <sup>δ</sup> tctcgacttgctacc	5736.36	5733.28
	1δ- <i>F</i>	a <sup>δ</sup> gtctcgacttgctacc-Flu	5996.35	5993.21
	2δ- <i>F</i>	a <sup>δ</sup> g <sup>δ</sup> tctcgacttgctacc-Flu	6303.85	6300.41
	1η	a <sup>η</sup> gtctcgacttgctacc	5408.87	5406.11
	2η	a <sup>η</sup> g <sup>η</sup> tctcgacttgctacc	5696.38	5693.34
	1η- <i>F</i>	a <sup>η</sup> gtctcgacttgctacc-Flu	5976.36	5973.24
	2η- <i>F</i>	a <sup>η</sup> g <sup>η</sup> tctcgacttgctacc-Flu	6263.87	6260.47
siRNA	EGFP	<i>Sense</i> CCUACGGCAAGCUGACCCUGA <i>Antisense</i> AGGGUCAGCUUGCCGUAGGUG	6698.10 6812.10	6695.84 6811.57
	RGMa	<i>Sense</i> CCAUCAUCUUCAAGAACUCC <i>Antisense</i> AAGUUCUUGAAGAUGAUGGUG	6563.99 6806.20	6565.38 6811.30
	2Δ-RGMa	<i>Sense</i> CCAUCAUCUUCAAGAACUUC <sup>δ</sup> C <sup>δ</sup> -p <sup>6</sup> <i>Antisense</i> AAGUUCUUGAAGAUGAUGGUG	7219.40 6806.20	7220.40 6811.30
	4Δ-RGMa	<i>Sense</i> CCAUCAUCUUCAAGAACUUC <sup>δ</sup> C <sup>δ</sup> -p <i>Antisense</i> AAGUUCUUGAAGAUGAUGG <sup>δ</sup> C <sup>δ</sup> -p	7219.40 7460.60	7220.40 7466.90
	2H-RGMa	<i>Sense</i> CCAUCAUCUUCAAGAACUUC <sup>η</sup> C <sup>η</sup> -p <i>Antisense</i> AAGUUCUUGAAGAUGAUGGUG	7260.00 6806.20	7262.90 6811.30
	4H-RGMa	<i>Sense</i> CCAUCAUCUUCAAGAACUUC <sup>η</sup> C <sup>η</sup> -p <i>Antisense</i> AAGUUCUUGAAGAUGAUGG <sup>η</sup> C <sup>η</sup> -p	7260.00 7501.20	7262.90 7502.00

<sup>1</sup> Lowercase letters indicate deoxyribonucleotides, uppercase letters—ribonucleotides, underlined uppercase—2'-*O*-methylribonucleotides.

<sup>2</sup> According to ESI LC-MS in the negative ions mode. <sup>3</sup> Control oligonucleotide with random sequence specially constructed not to affect gene expression in cells. <sup>4</sup> Flu—3'-terminal fluorescein label. <sup>5</sup> The symbols (δ), (η) mark the positions of dodecylphenyl or hexadecyl groups, respectively. <sup>6</sup> p—3'-terminal phosphate group.

4-Dodecylbenzenesulfonyl azide was obtained from a commercial supplier (Sigma-Aldrich Inc., St Louis, MO, USA). 1-Hexadecanesulfonyl azide was prepared from the corresponding sulfonyl chloride and a slight excess of sodium azide in acetonitrile—THF (1:1 *v/v*) as a white solid as described in [31]. IR, <sup>1</sup>H and <sup>13</sup>C NMR spectra of the compound were consistent with the structure.

## 2.2. Dynamic Light Scattering (DLS) Characterization

Lipid-oligonucleotide conjugates (LONs) were analyzed at room temperature by dynamic light scattering (DLS) using a Zetasizer Nano ZS90 (Malvern Instruments Ltd., Malvern, UK). Size was measured in a DTS0012 cuvette (Malvern Instruments Ltd., Malvern, UK). All the scattered photons were collected at a 173° scattering angle. Measurement conditions were: material protein (RI: 1.450; Absorption: 0.001), dispersant water (Viscosity: 0.8872 cP; RI: 1.330) temperature at 25 °C and equilibration time was 3 s. Each test was triplicated. Nanoparticle size was determined in deionized water at 5 μM oligonu-

cleotide concentration. The scattering intensity data was processed using the instrumental software to obtain the hydrodynamic diameter (Dh).

### 2.3. Cell Lines

RAW 264.7 cells were cultivated in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich Inc., St Louis, MO, USA) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (GE Healthcare, Chicago, IL, USA) and 1% antibiotic/antimycotic mix (100 unit/mL penicillin, 0.1 mg/mL streptomycin and 0.25 µg/mL amphotericin) at 37 °C in a humidified atmosphere with 5% CO<sub>2</sub> (standard conditions, SC) and passaged regularly to maintain exponential growth.

THP-1 cells were cultivated in Roswell Park Memorial Institute (RPMI) medium (Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% FBS, 1% antibiotic/antimycotic mix (100 unit/mL penicillin, 0.1 mg/mL streptomycin and 0.25 µg/mL amphotericin), 1% GlutaMAX (Thermo Fisher Scientific, Waltham, MA, USA), 1% HEPES (Thermo Fisher Scientific, Waltham, MA, USA) and 1% glucose (Sigma-Aldrich Inc., St Louis, MO, USA) under SC and passaged regularly to maintain exponential growth.

To generate human macrophages (Mφ) THP-1 cells were cultivated in complete RPMI medium in the presence of 200 nM PMA (Sigma-Aldrich Inc., St Louis, MO, USA) under SC for 48 h. Then, the culture medium was replaced with complete RPMI medium in the absence of PMA and cells were incubated for an additional 24 h under SC.

PK-59 human pancreatic cancer cells were cultivated in complete RPMI-1640 medium (Thermo Fisher Scientific, Waltham, MA, USA) under SC and passaged regularly to maintain exponential growth.

### 2.4. MTT Assay of Cell Viability

RAW264.7 cells were seeded in 96-well plates at a density of  $5 \times 10^3$  cells/well in complete DMEM and incubated for 24 h under SC to adhere. THP-1 cells were seeded in 96-well plates at a density of  $50 \times 10^3$  cells/well in complete DMEM medium supplemented with 200 nM PMA and incubated for 72 h under SC to generate human Mφ as described above. Then, RAW264.7 and Mφ cells were incubated in the corresponding complete culture media supplemented with 5, 10 and 20 µM of the respective lipid-oligonucleotide conjugate (LON) for an additional 24 h under SC. Cytotoxicity of LONs was assessed by MTT assay. Briefly, MTT solution (Sigma-Aldrich, Darmstadt, Germany) was added to the cells up to 0.5 mg/mL and the incubation was continued under SC for an additional 2 h. Then, the MTT-containing medium was aspirated and the formazan crystals formed in living cells were solubilized with 100 µL/well of DMSO. The absorbance of each well was read at the test and reference wavelengths (λ) of 570 and 620 nm, respectively, on a Multiscan RC plate reader (Thermo LabSystems, Helsinki, Finland). The percentage of living cells was calculated as follows:

$$\text{Living cells (\%)} = (\text{OD}_{570\text{exp}} - \text{OD}_{620\text{exp}}) / (\text{OD}_{570\text{cont}} - \text{OD}_{620\text{cont}}) \times 100, \quad (1)$$

where OD<sub>570exp</sub> and OD<sub>570cont</sub> correspond to optical density in experimental and control wells, respectively, at λ 570 nm, and OD<sub>620exp</sub> and OD<sub>620cont</sub> correspond to optical density in experimental and control wells, respectively, at λ 620 nm. All experimental points were run in triplicate for statistical analysis.

### 2.5. Flow Cytometry

Intracellular accumulation of the lipid-oligonucleotide conjugates labeled with fluorophores, namely δ-Flu, 2δ-Flu, η-Flu and 2η-Flu, was assessed in RAW264.7 and human Mφ by flow cytometry. RAW264.7 cells were seeded into 24-well plates at a density of  $80 \times 10^3$  cells/well in complete DMEM and cultivated overnight to adhere. THP-1 cells were seeded into 24-well plates at a density of  $100 \times 10^3$  cells/well in complete RPMI medium supplemented with 200 nM PMA and incubated for 72 h under SC to generate human Mφ as described above. Subsequently, RAW 264.7 and Mφ were incubated

in the corresponding complete media supplemented with 5  $\mu\text{M}$  of the respective lipid-oligonucleotide conjugate under SC for 24 h. After incubation, cells were detached from the plate with 2% trypsin (MP Biomedicals, Irvine, CA, USA), resuspended in a complete culture medium, centrifuged at  $200\times g$  for 5 min, washed with PBS and fixed with 2% formaldehyde in a PBS buffer (10 min, room temperature). Cells were analyzed on a NovoCyte flow cytometer (ACEA Biosciences, Santa Clara, CA, USA) and the data were processed with a NovoExpress (ACEA Biosciences, Santa Clara, CA, USA). All experiments were run in triplicate for statistical analysis. The transfection efficiency (TE) was characterized by two values: percentage of fluorescence-positive cells in a population and mean fluorescent intensity (MFI) of cells in a sample.

### 2.6. Confocal Microscopy

RAW264.7 cells were seeded on coverslips (Marienfeld, Lauda-Königshofen, Germany) in 24-well plates at a density of  $20 \times 10^3$  cells/well in complete DMEM medium and incubated for 24 h under SC to adhere. THP-1 cells were seeded on coverslips in 24-well plates at a density of  $100 \times 10^3$  cells/well in complete RPMI medium supplemented with 200 nM PMA to obtain human M $\phi$  as described above. Afterwards, RAW 264.7 cells and human M $\phi$  were incubated in the corresponding complete media in the presence of 5  $\mu\text{M}$  of the respective lipid-oligonucleotide conjugates under SC for 24 h. After incubation, coverslips with cells were washed twice with PBS, fixed with 3.7% formaldehyde in PBS buffer (15 min, 37 °C), and washed twice with PBS. The cells were stained with the Phalloidin-iFluor532 (Abcam, Cambridge, UK) or Phalloidin-TRITC (ECM Biosciences, Versailles, KY, USA) according to the manufacturer's protocols. After staining, the cells were washed twice with PBS and placed on slides in a drop of ProLong<sup>TM</sup> Glass Antifade Mountant with NucBlue<sup>TM</sup> (Thermo Fisher Scientific, Waltham, MA, USA). Mounted samples were allowed to cure on a flat, dry surface for 18–24 h at room temperature in the dark. Intracellular localization of lipid-oligonucleotide conjugates was assessed by confocal fluorescent microscopy on an LSM710 (Zeiss, Oberkochen, Germany) using an  $\alpha$ Plan-Apochromat 100 $\times$ /1.46 Oil DIC M27 objective. Analysis of intracellular accumulation of LONs and Z-stacks was performed using ZEN software (Zeiss, Oberkochen, Germany). Confocal microscopic analysis was performed in three channels (blue, green, red). Fluorescence in the blue channel corresponded to DAPI (nuclei staining); the green channel corresponded to fluorescence of LONs labeled with fluorescein (Flu), and the red channel corresponded to Phalloidin-iFluor532 or Phalloidin-TRITC (cytoskeleton staining).

### 2.7. Gene Silencing

Experiments on siRNA-mediated silencing of repulsive guidance molecule a (RGMa) were carried out in the RGMa-expressing PK-59 human pancreatic cancer cell line. On the day of transfection, PK-59 cells were separately seeded in 24-well flat-bottomed plates at a density of  $4 \times 10^5$  cells/well in 0.5 mL of growth medium without antibiotics. The cells were transfected with lipid-conjugated RGMa siRNAs carrying either two (passenger strand only) or four (both passenger and antisense strands) dodecylphenyl ( $\Delta$ ) or hexadecyl (H) groups at their 3'-end, or unmodified control RGMa siRNA (Table 1) at concentrations 1, 10, or 100 nM in a complex with Lipofectamine<sup>®</sup> 2000 (2  $\mu\text{L}$ /well) according to the manufacturer's instructions (Invitrogen Life Technologies Corporation, Frederick, MD, USA). The cells were then incubated at 37 °C in the atmosphere of 5% CO<sub>2</sub> in a humidified incubator. Cells were harvested and assayed 24 h after transfection.

The level of RGMa expression was quantified by real-time, two-step reverse transcriptase PCR (qRT-PCR) assay. Briefly, total RNA was extracted from the collected cells using ISOGENE (Nippon Gene, Ltd., Toyama, Japan), treated with DNase I<sup>®</sup> (ThermoFisher Scientific, Waltham, MA, USA) and retro-transcribed using Oligo (dT)<sub>20</sub> primer and ReverTra Ace<sup>®</sup> (Toyobo Co., Ltd., Japan). The resulting cDNA was quantified using the AriaMx Real-Time PCR System<sup>®</sup> (Agilent Technologies, Ltd., Santa Clara, CA, USA) using the following primers: 5'-TGGCCGCTCATCGACAATAA-3', forward; and 5'-

GCTTGTCCCCACCGTTCTTA–3', reverse. Quantifications of RGMA mRNA expressions were normalized by the Comparative Ct Method (Ct) using 18S rRNA as an endogenous control to ensure accuracy; primers: 5'–GTAACCCGTTGAACCCATT–3', forward; and 5'–CCATCCAATCGGTAGTAGCG–3', reverse [32]. All the data are the means of triplicate determinations. Each determination was calculated from triplicate measurements. Each error bar indicates a standard deviation of the triplicate determinations. Non-silencing siRNA targeting EGFP was used as a negative control, which was designed to have a minimum of 4 mismatches to all human, mouse and rat genes, and confirmed to have minimal targeting by genome-wide microarray analysis [33]. Significant differences in the results were determined by the two-tailed Student's t-test using GraphPad Prism. A probability value of  $p < 0.05$  was considered to be significant in the present study.

### 3. Results

#### 3.1. Synthesis of Lipid-Oligonucleotide Conjugates (LONs) Incorporating Long-Chain (Sulfonyl)Phosphoramidate Groups

The synthesis of lipid conjugates of oligodeoxynucleotides having one or two ((4-dodecylphenyl)sulfonyl)phosphoramidate ( $\delta$ ) or (hexadecylsulfonyl)phosphoramidate ( $\eta$ ) modifications of the phosphate linkages adjacent to the 5'-end (Figure 1A) was accomplished according to our previously reported protocols of automated solid-phase assembly via the phosphoramidite method employing the Staudinger reaction of the support-bound phosphite triester formed during the phosphoramidite coupling with 0.5 M solution of either 4-dodecylbenzenesulfonyl azide or 1-hexadecanesulfonyl azide in acetonitrile for 120 min at ambient temperature instead of the usual aqueous iodine oxidation [23,24]. Oligonucleotides containing 5'-dimethoxytrityl group (DMTr) were isolated by reverse-phased (RP) HPLC or, if necessary, purified to homogeneity by denaturing PAGE under the same conditions as normal phosphodiester oligodeoxynucleotides, followed by RP-HPLC to remove phosphodiester impurities. The molecular masses of the  $\delta$ - and  $\eta$ -oligodeoxynucleotides obtained have been verified by ESI LC-MS (Table 1).

Sequences of siRNAs targeting RGMA mRNA were obtained from literature [25]. Lipid-siRNA conjugates containing 2–4 dodecylphenyl ( $\delta$ ) or hexadecyl ( $\eta$ ) groups in the antisense and/or passenger chains were prepared according to the method above (Table 1). The 3'-overhangs of the siRNAs were modified by 2'-O-methyl diribonucleotides to increase enzymatic stability and avoid possible side reactions associated with the neighboring 2'-OH group when sulfonyl phosphoramidate groups were introduced into the overhang and at the 3'-terminal phosphate group (Figure 1B).

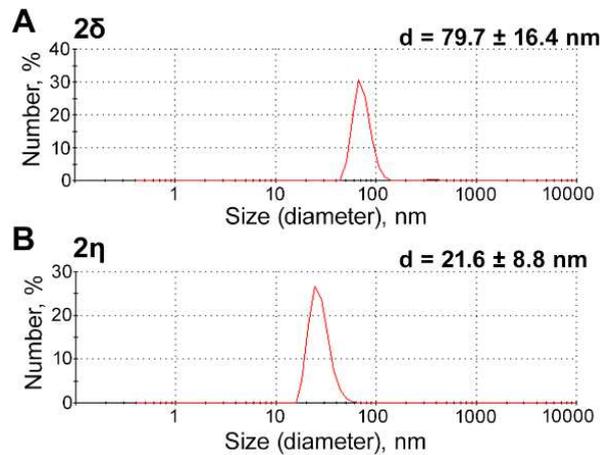
#### 3.2. Spontaneous Nanoparticle Formation by LONs in Aqueous Media

As amphiphiles, lipid-oligonucleotide conjugates (LONs) can be expected to spontaneously form micelles in aqueous solutions. Indeed, dynamic light scattering (DLS) analysis revealed that both 2 $\delta$  and 2 $\eta$  LONs formed nanoparticles in water characterized by mean diameters of  $21.6 \pm 8.8$  and  $79.7 \pm 16.4$  nm for 2 $\eta$  and 2 $\delta$ , respectively (Figure 2). Interestingly, longer aliphatic chain (16-carbon vs. 12-carbon) at the internucleotidic position favored the formation of smaller nanoparticles, which may be responsible for a better cellular uptake of a more hydrophobic 2 $\eta$  LON (Figure 2).

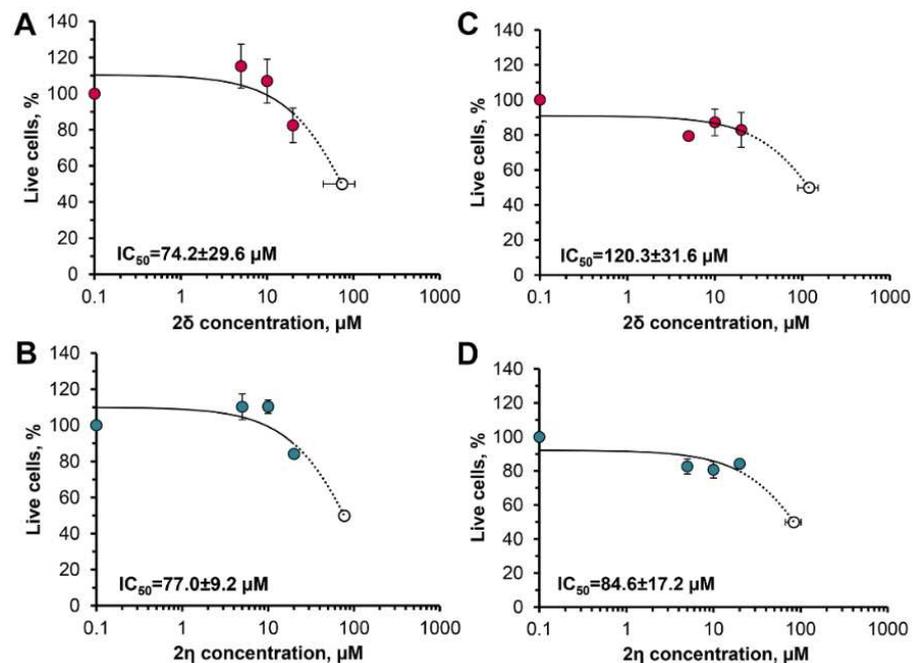
#### 3.3. LONs were Nontoxic to Murine or Human Macrophages at Therapeutic Concentrations

Introduction of lipophilic moieties into oligonucleotide chains can alter such properties of LONs as cytotoxicity. Cytotoxicity of the most hydrophobic LONs, namely, 2 $\delta$  and 2 $\eta$ , was analyzed with respect to murine RAW264.7 and human THP-1-derived macrophages. Cells were incubated in complete media supplemented with different concentrations of one of LONs for 24 h followed by estimation of cell viability using colorimetric MTT assay. As depicted in Figure 3, both 2 $\delta$  and 2 $\eta$  oligonucleotides were nontoxic to murine and human macrophages. IC<sub>50</sub> obtained by extrapolation of dose-response curves ranged from 74 to 120  $\mu$ M for 2 $\delta$  and from 77 to 85  $\mu$ M for 2 $\eta$  oligonucleotides, whereas the working

concentration of LONs in subsequent experiments was chosen to be 5  $\mu\text{M}$ . It is worth mentioning that murine macrophages were more sensitive to LONs in comparison with human THP-1-derived macrophages ( $\text{IC}_{50}$  were  $74.2 \pm 29.6 \mu\text{M}$  and  $120.3 \pm 31.6 \mu\text{M}$  for  $2\delta$  oligonucleotides, respectively;  $77.0 \pm 9.2 \mu\text{M}$  and  $84.6 \pm 17.2 \mu\text{M}$  for  $2\eta$  oligonucleotides, respectively) (Figure 3).



**Figure 2.** Hydrodynamic diameters of nanoparticles formed spontaneously in water by  $2\delta$  (A) and  $2\eta$  (B) lipid-oligonucleotide conjugates. Dynamic light scattering (DLS) data.



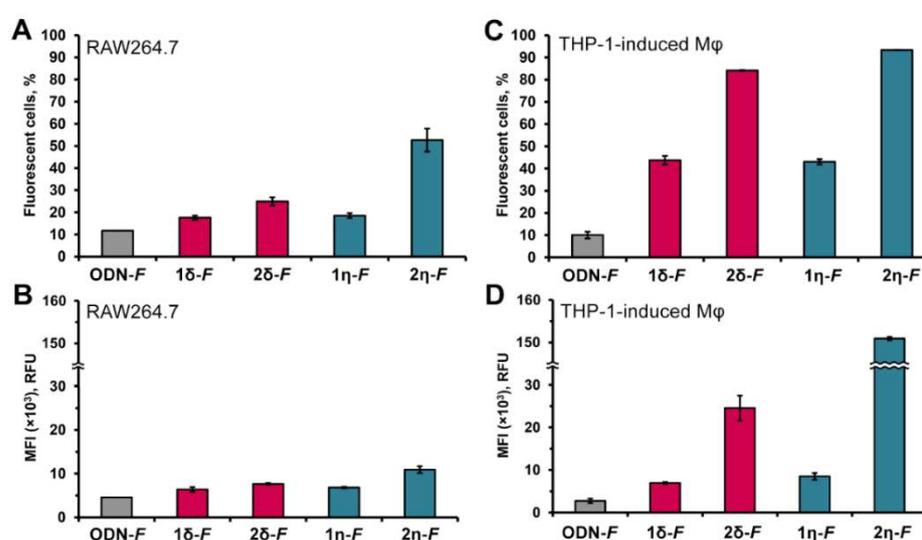
**Figure 3.** Dose-response curves of  $2\delta$  and  $2\eta$  LONs for RAW264.7 (A,B) and human THP-1 induced macrophages (C,D). Cells were incubated with LONs in a complete medium for 24 h. The results are shown as a percentage of viable cells observed after cell incubation with LON relative to untreated cells (100%). All experimental points were run in triplicate for statistical analysis. Data are presented as MEAN  $\pm$  SD. The dotted curve shows extrapolation to 50% of viable cells.

### 3.4. Conjugates with Longer Lipidic Chains Accumulated in Cells with Higher Efficiency

The ability of LONs to accumulate in eukaryotic cells in the absence of a transfection agent (gymnosis) was studied by flow cytometry using fluorescein-labeled LONs and murine macrophages RAW264.7 or THP-1-induced human macrophages. The efficiency

of intracellular accumulation of LONs was assessed by two parameters: the number of fluorescein-positive cells in the population and average level of the conjugate content in the cells (mean fluorescence intensity, MFI).

Accumulation of LONs in RAW264.7 and THP-1-induced macrophages differed considerably. LONs accumulated in RAW264.7 cells with moderate efficiency: 20–50% of cells were transfected with a MFI of  $6\text{--}11 \times 10^3$  RFU (Figure 4A,B). As expected, the cells incubated and control oligonucleotide ODN-F displayed the lowest efficiency: 10% of fluorescent cells with MFI of  $4 \times 10^3$  RFU, which corresponded to the baseline level of oligonucleotide taken up by the macrophages and/or adhered to the cell surface. We observed that more hydrophobic LONs, that is, the ones containing either two hydrocarbon chains instead of only one, or those having a longer aliphatic group: hexadecyl instead of dodecylphenyl, were showing more efficient accumulation in RAW264.7 cells (Figure 4A,B: compare conjugates having two vs. one group, and  $\eta$ - vs.  $\delta$ -modified LONs) than their less hydrophobic counterparts and unmodified control ODN-F.



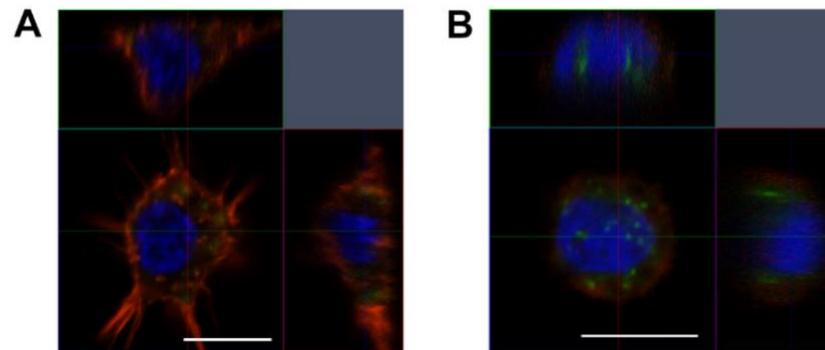
**Figure 4.** Accumulation of fluorescein-labeled LONs in murine macrophages RAW264.7 (A,B) and THP-1-induced human macrophages (C,D). The percentage of fluorescent cells (A,C) and mean fluorescence intensity (B,D) were measured by flow cytometry 24 h post transfection. Data are presented as MEAN  $\pm$  SD.

A similar but much more prominent accumulation of lipid-conjugated oligodeoxynucleotides was observed in human THP-1-induced macrophages. LONs accumulated in these cells with higher efficiency both in terms of the number of transfected cells and mean fluorescence intensity (Figure 4C,D). Conjugates 1 $\delta$ -F and 1 $\eta$ -F containing one lipophilic group displayed similar transfection efficiency regardless of the type of lipid chains, either dodecylphenyl ( $\delta$ ) or hexadecyl ( $\eta$ ): 43% cells were transfected with MFI of  $8 \times 10^3$  RFU, which was fourfold higher as compared to the control ODN-F (Figure 4C,D). Increasing the number of lipophilic groups from one to two resulted in a significant enhancement of transfection efficiency. 2 $\delta$  LON yielded 85% cells with fluorescence intensity  $25 \times 10^3$  RFU whereas 2 $\eta$  conjugate was taken up with the highest efficiency, accumulating in nearly 100% cells with fluorescence intensity  $150 \times 10^3$  RFU (Figure 4C,D), which is on par with the level achieved by liposome-mediated oligonucleotide delivery [34].

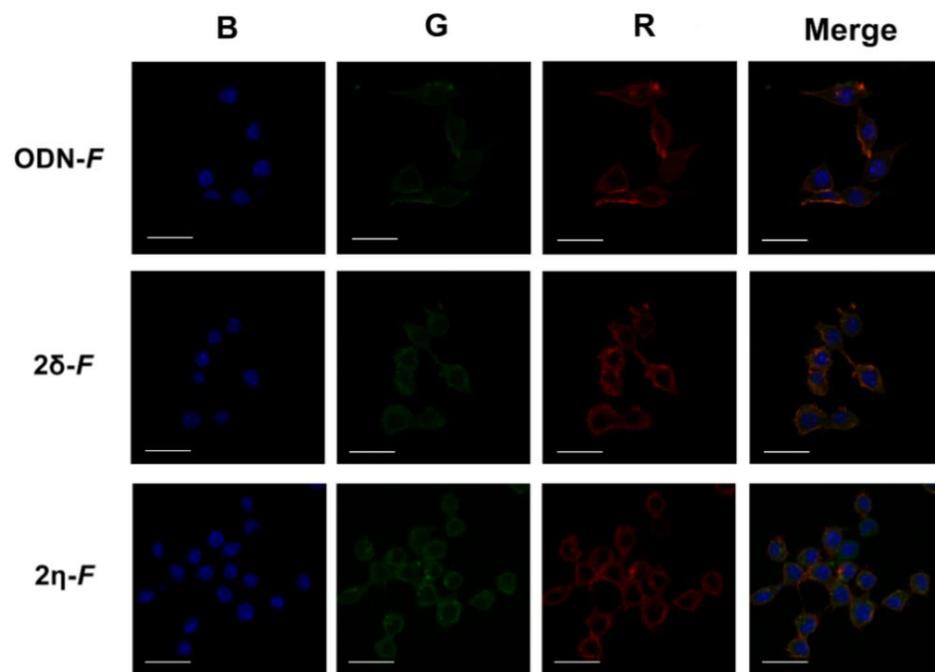
### 3.5. Lipid-Oligonucleotide Conjugates are Efficiently Taken up by Cells in the Absence of Transfection Agents and Localize Mostly in the Cytoplasm within Endosomes

Confocal microscopic analysis was carried out to verify and characterize intracellular localization of LONs. RAW 264.7 and human THP-1-induced macrophages were incubated with  $5.0 \mu\text{M}$  fluorescein-labeled LONs in complete medium for 24 h. After incubation, cells

were stained with Phalloidin-iFluor532 or Phalloidin-TRITC (actin microfilaments) and NucBlue (nuclei) (Figures 5–7).

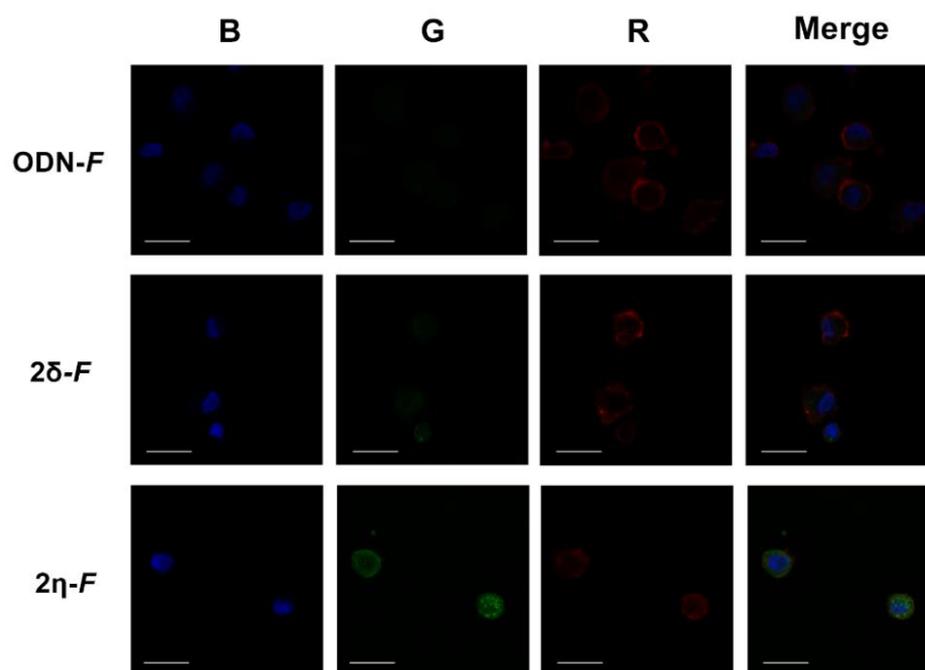


**Figure 5.** Intracellular localization of 2 $\eta$  LON in murine macrophages RAW264.7 (A) and THP-1-induced human macrophages (B). Analysis was performed by confocal fluorescent microscopy ( $\alpha$ Plan-Apochromat 100 $\times$ /1.46 Oil DIC M27 objective) 24 h post transfection. Data are presented as orthogonal projections. Blue signal is DAPI (nuclear staining), green signal is fluorescein-labeled 2 $\eta$  LON, red signal is Phalloidin-iFluor532 (A) or Phalloidin-TRITC (B). Scale bar 10  $\mu$ m.



**Figure 6.** Intracellular accumulation of LONs in RAW264.7. Analysis of samples was performed 24 h after addition of fluorescein-labeled conjugates (5  $\mu$ M) to the cells. Analysis was performed using  $\alpha$ Plan-Apochromat 100 $\times$ /1.46 Oil DIC M27 objective. Three-channel (BGR) pictures were obtained using staining by DAPI (nuclei staining) (B); fluorescein-labeled LONs (G); Phalloidin-iFluor532 (cytoskeleton staining) (R). Scale bars 20  $\mu$ m. ODN-F—control unmodified oligonucleotide labeled with fluorescein.

Firstly, to prove that LONs do not expose on the cell membranes but do penetrate into the cells, confocal microscopic analysis was performed and Z-stack images were built. It was demonstrated that LON 2 $\eta$ , the most efficient one according to flow cytometry data, was indeed localized in the cytosol of RAW264.7 cells and THP-1-induced macrophages (Figure 5). In addition, LON 1 $\eta$  was shown to efficiently accumulate in KB-8-5 cells (see Supplementary Materials, Figure S4).



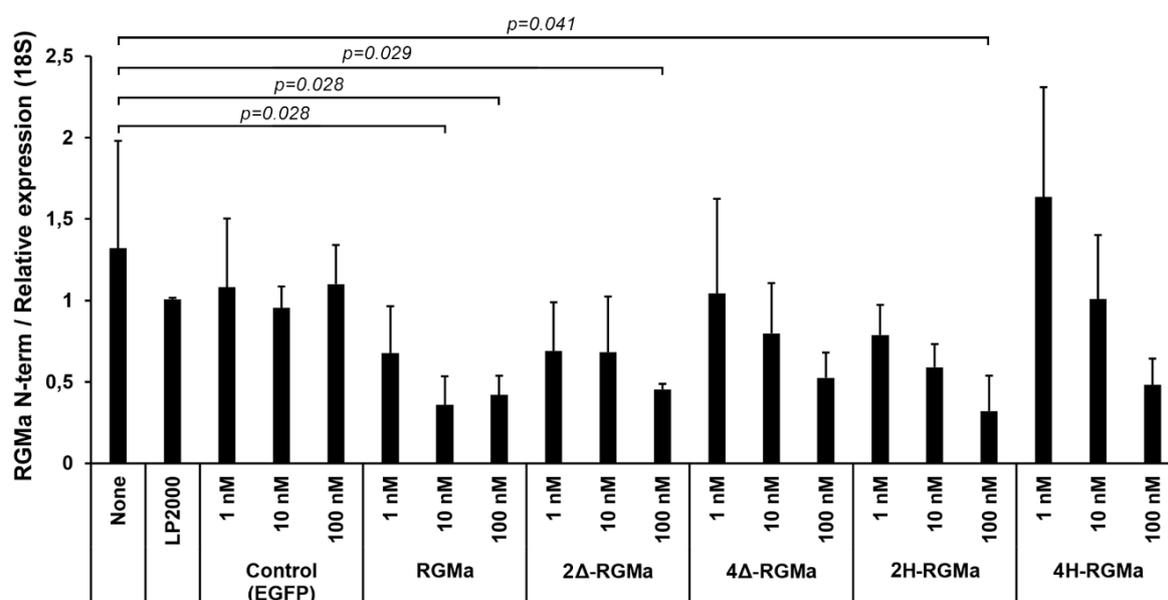
**Figure 7.** Intracellular accumulation of LONs in human THP-1-induced macrophages. Analysis of samples was performed 24 h after addition of fluorescein-labeled conjugates ( $5 \mu\text{M}$ ) to the cells. Analysis was performed using  $\alpha\text{Plan-Apochromat } 100\times/1.46 \text{ Oil DIC M27}$  objective. Three-channel (BGR) pictures were obtained using staining by DAPI (nuclear staining) (B); fluorescein-labeled LONs (G); Phalloidin-TRITC (cytoskeleton staining) (R). Scale bars  $20 \mu\text{m}$ . ODN-F—control unmodified oligonucleotide labeled with fluorescein.

It can be seen that the efficiency of intracellular accumulation of LONs according to confocal microscopy (Figures 6 and 7) correlate well with the obtained flow cytometry data (Figure 4). Oligonucleotides containing two lipophilic groups accumulated in cells with much higher efficiency in comparison with LONs containing only one group, LON  $2\eta$  with two hexadecyl groups being the best penetrator. LONs  $1\delta$ -F and  $1\eta$ -F containing only one lipophilic modification accumulated in cells with efficiency not significantly higher than that of the control unmodified oligonucleotide ODN-F. However, these LONs penetrated into the cells whereas ODN-F mainly adhered to the cell membrane (Figure S5, Supplementary Materials).

In detail, LONs were evenly distributed in the cytosol, localizing mainly in endosomes; this was conjectured from the presence of numerous bright green fluorescent dots containing LONs in the cytosol of both cell lines tested (Figures 6 and 7).

### 3.6. Lipid-Conjugated RGMA siRNAs Demonstrated Gene-Silencing Ability Similar to Unmodified RGMA siRNA in PK-59 Human Pancreatic Carcinoma Cell Culture

The gene silencing ability of lipid-conjugated siRNAs in comparison with unconjugated siRNA was elucidated in human pancreatic carcinoma PK-59 cells expressing repulsive guidance molecule a (RGMA). RGMA is a glycoprotein, which is a potent inhibitor of axonal re-myelination and, thus, is regarded as a potential therapeutic target for multiple sclerosis (MS) treatment [35]. Inhibition of RGMA by human monoclonal antibodies promotes neuronal regeneration and repair [36]. RGMA siRNA sequence was obtained from literature [29]. Lipid-conjugated RGMA siRNAs were modified at the 3'-end by either two (passenger strand only) or four (both passenger and antisense strands) dodecylphenyl ( $\Delta$ ) or hexadecyl (H) groups. To compare silencing activity under similar conditions, transfection was carried out at concentrations of siRNA 1, 10 and 100 nM in the presence of Lipofectamine<sup>®</sup> 2000. The levels of RGMA expression as assessed by RT-PCR are shown in Figure 8.



**Figure 8.** Inhibition of RGMa expression in PK-59 human pancreatic carcinoma cell culture by lipid conjugates of RGMa siRNAs vs. unmodified siRNA (RGMa). The conjugates were modified near the 3'-end by two (passenger chain only) or four (both chains) lipophilic groups ( $\Delta$ —dodecylphenyl, H—hexadecyl).

The gene silencing effect of siRNA conjugates 2 $\Delta$ -RGMa and 2H-RGMa that carried two lipophilic groups at the 3'-end of the passenger strand was comparable to that of the unconjugated RGMa siRNA at a concentration of 100 nM. At the same time, the activity of the conjugates 4 $\Delta$ -RGMa and 4H-RGMa containing four lipophilic groups at both 3'-ends of passenger and antisense strands was reduced as compared to the unmodified control siRNA.

#### 4. Discussion

A previously developed methodology of Staudinger reaction-mediated in-line oligonucleotide modification [26–28] was employed herein to obtain lipophilic conjugates of oligodeoxynucleotides and siRNAs carrying one or more extended hydrocarbon chains, 4-dodecylphenyl or hexadecyl, attached to the internucleotidic phosphate groups via the (sulfonyl)phosphoramidate linkage, which is hydrolytically stable and carries negative charge under physiological conditions (Figure 1A,B). The scheme allows one to modify any internucleotidic or terminal phosphate group within a DNA or 2'-OMe RNA chain from the 5'- to 3'-end by a facile chemical reaction with a stable sulfonyl azide precursor during automated DNA/RNA assembly without recourse to a special phosphoramidite monomer, which is usually obtained by a multi-step synthesis [18,20,21]. The obtained lipid-oligonucleotide conjugates (LONs) can be conveniently isolated by the usual C18 reverse-phased HPLC, due to significant difference in retention times with unmodified oligonucleotides (see Supplementary Materials) for HPLC profiles and ESI-MS spectra). Our preliminary data have shown that such lipophilic groups are fully compatible with other phosphate group modifications such as phosphorothioate, mesyl phosphoramidate [27,37,38] and phosphoryl guanidine [39].

Notably, we have observed spontaneous nanoparticle formation in aqueous solution by LONs containing two adjacent hydrophobic chains, with nanoparticle size dependent on the type of the side-chain (Figure 2). Nanoparticle formation by oligonucleotides is thought to be essential for good cellular uptake [40], structured RNA studies [41] and design of stimuli-responsive drug delivery systems [42].

Low cytotoxicity is another beneficial characteristic of such LONs, which allows one to vary the concentration of an oligonucleotide over a wide range to optimize its ther-

apeutic dosage for further *in vitro* and *in vivo* studies. We have confirmed that LONs used in this study evinced low cytotoxicity. Calculated  $IC_{50}$  of studied LONs fell within the 75 to 120  $\mu\text{M}$  range, which was significantly higher than concentrations of antisense oligonucleotides targeting therapeutically relevant RNAs commonly used *in vitro* and *in vivo*: nanomolar concentrations [43–46] or  $<5 \mu\text{M}$  [47] or therapeutic doses of siRNAs and antisense oligonucleotides approved by FDA and EMA for clinical practice: for siRNAs 21 nmol/kg for Patisiran [48] and 153 nmol/kg for Givosiran [49]; for antisense oligonucleotides  $<5 \mu\text{moles/kg}$  for Golodirsen [50], Volanesorsen [51] and Inotersen [52]. It is worth noting that the increase in hydrophobicity of the conjugates have no influence on their cytotoxicity (Figure 3). This result prompts us to expect that the increase of the number of side-chains or introduction of more hydrophobic groups, for example, hydrocarbon chains longer than 16 carbon atoms into LONs will not increase significantly the overall cytotoxicity of conjugates.

We demonstrated that LONs accumulated in the cells with intracellular distribution analogous to that of oligonucleotides delivered by liposomal formulations (Figures 4 and 5) [53–55]. It appears that LONs penetrated through the outer membrane and localized in cytosol predominantly in endosomes 24 h after addition (Figures 6 and 7). It is known that endosomal escape of therapeutic oligonucleotides is one of the rate-limiting steps of their delivery and the key to successful intracellular performance of nucleic acid therapeutics [56]. Although longer term kinetics of LON accumulation ( $>24$  h) was not studied, it can be assumed that such LONs will require longer times for their endosomal escape, similarly to cholesterol-conjugated siRNAs [57].

Cytoplasmic distribution of the developed LONs may be advantageous as far as the targets of therapeutic oligonucleotides (antisense oligonucleotides, siRNAs and anti-miRs) are localized in the cytosol such as mRNAs and many miRNAs. Application of the developed LONs to modulate splicing of pre-mRNA requires additional investigation including optimization of the number and length of the side-chains, selection of optimal delivery conditions etc. to achieve nuclear delivery of the conjugates. We demonstrated that dodecylphenyl and hexadecyl modifications did not significantly impede gene-silencing activity of lipophilic siRNAs delivered into cells by Lipofectamine (Figure 8). However, further studies on elucidating a wider spectrum of biological activities of siRNAs and antisense oligonucleotides carrying such (sulfonyl)phosphoramidate-based lipophilic groups are required.

One potential application for lipid-oligonucleotide conjugates as described herein could be the design of therapeutic oligonucleotides for the treatment of multiple sclerosis (MS), a severe neurodegenerative disease characterized by the destruction of the myelin sheaths of neurons [58]. Kitayama et al. [29] uncovered a link between the destruction of axonal myelin sheaths during MS and repulsive guidance molecule a (RGMa): a glycoprotein, which is a powerful inhibitor of neuronal regeneration. Inhibition of RGMa by human monoclonal antibodies stimulate re-myelination and regeneration of damaged neurons in the laboratory *in vivo* model of MS [36]. Thus, reducing the level of RGMa by downregulation of the corresponding gene by antisense oligonucleotide or siRNA could be a promising strategy for therapeutic intervention in MS.

We studied RGMa inhibition in the RGMa-expressing PK-59 pancreatic cancer cell line using lipid-siRNA conjugates containing either two or four dodecylphenyl or hexadecyl chains at the 3'-end of one or both of the complementary RNA strands, respectively (Figure 1B; Table 1). It was evident that siRNA conjugates containing two adjacent lipid groups in the passenger strand were able to downregulate RGMa mRNA with a similar efficiency to that of the unmodified control RGMa siRNA at 100 nM concentration while those carrying four of either dodecylphenyl or hexadecyl groups were significantly inferior (Figure 8). We foresee that such therapeutic lipid-oligonucleotide conjugates (LONs) displaying low toxicity and improved cellular uptake may demonstrate high efficiency and minimal side effects in future experiments *in vivo*.

## 5. Conclusions

To conclude, we have obtained lipid conjugates of oligodeoxynucleotides and siRNAs carrying one, two or four hydrophobic ((4-dodecylphenyl)sulfonyl)phosphoramidate or (hexadecylsulfonyl)phosphoramidate groups at either 5'- or 3'-internucleotidic positions by a modified protocol of the phosphoramidite synthesis replacing usual aqueous iodine oxidation by the Staudinger reaction between a solid-supported phosphite triester and the corresponding long-chain sulfonyl azide.

The oligodeoxynucleotide conjugates were characterized by low cytotoxicity in either murine or human macrophages and improved cell uptake into the same in the absence of a transfection agent. Spontaneous formation of nanoparticles was observed in aqueous solutions of doubly-conjugated oligonucleotides having either of the two lipophilic modifications at the 5'-end, which may be responsible for the improved cellular uptake of the respective conjugates. Notably, the size of the nanoparticles decreased with the increase in the length of the hydrocarbon chain from dodecylphenyl to hexadecyl. Flow cytometry revealed the superior uptake of hexadecyl conjugates over the dodecyl ones into human macrophages. Intracellular localization of the conjugates according to confocal microscopy was predominantly cytosolic with the majority of the fluorescence within endosomes.

Finally, two adjacent lipophilic dodecylphenyl or hexadecyl groups placed within the 3'-terminal overhang and at the 3'-terminal phosphate of the passenger strand did not significantly impair the gene-silencing ability of siRNAs targeting repulsive guidance molecule (RGMA) mRNA in PK-59 cell culture, whereas the siRNAs carrying four of the same at the 3'-ends of both passenger and antisense strands were considerably less active. Thus, lipid-siRNA conjugates with optimized number and position of the lipid chains could be a prospective candidate for future experiments in an in vivo model of multiple sclerosis and, potentially, other diseases when RNA targeting is applicable.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/2076-3417/11/3/1174/s1>, Figures S1–S3: IR, <sup>1</sup>H and <sup>13</sup>C spectra of 1-hexadecanesulfonyl azide; reverse-phased HPLC profiles and ESI mass spectra of modified oligonucleotides, Figure S4: Intracellular accumulation of η-F conjugate in KB-8-5 cells, Figure S5: Intracellular accumulation of δ-F and η-F conjugates in RAW264.7 cells and THP-1-induced macrophages.

**Author Contributions:** The first three authors (A.D., O.M. and A.F.) made equal contributions to the work. Conceptualization, D.S.; methodology, M.Z., M.F., D.S. and T.Z.; software, O.M.; validation, D.S., M.F. and M.Z.; formal analysis, O.M. and Y.S.; investigation, A.D. and Y.S.; resources, M.Z. and T.Z.; data curation, O.M., A.F. and Y.S.; writing—original draft preparation, O.M. and D.S.; writing—review and editing, O.M., M.Z. and D.S.; visualization, O.M. and A.D.; supervision, O.M. and A.F.; project administration, M.Z., M.F. and D.S.; funding acquisition, M.Z., D.S. and M.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Russian Science Foundation, grant No. 19-74-30011 (cytotoxicity, confocal microscopy, flow cytometry, dynamic light scattering), Russian Foundation for Basic Research, grant No. 18-515-57006 (siRNA synthesis), the Ministry of Science and Higher Education of the Russian Federation, project No. FSUS-2020-0035 to Novosibirsk State University (oligodeoxynucleotide synthesis), and Japan Medical Research Foundation, grant No. 2018JP002 (biological activity).

**Data Availability Statement:** Data sharing is not applicable to this article.

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

## References

1. Crooke, S.T. Molecular mechanisms of antisense oligonucleotides. *Nucleic Acid Ther.* **2017**, *27*, 70–77. [[CrossRef](#)]
2. Shen, X.; Corey, D.R. Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs. *Nucleic Acids Res.* **2018**, *46*, 1584–1600. [[CrossRef](#)]
3. Smith, C.I.E.; Zain, R. Therapeutic Oligonucleotides: State of the Art. *Annu. Rev. Pharmacol. Toxicol.* **2019**, *59*, 605–630. [[CrossRef](#)] [[PubMed](#)]

4. Evers, M.M.; Toonen, L.J.A.; van Roon-Mom, W.M.C. Antisense oligonucleotides in therapy for neurodegenerative disorders. *Adv. Drug Deliv. Rev.* **2015**, *87*, 90–103. [[CrossRef](#)] [[PubMed](#)]
5. Aung-Htut, M.T.; McIntosh, C.S.; Ham, K.A.; Pitout, I.L.; Flynn, L.L.; Greer, K.; Fletcher, S.; Wilton, S.D. Systematic Approach to Developing Splice Modulating Antisense Oligonucleotides. *Int. J. Mol. Sci.* **2019**, *20*, 5030. [[CrossRef](#)] [[PubMed](#)]
6. Bennett, C.F.; Baker, B.F.; Pham, N.; Swayze, E.; Geary, R.S. Pharmacology of Antisense Drugs. *Annu. Rev. Pharmacol. Toxicol.* **2017**, *57*, 81–105. [[CrossRef](#)]
7. Castanotto, D.; Stein, C.A. Antisense oligonucleotides in cancer. *Curr. Opin. Oncol.* **2014**, *26*, 584–589. [[CrossRef](#)]
8. Wagner, A.; Bock, C.T.; Fechner, H.; Kurreck, J. Application of modified antisense oligonucleotides and siRNAs as antiviral drugs. *Future Med. Chem.* **2015**, *7*, 1637–1642. [[CrossRef](#)]
9. Hegarty, J.P.; Stewart, D.B. Advances in therapeutic bacterial antisense biotechnology. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 1055–1065. [[CrossRef](#)]
10. Shimo, T.; Maruyama, R.; Yokota, T. Designing Effective Antisense Oligonucleotides for Exon Skipping. In *Duchenne Muscular Dystrophy. Methods and Protocols*; Bernardini, C., Ed.; Humana Press: New York, NY, USA, 2018.
11. Ma, C.-C.; Wang, Z.-L.; Xu, T.; He, Z.-Y.; Wei, Y.-Q. The approved gene therapy drugs worldwide: From 1998 to 2019. *Biotechnol. Adv.* **2020**, *40*, 107502. [[CrossRef](#)]
12. Juliano, R.L. The delivery of therapeutic oligonucleotides. *Nucleic Acids Res.* **2016**, *44*, 6518–6548. [[CrossRef](#)] [[PubMed](#)]
13. Roberts, T.C.; Langer, R.; Wood, M.J.A. Advances in oligonucleotide drug delivery. *Nat. Rev. Drug Discov.* **2020**, *19*, 673–694. [[CrossRef](#)] [[PubMed](#)]
14. Pattanayak, S.; Khatra, H.; Saha, S.; Sinha, S. A cationic morpholino antisense oligomer conjugate: Synthesis, cellular uptake and inhibition of Gli1 in the hedgehog signalling pathway. *RSC Adv.* **2014**, *4*, 1951–1954. [[CrossRef](#)]
15. Hnedzko, D.; McGee, D.W.; Karamitas, Y.A.; Rozners, E. Sequence-selective recognition of double-stranded RNA and enhanced cellular uptake of cationic nucleobase and backbone-modified peptide nucleic acids. *RNA* **2017**, *23*, 58–69. [[CrossRef](#)] [[PubMed](#)]
16. Wang, M.; Wu, B.; Shah, S.N.; Lu, P.; Lu, Q. Polyquaternium-mediated delivery of morpholino oligonucleotides for exon-skipping in vitro and in mdx mice. *Drug Deliv.* **2017**, *24*, 952–961. [[CrossRef](#)] [[PubMed](#)]
17. Morcos, P.A.; Li, Y.; Jiang, S. Vivo-Morpholinos: A non-peptide transporter delivers Morpholinos into a wide array of mouse tissues. *Biotechniques* **2018**, *45*, 613–623. [[CrossRef](#)]
18. Craig, K.; Abrams, M.; Amiji, M. Recent preclinical and clinical advances in oligonucleotide conjugates. *Expert Opin. Drug Deliv.* **2018**, *15*, 629–640. [[CrossRef](#)]
19. Ren, T.; Song, Y.K.; Zhang, G.; Liu, D. Structural basis of DOTMA for its high intravenous transfection activity in mouse. *Gene Ther.* **2000**, *7*, 764–768. [[CrossRef](#)]
20. Patwa, A.; Gissot, A.; Bestel, I.; Barthélémy, P. Hybrid lipid oligonucleotide conjugates: Synthesis, self-assemblies and biomedical applications. *Chem. Soc. Rev.* **2011**, *40*, 5844–5854. [[CrossRef](#)]
21. Zhao, B.; Tian, Q.; Bagheri, Y.; You, M. Lipid-oligonucleotide conjugates for simple and efficient cell membrane engineering and bioanalysis. *Curr. Opin. Biomed. Eng.* **2020**, *13*, 76–83. [[CrossRef](#)]
22. Li, X.; Feng, K.; Li, L.; Yang, L.; Pan, X.; Yazd, H.S.; Cui, C.; Li, J.; Moroz, L.; Sun, Y.; et al. Lipid-oligonucleotide conjugates for bioapplications. *Nat. Sci. Rev.* **2020**, *7*, 1933–1953. [[CrossRef](#)]
23. Osborn, M.F.; Khvorova, A. Improving siRNA Delivery In Vivo Through Lipid Conjugation. *Nucleic Acid Ther.* **2018**, *28*, 128–136. [[CrossRef](#)] [[PubMed](#)]
24. Østergaard, M.E.; Jackson, M.; Low, A.; Chappell, A.E.; Lee, R.G.; Peralta, R.Q.; Yu, J.; Kinberger, G.A.; Dan, A.; Carty, R.; et al. Conjugation of hydrophobic moieties enhances potency of antisense oligonucleotides in the muscle of rodents and non-human primates. *Nucleic Acids Res.* **2019**, *47*, 6045–6058. [[CrossRef](#)] [[PubMed](#)]
25. Chernikov, I.V.; Meschaninova, M.L.; Chernolovskaya, E.L. Preparation, Determination of Activity, and Biodistribution of Cholesterol-Containing Nuclease-Resistant siRNAs In Vivo. In *RNA Interference and CRISPR Technologies. Technical Advances and New Therapeutic Opportunities*; Sioud, M., Ed.; Humana Press: New York, NY, USA, 2020.
26. Prokhorova, D.V.; Chelobanov, B.P.; Burakova, E.A.; Fokina, A.A.; Stetsenko, D.A. New oligodeoxyribonucleotide derivatives bearing internucleotide *N*-tosyl phosphoramidate groups: Synthesis and complementary binding to DNA and RNA. *Russ. J. Bioorg. Chem.* **2017**, *43*, 38–42. [[CrossRef](#)]
27. Chelobanov, B.P.; Burakova, E.A.; Prokhorova, D.V.; Fokina, A.A.; Stetsenko, D.A. New oligodeoxynucleotide derivatives containing *N*-(methanesulfonyl)-phosphoramidate (mesyl phosphoramidate) internucleotide group. *Russ. J. Bioorg. Chem.* **2017**, *43*, 664–668. [[CrossRef](#)]
28. Burakova, E.A.; Derzhalova, A.S.; Chelobanov, B.P.; Fokina, A.A.; Stetsenko, D.A. New Oligodeoxynucleotide Derivatives Containing *N*-(Sulfonyl)-Phosphoramidate Groups. *Russ. J. Bioorg. Chem.* **2017**, *45*, 662–668. [[CrossRef](#)]
29. Kitayama, M.; Ueno, M.; Itakura, T.; Yamashita, T. Activated Microglia Inhibit Axonal Growth through RGMA. *PLoS ONE* **2011**, *6*, e25234. [[CrossRef](#)]
30. Farzan, V.M.; Ulashchik, E.A.; Martynenko-Makaev, Y.V.; Kvach, M.V.; Aparin, I.O.; Brylev, V.A.; Prikazchikova, T.A.; Maklakova, S.Y.; Majouga, A.G.; Ustinov, A.V.; et al. Automated Solid-Phase Click Synthesis of Oligonucleotide Conjugates: From Small Molecules to Diverse *N*-Acetylgalactosamine Clusters. *Bioconjug. Chem.* **2017**, *28*, 2599–2607. [[CrossRef](#)]
31. Matano, Y.; Ohkubo, H.; Honsho, Y.; Saito, A.; Seki, S.; Imahori, H. Synthesis and Charge-Carrier Transport Properties of Poly(phosphole *P*-alkanesulfonylimide)s. *Org. Lett.* **2013**, *15*, 932–935. [[CrossRef](#)]

32. Kamerkar, S.; LeBleu, V.S.; Sugimoto, H.; Yang, S.; Ruivo, C.F.; Melo, S.A.; Lee, J.J.; Kalluri, R. Exosomes facilitate therapeutic targeting of oncogenic KRAS in pancreatic cancer. *Nature* **2017**, *546*, 498–503. [[CrossRef](#)]
33. Song, E.; Lee, S.K.; Wang, J.; Ince, N.; Ouyang, N.; Min, J.; Chen, J.; Lieberman, S.P. RNA interference targeting Fas protects mice from fulminant hepatitis. *J. Nat. Med.* **2003**, *9*, 347–351. [[CrossRef](#)] [[PubMed](#)]
34. Markov, O.V.; Filatov, A.V.; Kupryushkin, M.S.; Chernikov, I.V.; Patutina, O.A.; Strunov, A.A.; Chernolovskaya, E.L.; Vlassov, V.V.; Pyshnyi, D.V.; Zenkova, M.A. Transport Oligonucleotides—A Novel System for Intracellular Delivery of Antisense Therapeutics. *Molecules* **2020**, *25*, 3663. [[CrossRef](#)] [[PubMed](#)]
35. Kubo, T.; Tokita, S.; Yamashita, T. Repulsive guidance molecule-a and demyelination: Implications for multiple sclerosis. *J. Neuroimmune Pharmacol.* **2012**, *7*, 524–528. [[CrossRef](#)] [[PubMed](#)]
36. Mothe, A.J.; Tassew, N.G.; Shabanzadeh, A.P.; Penheiro, R.; Vigouroux, R.J.; Huang, L.; Grinnell, C.; Cui, Y.F.; Fung, E.; Monnier, P.P.; et al. RGMa inhibition with human monoclonal antibodies promotes regeneration, plasticity and repair, and attenuates neuropathic pain after spinal cord injury. *Sci. Rep.* **2017**, *7*, 10529. [[CrossRef](#)] [[PubMed](#)]
37. Miroshnichenko, S.K.; Patutina, O.A.; Burakova, E.A.; Chelobanov, B.P.; Fokina, A.A.; Vlassov, V.V.; Altman, S.; Zenkova, M.A.; Stetsenko, D.A. Methyl phosphoramidate antisense oligonucleotides as an alternative to phosphorothioates: Improved biochemical and biological properties. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 1229–1234. [[CrossRef](#)] [[PubMed](#)]
38. Patutina, O.A.; Gaponova (Miroshnichenko), S.K.; Sen'kova, A.V.; Savin, I.A.; Gladkikh, D.V.; Burakova, E.A.; Fokina, A.A.; Maslov, M.A.; Shmendel', E.V.; Wood, M.J.A.; et al. Methyl phosphoramidate backbone modified antisense oligonucleotides targeting miR-21 with enhanced in vivo therapeutic potency. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, in press. [[CrossRef](#)]
39. Kupryushkin, M.S.; Pyshnyi, D.V.; Stetsenko, D.A. Phosphoryl Guanidines: A New Type of Nucleic Acid Analogues. *Acta Nat.* **2014**, *6*, 116–118. [[CrossRef](#)]
40. Ezzat, K.; Aoki, Y.; Koo, T.; McClorey, G.; Benner, L.; Coenen-Stass, A.; O'Donovan, L.; Lehto, T.; Garcia-Guerra, A.; Nordin, J.; et al. Self-Assembly into Nanoparticles Is Essential for Receptor Mediated Uptake of Therapeutic Antisense Oligonucleotides. *Nano Lett.* **2015**, *15*, 4364–4373. [[CrossRef](#)]
41. Deo, S.; Patel, T.R.; Dzananovic, E.; Booy, E.P.; Zeid, K.; McEleney, K.; Harding, S.E.; McKenna, S.A. Activation of 2'5'-Oligoadenylate Synthetase by Stem Loops at the 5'-End of the West Nile Virus Genome. *PLoS ONE* **2014**, *9*, e92545. [[CrossRef](#)]
42. Awino, J.K.; Gudipati, S.; Hartmann, A.K.; Santiana, J.J.; Cairns-Gibson, D.F.; Gomez, N.; Rouge, J.L. Nucleic Acid Nanocapsules for Enzyme-Triggered Drug Release. *J. Am. Chem. Soc.* **2017**, *139*, 6278–6281. [[CrossRef](#)]
43. Soutschek, J.; Akinc, A.; Bramlage, B.; Charisse, K.; Constien, R.; Donoghue, M.; Elbashir, S.; Geick, A.; Hadwiger, P.; Harborth, J.; et al. Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **2004**, *432*, 173–178. [[CrossRef](#)] [[PubMed](#)]
44. Karaki, S.; Benizri, S.; Mejias, R.; Baylot, V.; Branger, N.; Nguyen, T.; Vialet, B.; Oumzil, K.; Barthélémy, P.; Rocchi, P. Lipid-oligonucleotide conjugates improve cellular uptake and efficiency of TCTP-antisense in castration-resistant prostate cancer. *J. Control. Release* **2017**, *258*, 1–9. [[CrossRef](#)] [[PubMed](#)]
45. Zhou, Y.; Zhu, F.; Liu, Y.; Zheng, M.; Wang, Y.; Zhang, D.; Anraku, Y.; Zou, Y.; Li, J.; Wu, H.; et al. Blood-brain barrier-penetrating siRNA nanomedicine for Alzheimer's disease therapy. *Sci. Adv.* **2020**, *6*, eabc7031. [[CrossRef](#)] [[PubMed](#)]
46. Sathyaprakash, C.; Manzano, R.; Varela, M.A.; Hashimoto, Y.; Wood, M.J.A.; Talbot, K.; Aoki, Y. Development of LNA Gapmer Oligonucleotide-Based Therapy for ALS/FTD Caused by the C9orf72 Repeat Expansion. In *Gapmers: Methods and Protocols*; Yokota, T., Maruyama, R., Eds.; Springer: New York, NY, USA, 2020; pp. 185–208.
47. Scoles, D.R.; Meera, P.; Schneider, M.D.; Paul, S.; Dansithong, W.; Figueroa, K.P.; Hung, G.; Rigo, F.; Bennett, C.F.; Otis, T.S.; et al. Antisense oligonucleotide therapy for spinocerebellar ataxia type 2. *Nature* **2017**, *544*, 362–366. [[CrossRef](#)] [[PubMed](#)]
48. Adams, D.; Gonzalez-Duarte, A.; O'Riordan, W.D.; Yang, C.C.; Ueda, M.; Kristen, A.V.; Tournev, I.; Schmidt, H.H.; Coelho, T.; Berk, J.L.; et al. Patisiran, an RNAi therapeutic, for hereditary transthyretin amyloidosis. *N. Engl. J. Med.* **2018**, *379*, 11–21. [[CrossRef](#)] [[PubMed](#)]
49. Scott, L.J. Givosiran: First Approval. *Drugs* **2020**, *80*, 335–339. [[CrossRef](#)]
50. Heo, Y.-A. Golodirsén: First Approval. *Drugs* **2020**, *80*, 329–333. [[CrossRef](#)]
51. Paik, J.; Duggan, S. Volanesorsen: First Global Approval. *Drugs* **2019**, *79*, 1349–1354. [[CrossRef](#)]
52. Keam, S.J. Inotersen: First Global Approval. *Drugs* **2018**, *78*, 1371–1376. [[CrossRef](#)]
53. Remaut, K.; Lucas, B.; Braeckmans, K.; Demeester, J.; De Smedt, S.C. Pegylation of liposomes favours the endosomal degradation of the delivered phosphodiester oligonucleotides. *J. Control. Release* **2007**, *117*, 256–266. [[CrossRef](#)]
54. Chen, J.; Yu, Z.; Chen, H.; Gao, J.; Liang, W. Transfection efficiency and intracellular fate of polycation liposomes combined with protamine. *Biomaterials* **2011**, *32*, 1412–1418. [[CrossRef](#)] [[PubMed](#)]
55. Ara, M.N.; Matsuda, T.; Hyodo, M.; Sakurai, Y.; Hatakeyama, H.; Ohga, N.; Hida, K.; Harashima, H. An aptamer ligand based liposomal nanocarrier system that targets tumor endothelial cells. *Biomaterials* **2014**, *35*, 7110–7120. [[CrossRef](#)] [[PubMed](#)]
56. Varkouhi, A.K.; Scholte, M.; Storm, G.; Haisma, H.J. Endosomal escape pathways for delivery of biologicals. *J. Control. Release* **2011**, *151*, 220–228. [[CrossRef](#)] [[PubMed](#)]
57. Petrova, N.S.; Chernikov, I.V.; Meschaninova, M.I.; Dovydenko, I.S.; Venyaminova, A.G.; Zenkova, M.A.; Vlassov, V.V.; Chernolovskaya, E.L. Carrier-free cellular uptake and the gene-silencing activity of the lipophilic siRNAs is strongly affected by the length of the linker between siRNA and lipophilic group. *Nucleic Acids Res.* **2012**, *40*, 2330–2344. [[CrossRef](#)] [[PubMed](#)]
58. Dobson, R.; Giovannoni, G. Multiple sclerosis—A review. *Eur. J. Neurol.* **2019**, *26*, 27–40. [[CrossRef](#)]



Communication

# Investigation of the Characteristics of NLS-PNA: Influence of NLS Location on Invasion Efficiency

Yuichiro Aiba <sup>\*,†</sup>, Gerardo Urbina <sup>†</sup>, Masanari Shibata and Osami Shoji <sup>\*</sup>

Department of Chemistry, Graduate School of Science, Nagoya University, Furo-Cho, Chikusa-Ku, Nagoya, Aichi 464-8602, Japan; gerardoaus@gmail.com (G.U.); shibata.masanari@b.mbox.nagoya-u.ac.jp (M.S.)

<sup>\*</sup> Correspondence: aiba.yuichiro@i.mbox.nagoya-u.ac.jp (Y.A.); shoji.osami@a.mbox.nagoya-u.ac.jp (O.S.)

<sup>†</sup> These authors equally contributed to this work.

Received: 1 November 2020; Accepted: 30 November 2020; Published: 3 December 2020

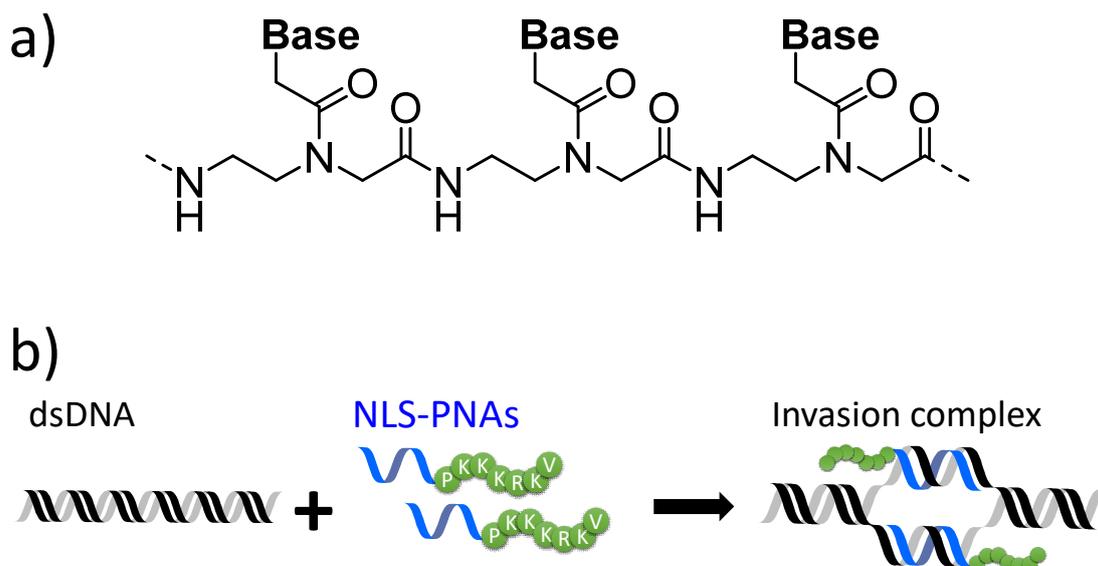


**Abstract:** Peptide nucleic acid can recognise sequences in double-stranded DNA (dsDNA) through the formation of a double-duplex invasion complex. This double-duplex invasion is a promising method for the recognition of dsDNA in cellula because peptide nucleic acid (PNA) invasion does not require the prior denaturation of dsDNA. To increase its applicability, we developed PNAs modified with a nuclear localisation signal (NLS) peptide. In this study, the characteristics of NLS-modified PNAs were investigated for the future design of novel peptide-modified PNAs.

**Keywords:** PNA; invasion; DNA; NLS

## 1. Introduction

The sequence-specific recognition of double-stranded DNA (dsDNA) has been an important research subject owing to its wide range of potential applications [1–7]. For this purpose, various methods have been developed, including those that make use of DNA-binding proteins [1,3–5,8–12], small molecules such as minor groove binders [13–16], and artificial DNA [17–24]. In addition, the recognition of dsDNA by peptide nucleic acid (PNA), an artificial nucleic acid mimic, has been reported by Nielsen et al. [7,21,24–31]. In 1991, PNA was first designed and synthesised as an analogue of natural nucleic acids, where the negatively charged sugar-phosphate backbone of DNA was substituted with an electrostatically neutral artificial *N*-(2-aminoethyl)glycine backbone (Figure 1a) [25]. Consequently, electrostatic repulsion between PNA and DNA is absent, enabling PNA to form more stable duplexes with complementary DNA via Watson–Crick base pairing than those between complementary DNA strands [31]. Furthermore, when pseudo-complementary PNAs (pcPNAs: see Figure S1)—where conventional adenine (A) and thymine (T) have been replaced by 2,6-diaminopurine (D) and 2-thiouracil (U)—are used, an effective invasion into dsDNA becomes possible, with two strands of pcPNA forming a double-duplex invasion complex (or, simply, invasion complex, Figure 1b) with dsDNA [26,32]. The introduction of D and U into PNA destabilises the formation of PNA/PNA duplexes through steric repulsions between the amino group of D and the thioketone group of U. The thermal stability of the PNA/DNA duplex is enhanced by the formation of stable base pairs between D/U and the natural nucleobases. PNA's unique DNA-recognition mode has been used in many applications, including site-directed mutagenesis [2,7], inhibition of enzymatic activity [26,33,34], and the development of artificial DNA cutters [35], with all of these depending on the sequence specificity of the invasion complex. The pseudopeptide backbone of PNA facilitates easy oligomer synthesis and modification for further improvements and imparts enhanced resistance to enzymatic degradation. Thus, double-duplex invasion is a promising method for the targeting of dsDNA in cellula.



**Figure 1.** (a) Chemical structure of peptide nucleic acid (PNA). (b) Stable invasion complex formation using nuclear-localisation-signal-modified PNAs (NLS-PNAs).

However, to realise such applications, an important issue must be overcome—namely, the high salt concentrations encountered in the cellular environment. High salt concentrations increase the stability of the DNA/DNA duplexes whilst slightly destabilising the PNA/DNA duplexes [36], which results in an overall decrease in invasion efficiency [37]. Several approaches have been attempted to enhance invasion efficiency in physiological salt concentrations, such as chemical modification of PNA and conjugation with functional molecules [37–41]. Our group developed conjugates of PNA with Ru-complexes and PNAs modified with a nuclear localisation signal (NLS) peptide to increase their applicability for DNA recognition at high salt concentrations (Figure 1b). These two modified PNAs displayed higher DNA affinity than corresponding unmodified pcPNAs and effectively formed an invasion complex even at physiological salt concentrations, whereunder unmodified PNAs stop functioning efficiently. It is important to understand the properties and limitations of these modified PNAs in order to further improve PNAs. Of these two methods, NLS modification is the simpler PNA modification method as the resulting NLS-modified PNAs (NLS-PNAs) can be procured by simply conjugating PNA with the functional peptide, and can be readily introduced during the synthesis of PNAs. Thus, in this study, we focused on NLS-PNAs and investigated their characteristics to obtain insight into how they can be improved.

## 2. Materials and Methods

### 2.1. Materials

Solvents and reagents were purchased from FUJIFILM Wako Pure Chemical Co. (Tokyo, Japan); Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan); Merck (Darmstadt, Germany); Sigma-Aldrich Co., LLC. (St. Louis, MO, USA); nacalai tesque Inc. (Kyoto, Japan); and Kanto Chemical Co., Inc. (Tokyo, Japan); and were used without further purification. PNA monomers were purchased from ASM Research Chemicals (Hannover, Germany). Oligonucleotides were purchased from Fasmac (Kanagawa, Japan).

PNAs with or without an NLS peptide were synthesised by standard Boc-chemistry-based solid-phase peptide synthesis according to a literature procedure [42], purified by reversed-phase high-performance liquid chromatography (HPLC) (Jasco Co.; Tokyo, Japan), and characterised by matrix-assisted laser desorption/ionization–time of flight mass spectrometer (MALDI-TOF MS) (ultraflex III, Bruker Daltonics; Billerica, MA, USA) (Figures S2 and S3). The concentration of PNAs and DNAs was determined based on their absorbance at 260 nm by using a UV-visible spectrophotometer

(molar extinction coefficients for PNA monomers (in  $M^{-1} \text{ cm}^{-1}$ ):  $\epsilon(D) = 7600$ ;  $\epsilon(C) = 6600$ ;  $\epsilon(G) = 11,700$ ;  $\epsilon(U) = 10,200$ ).

## 2.2. Invasion Experiments

Target 130-bp DNA from pBR322 was incubated with a pair of pcPNAs in a 5 mM HEPES buffer (pH 7.0) at 50 °C for 1 h. After incubation, the solutions were subjected to a microchip electrophoresis system (MultiNA, Shimadzu; Kyoto, Japan) and the invasion efficiency was evaluated.

## 3. Results and Discussion

### 3.1. A Positive Effect on Invasion Efficiency by the Introduction of an NLS into Short PNAs

In a previous report, we successfully demonstrated that conjugation of an NLS to pcPNAs results in enhanced invasion efficiency [37]. Specifically, an NLS peptide with the sequence PKKKRKV found in the SV40 virus, which has been widely studied, was conjugated at the C-terminus of a pentadecamer PNA. If the introduction of an NLS, even in short PNAs, exerts a positive effect upon invasion efficiency, then the properties of NLS-PNAs should be more easily elucidated by examining them with short PNAs, whose synthesis requires less effort than longer PNAs. Thus, we checked the effectiveness of the NLS in decamer PNAs and employed two types of PNA: (1) unmodified PNAs (U-PNAs), i.e., a pair of decamer pcPNAs without an NLS (with only a few lysine residues for solubility) and (2) C-PNAs, which are identical in sequence to U-PNAs but with an NLS at their C-termini.

U-PNAs and C-PNAs were synthesised by standard *tert*-butoxycarbonyl (Boc)-based solid-phase peptide synthesis. The PNA sequences are listed in Table 1. PNA1 and PNA2 represent each strand of pcPNA that is complementary to 10 bp in the pBR322 plasmid (PNA1: GUUDCUGDUG and PNA2: CDUCDGUDDC). All PNAs have a free N-terminal amino group, with the C-terminal carboxylic acid being converted to an amide. The efficiency of invasion complex formation was evaluated by an electrophoresis mobility shift assay (EMSA). The formation of the invasion complex results in a lower electrophoretic mobility, mainly due to changes in the local structure of the dsDNA at the invasion site [43]. This allows the separation of the invasion complex from the free dsDNA. An EMSA of the invasion complex formation in the presence of 75 mM NaCl is shown in Figure 2.

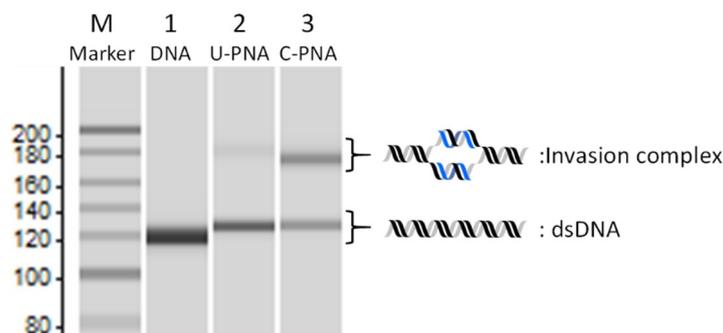
**Table 1.** PNAs synthesised and used in this research.

Name	Sequences <sup>[a]</sup> (N to C)
U-PNA1	H- <b>K</b> GUUDCUGDUG <b>KK</b> -NH <sub>2</sub>
U-PNA2	H- <b>K</b> CDUCDGUDDC <b>KK</b> -NH <sub>2</sub>
N-PNA1	H- <i><b>PKKKRKV</b></i> GUUDCUGDUG <b>KK</b> -NH <sub>2</sub>
N-PNA2	H- <i><b>PKKKRKV</b></i> CDUCDGUDDC <b>KK</b> -NH <sub>2</sub>
C-PNA1	H- <b>K</b> GUUDCUGDUG <i><b>PKKKRKV</b></i> -NH <sub>2</sub>
C-PNA2	H- <b>K</b> CDUCDGUDDC <i><b>PKKKRKV</b></i> -NH <sub>2</sub>
N-Pro-PNA1	H- <i><b>PKKKRKVP</b></i> GUUDCUGDUG <b>KK</b> -NH <sub>2</sub>
N-Pro-PNA2	H- <i><b>PKKKRKVP</b></i> CDUCDGUDDC <b>KK</b> -NH <sub>2</sub>

<sup>[a]</sup> Bold: PNAs; Italic: amino acids; K = lysine; P = proline; R = arginine; V = valine; D = 2,6-diaminopurine; U = 2-thiouracil.

In lane 1, a single band corresponding to free dsDNA was observed between 100 and 150 bp. In the presence of U-PNA, the invasion complex was barely formed under the high-salt conditions employed, giving only a faint band with lower mobility between 180 and 200 bp (lane 2), which was assignable to the invasion complex. However, the C-PNAs with NLS at their C-termini gave an intense band corresponding to the invasion complex (lane 3). It is clear from these results that the invasion efficiency was improved by the conjugation of an NLS to the short PNAs, leading to the formation of a

stable invasion complex even at elevated salt concentrations. As a result, it was confirmed that the NLS modifications in short PNAs showed a beneficial effect on invasion. Thus, we decided to investigate the properties of NLS-PNAs by using these short decamer PNAs. The effect of NLS modification appears to be mainly attributable to electrostatic interactions. However, the conjugation of PNA with repeating cationic peptides (e.g., polylysine) also increases non-specific interactions with the DNA backbone, resulting in a negative effect on the formation of the invasion complex [44]. Presumably, in biological motifs, unfavourable sequences (e.g., those provoking non-specific binding to DNA) have been evolutionarily excluded. Thus, we can expect that NLS peptides possess an ideal amino acid sequence that shows a moderate affinity to DNA as well as a positive effect on invasion efficiency.



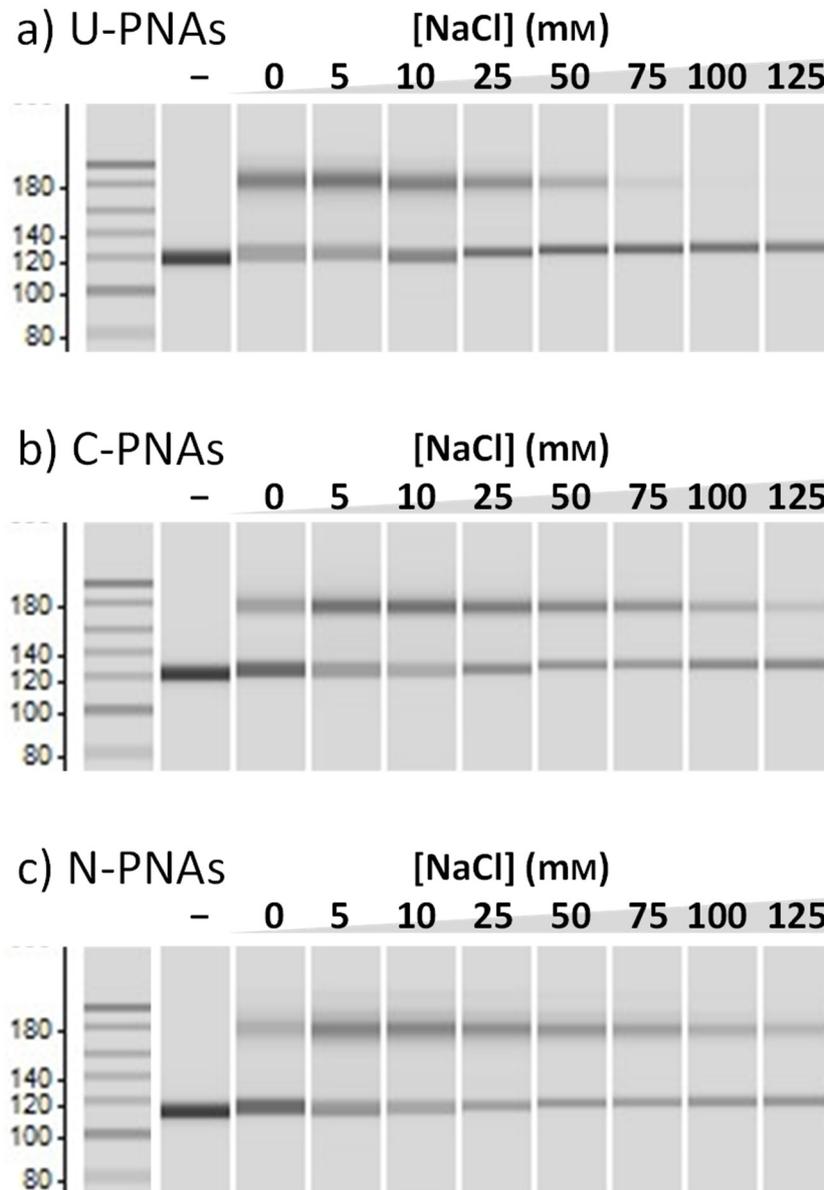
**Figure 2.** Comparison of U-PNA and C-PNA invasion efficiency in the presence of 75 mM NaCl. M: 20-bp DNA Ladder (TaKaRa); lane 1: 130-bp DNA only; lane 2 U-PNA; lane 3: C-PNA. Invasion conditions: [DNA] = [each PNA] = 50 nM, [HEPES (pH 7.0)] = 5 mM, and [NaCl] = 75 mM at 50 °C for 1 h.

### 3.2. Changing the Location of the NLS from the C-Terminus to the N-Terminus of the PNAs

As mentioned above, in the previous report on NLS-PNAs, the NLS peptide was conjugated to the C-termini of the PNAs (C-PNAs). The synthesis of NLS-PNA was carried out from the C-terminus to the N-terminus by solid-phase peptide synthesis, and the NLS peptide was introduced to the resin first, followed by the PNA. When examining various functional peptides as alternatives to NLS for future research, N-terminally modified PNA conjugates, rather than C-terminally modified PNA conjugates, would significantly reduce the time and effort required for investigation. This is because it is possible to synthesise a large number of peptide-modified PNAs at once by elongating the PNAs first and then splitting the resin followed by peptide conjugation. Therefore, we evaluated whether changing the location of the NLS from the C-terminus to the N-terminus of the PNA (N-PNA) would have a detrimental effect upon invasion efficiency. PNAs modified with NLS peptide at the N-termini (N-PNAs) were synthesised using the same method as C-PNAs, and the invasion experiments with a series of PNAs were carried out under a broad range of salt concentrations, ranging from 0 to 125 mM NaCl.

As a control, the invasion efficiency of U-PNAs was also examined at increasing salt concentrations. Figure 3 shows that the band intensity of the invasion complex with U-PNA decreased with increasing salt concentration and almost disappeared when the NaCl concentration exceeded 50 mM. This result reaffirms the aforementioned phenomenon, wherein elevated salt concentrations stabilise dsDNA whilst weakening PNA/DNA interaction, resulting in a lower invasion efficiency. When the same experiment was conducted using C-PNAs, however, a high invasion efficiency beyond 50 mM NaCl was observed. Unlike in our previous report with pentadecamer NLS-PNAs, in the present study, a decrease in invasion efficiency was observed above 75 mM NaCl. This is due to the lower concentration of PNA, which we selected herein to amplify any effect. The introduction of NLS, even in shorter PNAs, was once more confirmed to be beneficial for invasion at high salt concentrations. Next, the effect of NLS position on invasion efficiency was investigated using N-PNAs. At first glance, both C-PNAs and N-PNAs exhibited similar behaviour in terms of invasion efficiency. The invasion efficiencies of both NLS-PNAs remained high, up to about 75 mM NaCl, although this value gradually decreased with

increasing salt concentration. This comparison confirms that, regardless of the location of the NLS, PNA-NLS conjugates display similar performance. However, N-PNAs were slightly better in terms of resistance to higher salt concentrations (over 75 mM) than C-PNAs, with a more pronounced effect at very high salt concentrations of 125 mM.

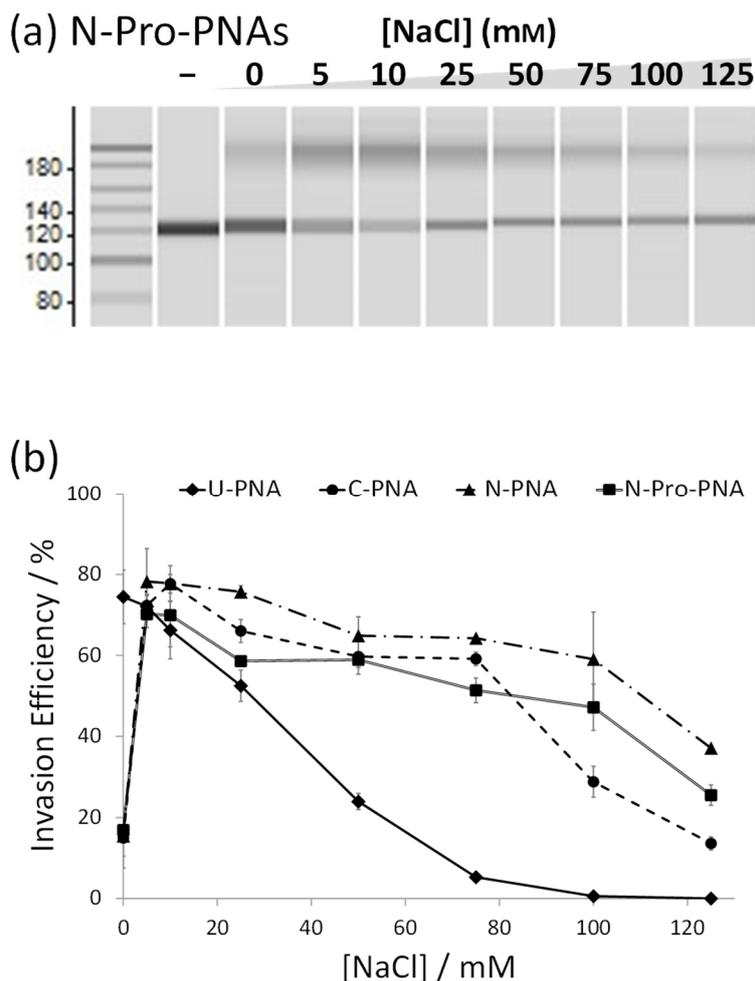


**Figure 3.** Comparison of (a) U-PNAs', (b) C-PNAs', and (c) N-PNAs' invasion efficiency. Invasion conditions: [DNA] = [each PNA] = 50 nM, [HEPES (pH 7.0)] = 5 mM, and [NaCl] = 0–125 mM at 50 °C for 1 h.

### 3.3. Insertion of Proline between the PNAs and Valine of the NLS

Besides the location of the NLS itself, there is a further difference between C-PNAs and N-PNAs. In order to maintain the directionality of the NLS, in C-PNAs, proline (P) of the NLS connects to the PNA, whereas in N-PNAs, the NLS is connected via valine (V). Unlike valine, proline possesses a unique cyclic structure bestowing it with a much higher backbone rigidity than any of the other canonical amino acids. In order to verify its significance, we synthesised N-Pro-PNAs, which are identical to N-PNAs except that a proline residue has been inserted between the PNA and valine. Interestingly, the insertion of proline proved unfavourable for invasion efficiency, and N-Pro-PNAs

displayed invasion properties between those of N-PNAs and C-PNAs (Figure 4). These results strongly indicate that the rigid amino acid connecting the NLS to PNA may unfavourably affect invasion.



**Figure 4.** (a) Invasion experiments for N-Pro-PNAs to evaluate the effect of an amino acid between NLS and PNA. Invasion conditions: [DNA] = [each PNA] = 50 nM, [HEPES (pH 7.0)] = 5 mM, and [NaCl] = 0–125 mM at 50 °C for 1 h. (b) Effect of salt concentration on invasion efficiency with various PNAs (U-, C-, N-, and N-Pro-PNAs). The value for invasion efficiency was obtained by comparing the band intensities in a microchip electrophotogram.

#### 4. Conclusions

In summary, we investigated the characteristics of NLS-PNA by employing a series of NLS-modified PNAs. These NLS-PNA conjugates were found to be superior to unmodified PNAs at elevated salt concentrations. The difference between C-terminally modified and N-terminally modified NLS-PNAs was not very significant, signifying that N-terminally modified PNA-NLS conjugates can be readily used for research purposes using the same method as previously reported C-terminally modified conjugates. This facilitates the development of new functional peptide-modified PNA conjugates, since solid-phase peptide synthesis requires synthesis from the C-terminus to the N-terminus. As an example, for C-terminally modified PNAs, when multiple functional peptides bound to the same PNA need to be screened, each functional peptide must be synthesised separately. In the case of N-terminal modification, the synthesis of peptide-modified PNAs can be simplified by synthesising the individual PNAs only once, separating the resin beads into as many samples as needed and continuing the synthesis of each NLS. In addition to the finding that the position of the NLS has some influence on

invasion, we demonstrated that the amino acids between the NLS and PNA also play an important role. These findings also provide new perspectives for the future design of novel peptide-modified PNAs.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/23/8663/s1>, Figure S1: chemical structures of 2,6-diaminopurine (D) and 2-thouracil (U), Figure S2: HPLC charts of purified PNAs, Figure S3: MALDI-TOF MS spectra of purified PNAs.

**Author Contributions:** Investigation, G.U. and M.S.; writing—original draft preparation, Y.A. and G.U.; writing—review and editing, Y.A., G.U., M.S. and O.S.; supervision, Y.A. and O.S.; project administration, Y.A.; funding acquisition, Y.A. and O.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by JSPS KAKENHI grant no. 19K05730 to Y.A. from the Ministry of Education, Culture, Sports, Science, and Technology (Japan). This work was also supported in part by the Izumi Science and Technology Foundation to Y.A., Kato Memorial Bioscience Foundation to Y.A., and Integrated Research Consortium on Chemical Sciences to Y.A.

**Acknowledgments:** We thank Joshua Kyle Stanfield for checking the English of this manuscript.

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

## References

1. Kim, Y.G.; Cha, J.; Chandrasegaran, S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1156–1160. [[CrossRef](#)] [[PubMed](#)]
2. Kim, K.H.; Nielsen, P.E.; Glazer, P.M. Site-directed gene mutation at mixed sequence targets by psoralen-conjugated pseudo-complementary peptide nucleic acids. *Nucleic Acids Res.* **2007**, *35*, 7604–7613. [[CrossRef](#)] [[PubMed](#)]
3. Urnov, F.D.; Rebar, E.J.; Holmes, M.C.; Zhang, H.S.; Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **2010**, *11*, 636–646. [[CrossRef](#)] [[PubMed](#)]
4. Christian, M.; Cermak, T.; Doyle, E.L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A.J.; Voytas, D.F. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* **2010**, *186*, 757–761. [[CrossRef](#)] [[PubMed](#)]
5. Qi, L.S.; Larson, M.H.; Gilbert, L.A.; Doudna, J.A.; Weissman, J.S.; Arkin, A.P.; Lim, W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **2013**, *152*, 1173–1183. [[CrossRef](#)]
6. Komiyama, M.; Yoshimoto, K.; Sisido, M.; Ariga, K. Chemistry Can Make Strict and Fuzzy Controls for Bio-Systems: DNA Nanoarchitectonics and Cell-Macromolecular Nanoarchitectonics. *Bull. Chem. Soc. Jpn.* **2017**, *90*, 967–1004. [[CrossRef](#)]
7. Economos, N.G.; Oyaghire, S.; Quijano, E.; Ricciardi, A.S.; Saltzman, W.M.; Glazer, P.M. Peptide Nucleic Acids and Gene Editing: Perspectives on Structure and Repair. *Molecules* **2020**, *25*, 735. [[CrossRef](#)]
8. Wolfe, S.A.; Nekludova, L.; Pabo, C.O. DNA recognition by Cys(2)His(2) zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 183–212. [[CrossRef](#)]
9. Jantz, D.; Amann, B.T.; Gatto, G.J.; Berg, J.M. The design of functional DNA-binding proteins based on zinc finger domains. *Chem. Rev.* **2004**, *104*, 789–799. [[CrossRef](#)]
10. Dhanasekaran, M.; Negi, S.; Sugiura, Y. Designer zinc finger proteins: Tools for creating artificial DNA-binding functional proteins. *Acc. Chem. Res.* **2006**, *39*, 45–52. [[CrossRef](#)]
11. Boch, J.; Scholze, H.; Schornack, S.; Landgraf, A.; Hahn, S.; Kay, S.; Lahaye, T.; Nickstadt, A.; Bonas, U. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* **2009**, *326*, 1509–1512. [[CrossRef](#)] [[PubMed](#)]
12. Xu, X.S.; Oi, L.S. A CRISPR-dCas Tool for Genetic Engineering and Synthetic Biology. *J. Mol. Biol.* **2019**, *431*, 34–47. [[CrossRef](#)] [[PubMed](#)]
13. Dervan, P.B.; Burli, R.W. Sequence-specific DNA recognition by polyamides. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688–693. [[CrossRef](#)]
14. Dervan, P.B. Molecular recognition of DNA by small molecules. *Biorg. Med. Chem.* **2001**, *9*, 2215–2235. [[CrossRef](#)]
15. Dervan, P.B.; Edelson, B.S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284–299. [[CrossRef](#)]

16. Bando, T.; Sugiyama, H. Synthesis and biological properties of sequence-specific DNA-alkylating pyrrole-imidazole polyamides. *Acc. Chem. Res.* **2006**, *39*, 935–944. [[CrossRef](#)]
17. Thuong, N.T.; Hélène, C. Sequence-Specific Recognition and Modification of Double-Helical DNA by Oligonucleotides. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666–690. [[CrossRef](#)]
18. Praseuth, D.; Guieysse, A.L.; Hélène, C. Triple helix formation and the antigene strategy for sequence-specific control of gene expression. *Biochim. Biophys. Acta Gene Struct. Expr.* **1999**, *1489*, 181–206. [[CrossRef](#)]
19. Fox, K.R. Targeting DNA with triplexes. *Curr. Med. Chem.* **2000**, *7*, 17–37. [[CrossRef](#)]
20. Rapireddy, S.; He, G.; Roy, S.; Armitage, B.A.; Ly, D.H. Strand invasion of mixed-sequence B-DNA by acridine-linked, gamma-peptide nucleic acid (gamma-PNA). *J. Am. Chem. Soc.* **2007**, *129*, 15596–15600. [[CrossRef](#)]
21. Vilaivan, T. Pyrrolidiny PNA with alpha/beta-Dipeptide Backbone: From Development to Applications. *Acc. Chem Res.* **2015**, *48*, 1645–1656. [[CrossRef](#)] [[PubMed](#)]
22. Hrdlicka, P.J.; Karmakar, S. 25 years and still going strong: 2'-O-(pyren-1-yl) methylribonucleotides-versatile building blocks for applications in molecular biology, diagnostics and materials science. *Org. Biomol. Chem.* **2017**, *15*, 9760–9774. [[CrossRef](#)] [[PubMed](#)]
23. Nakamura, S.; Kawabata, H.; Fujimoto, K. Double duplex invasion of DNA induced by ultrafast photo-cross-linking using 3-cyanovinylcarbazole for antigene methods. *Chem. Commun.* **2017**, *53*, 7616–7619. [[CrossRef](#)] [[PubMed](#)]
24. Muangkaew, P.; Vilaivan, T. Modulation of DNA and RNA by PNA. *Bioorg. Med. Chem. Lett.* **2020**, *30*, 127064. [[CrossRef](#)] [[PubMed](#)]
25. Nielsen, P.E.; Egholm, M.; Berg, R.H.; Buchardt, O. Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science* **1991**, *254*, 1497–1500. [[CrossRef](#)] [[PubMed](#)]
26. Lohse, J.; Dahl, O.; Nielsen, P.E. Double duplex invasion by peptide nucleic acid: A general principle for sequence-specific targeting of double-stranded DNA. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11804–11808. [[CrossRef](#)] [[PubMed](#)]
27. Demidov, V.V.; Frank-Kamenetskii, M.D. Two sides of the coin: Affinity and specificity of nucleic acid interactions. *Trends Biochem. Sci.* **2004**, *29*, 62–71. [[CrossRef](#)]
28. Nielsen, P.E. Peptide Nucleic Acids (PNA) in Chemical Biology and Drug Discovery. *Chem. Biodivers.* **2010**, *7*, 786–804. [[CrossRef](#)]
29. Sharma, C.; Awasthi, S.K. Versatility of peptide nucleic acids (PNAs): Role in chemical biology, drug discovery, and origins of life. *Chem. Biol. Drug Des.* **2017**, *89*, 16–37. [[CrossRef](#)]
30. Manicardi, A.; Rozzi, A.; Korom, S.; Corradini, R. Building on the peptide nucleic acid (PNA) scaffold: A biomolecular engineering approach. *Supramol. Chem.* **2017**, *29*, 784–795. [[CrossRef](#)]
31. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S.M.; Driver, D.A.; Berg, R.H.; Kim, S.K.; Norden, B.; Nielsen, P.E. PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules. *Nature* **1993**, *365*, 566–568. [[CrossRef](#)] [[PubMed](#)]
32. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P.E. Peptide Nucleic Acids (PNAs) Containing Thymine Monomers Derived from Chiral Amino Acids: Hybridization and Solubility Properties of D-Lysine PNA. *Angew. Chem. Int. Ed.* **1996**, *35*, 1939–1942.
33. Izvol'sky, K.I.; Demidov, V.V.; Nielsen, P.E.; Frank-Kamenetskii, M.D. Sequence-specific protection of duplex DNA against restriction and methylation enzymes by pseudocomplementary PNAs. *Biochemistry* **2000**, *39*, 10908–10913. [[CrossRef](#)] [[PubMed](#)]
34. Protozanova, E.; Demidov, V.V.; Nielsen, P.E.; Frank-Kamenetskii, M.D. Pseudocomplementary PNAs as selective modifiers of protein activity on duplex DNA: The case of type II restriction enzymes. *Nucleic Acids Res.* **2003**, *31*, 3929–3935. [[CrossRef](#)] [[PubMed](#)]
35. Komiyama, M.; Aiba, Y.; Yamamoto, Y.; Sumaoka, J. Artificial restriction DNA cutter for site-selective scission of double-stranded DNA with tunable scission site and specificity. *Nat. Protoc.* **2008**, *3*, 655–662. [[CrossRef](#)]
36. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P.E.; Norden, B.; Graslund, A. Ionic effects on the stability and conformation of peptide nucleic acid complexes. *J. Am. Chem. Soc.* **1996**, *118*, 5544–5552. [[CrossRef](#)]
37. Aiba, Y.; Honda, Y.; Komiyama, M. Promotion of Double-Duplex Invasion of Peptide Nucleic Acids through Conjugation with Nuclear Localization Signal Peptide. *Chem. Eur. J.* **2015**, *21*, 4021–4026. [[CrossRef](#)]

38. Aiba, Y.; Hamano, Y.; Kameshima, W.; Araki, Y.; Wada, T.; Accetta, A.; Sforza, S.; Corradini, R.; Marchelli, R.; Komiyama, M. PNA-NLS conjugates as single-molecular activators of target sites in double-stranded DNA for site-selective scission. *Org. Biomol. Chem.* **2013**, *11*, 5233–5238. [[CrossRef](#)]
39. Aiba, Y.; Ohyama, J.; Komiyama, M. Transfection of PNA-NLS Conjugates into Human Cells Using Partially Complementary Oligonucleotides. *Chem. Lett.* **2015**, *44*, 1547–1549. [[CrossRef](#)]
40. Hibino, M.; Aiba, Y.; Watanabe, Y.; Shoji, O. Peptide Nucleic Acid Conjugated with Ruthenium-Complex Stabilizing Double-Duplex Invasion Complex Even under Physiological Conditions. *ChemBioChem* **2018**, *19*, 1601–1604. [[CrossRef](#)]
41. Hibino, M.; Aiba, Y.; Shoji, O. Cationic guanine: Positively charged nucleobase with improved DNA affinity inhibits self-duplex formation. *Chem. Commun.* **2020**, *56*, 2546–2549. [[CrossRef](#)] [[PubMed](#)]
42. Komiyama, M.; Aiba, Y.; Ishizuka, T.; Sumaoka, J. Solid-phase synthesis of pseudo-complementary peptide nucleic acids. *Nat. Protoc.* **2008**, *3*, 646–654. [[CrossRef](#)] [[PubMed](#)]
43. Kuhn, H.; Cherny, D.I.; Demidov, V.V.; Frank-Kamenetskii, M.D. Inducing and modulating anisotropic DNA bends by pseudocomplementary peptide nucleic acids. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7548–7553. [[CrossRef](#)] [[PubMed](#)]
44. Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. Chiral introduction of positive charges to PNA for double-duplex invasion to versatile sequences. *Nucleic Acids Res.* **2008**, *36*, 1464–1471. [[CrossRef](#)] [[PubMed](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Article

# Construction of an Enzymatically-Conjugated DNA Aptamer–Protein Hybrid Molecule for Use as a BRET-Based Biosensor

Masayasu Mie \*, Rena Hirashima, Yasumasa Mashimo and Eiry Kobatake

Department of Life Science and Technology, School of Life Science and Technology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8502, Japan; nrmsrh410@gmail.com (R.H.); mashimo.y.ab@m.titech.ac.jp (Y.M.); kobatake.e.aa@m.titech.ac.jp (E.K.)

\* Correspondence: mie.m.aa@m.titech.ac.jp

Received: 28 September 2020; Accepted: 27 October 2020; Published: 29 October 2020



**Abstract:** DNA-protein conjugates are useful molecules for construction of biosensors. Herein, we report the development of an enzymatically-conjugated DNA aptamer–protein hybrid molecule for use as a bioluminescence resonance energy transfer (BRET)-based biosensor. DNA aptamers were enzymatically conjugated to a fusion protein via the catalytic domain of porcine circovirus type 2 replication initiation protein (PCV2 Rep) comprising residues 14–109 (tpRep), which was truncated from the full catalytic domain of PCV2 Rep comprising residues 1–116 by removing the flexible regions at the N- and C-terminals. For development of a BRET-based biosensor, we constructed a fusion protein in which tpRep was positioned between NanoLuc luciferase and a fluorescent protein and conjugated to single-stranded DNA aptamers that specifically bind to either thrombin or lysozyme. We demonstrated that the BRET ratios depended on the concentration of the target molecules.

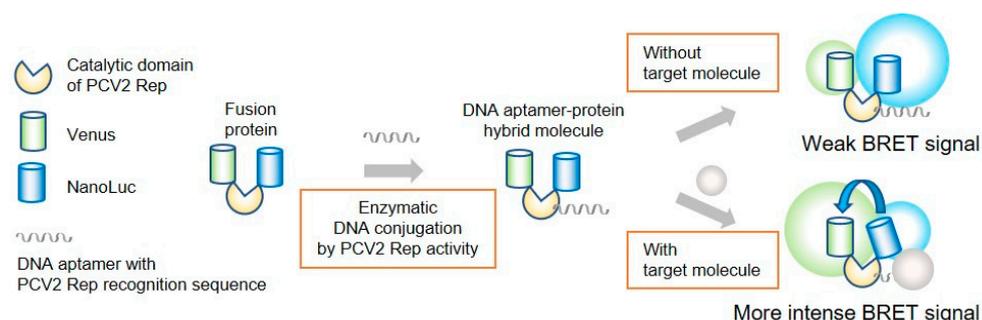
**Keywords:** DNA-protein conjugate; replication initiation protein; DNA aptamer; BRET-based biosensor

## 1. Introduction

DNA-protein conjugates are useful materials for biosensing applications [1]. Use of DNA in these conjugate molecules allows for both signal amplification as well as molecular recognition (e.g., using DNA aptamers) [2–5]. DNA aptamers bind to specific molecules with high affinity and specificity. Therefore, DNA aptamer-based biosensing systems have been constructed as conjugates of DNA aptamers and reporter molecules [6–8]. Single-stranded DNA (ssDNA) containing aptamer sequences are modified at the 5' and/or 3' ends with a fluorophore for use in construction of DNA aptamer-based biosensors [9–12]. In addition to modification of ssDNA with a fluorophore, a reporter protein can also be conjugated to ssDNA [3–5]. However, conventional methods for conjugation of ssDNA to a reporter protein require cumbersome procedures. To overcome these limitations, we developed a method for conjugation of ssDNA to a protein of interest fused with a replication initiation protein (Rep) [5,13]. In this method, a protein fused with Rep can be covalently linked to ssDNA via enzymatic reaction, without the need for any chemical modification of the ssDNA. Recently, the catalytic domain of porcine circovirus type 2 Rep comprising residues 1–116 (pRep) was employed to construct DNA-NanoLuc luciferase (NanoLuc) conjugates [5]. DNA-NanoLuc conjugates were applied for use in a DNA aptamer-sandwich assay system. Moreover, pRep retains its DNA binding activity regardless of whether the protein of interest is fused to the N-terminus or C-terminus.

Bioluminescence resonance energy transfer (BRET)-based biosensors are frequently used in biosensing methods [14–16]. For the design of intramolecular BRET sensors, sensory domains that lead to conformational changes are located between donor and acceptor molecules, such as luciferase

and a fluorescent protein. Therefore, we decided to design a fusion protein for application as a BRET biosensor, in which pRep was positioned between luciferase and a fluorescent protein, and then a DNA aptamer was conjugated via pRep. It was anticipated that our designed fusion protein conjugates containing DNA aptamers might undergo conformational changes upon binding of a target molecule to the DNA aptamer (Figure 1). BRET efficiency is dependent on not only the distance between the donor and acceptor molecules, but also their relative orientations. In our design, it was speculated that larger change in the distance between donor and acceptor molecules would not be occurred even when a target molecule binds to the DNA aptamer. However, the BRET efficiency was expected to change through alteration of their orientation upon binding of a target molecule to the DNA aptamer.



**Figure 1.** BRET-based biosensor with DNA-protein conjugates. The catalytic domain PCV2 Rep was fused with Venus and NanoLuc. A DNA aptamer with the PCV2 Rep recognition sequence was enzymatically conjugated to the fusion protein via Rep.

Herein, we tested this idea by constructing a fusion protein in which pRep was positioned between Venus as an acceptor fluorophore and NanoLuc as an energy donor. Overlap of the bioluminescent emission spectrum with the excitation spectrum of the acceptor fluorophore is important for selection of a BRET molecule pair. The combination of Venus and NanoLuc has been previously applied in several BRET-based sensors [15,16]. Energy transfer occurs when the donor and acceptor molecules are within 10 nm of each other in proximity [14,17]. Based on the known structure of pRep, we expected that BRET between NanoLuc and Venus would occur in the designed fusion protein. As noted above, the BRET efficiency depends not only on the distance between the donor and acceptor molecules, but also their relative orientations. To overcome limitations associated with molecular orientation, linkers are generally inserted between the sensory domain and the BRET molecules to allow for a certain degree of movement of the BRET molecules [14]. pRep exhibits flexible regions at the N- and C-terminals [18]. In the present study, we expected the orientation change upon binding of target molecules to sensor. To allow movement of the BRET molecules only when the target molecules bind to the DNA aptamer conjugated to the fusion protein, we constructed a truncated pRep variant from the catalytic domain of PCV2 Rep comprising residues 14–109 (tpRep). By using tpRep, the BRET efficiency is expected to change through alteration of the orientation upon binding of the target molecule to the DNA aptamer. Herein, truncated pReps were expressed in *Escherichia coli* and subsequently purified, and the DNA binding abilities were evaluated. Finally, tpRep was used to construct DNA-protein conjugates with thrombin- and lysozyme-binding DNA aptamers, and the resultant DNA-protein conjugates were evaluated for use as a BRET-based biosensor.

## 2. Materials and Methods

### 2.1. Construction of Plasmids

For expression of Rep mutants, the plasmids pET-His-pRep<sub>1-109</sub>, pET-His-pRep<sub>14-116</sub> and pET-His-tpRep, were constructed as follows. pReps with the flexible N- and/or C-terminus truncated were constructed by site-directed mutagenesis. First, truncated pRep without the flexible 7-amino acid C-terminus, namely pRep<sub>1-109</sub>, was constructed by site-directed mutagenesis using the

primer set 5'-TGGAGCTGTCGACTAAGGTACCCTC-3' and 5'-TAGTCGACAGCTCCACACTCCATTA-3'. pET-His-pRep, which was previously constructed in our laboratory, was used as a template [5]. The resultant plasmid was named pET-His-pRep<sub>1-109</sub>. Truncated pRep without the flexible 13-amino acid N-terminus, namely pRep<sub>14-116</sub>, and tpRep were constructed by the same procedure using the primer set 5'-CGAATTCCACAAACGTTGGGTCTTC-3' and 5'-ACGTTTGTGGAATTCGCCCATGGCATG-3'. pET-His-pRep and pET-His-pRep<sub>1-109</sub> were used as templates, respectively. The resultant plasmids were named pET-His-pRep<sub>14-116</sub> and pET-His-tpRep, respectively.

The pET28-VenusΔ10-pRep-NanoLuc-His plasmid for expression of the fusion protein VenusΔ10-pRep-NanoLuc-His (VpRNH), in which the pRep protein was fused to the N-terminus of NanoLuc and the C-terminus of VenusΔ10, was constructed as follows. The VenusΔ10 fragment was amplified by PCR using the primer set 5'-tacgaattcagtaaaggagaagaacttttc-3' and 5'-ggcgaattcggtaacccccagcagctgttac-3', and Nano-lantern(cAMP-1.6)/pRSETB (Addgene #53591) was used as a template. The amplified fragment was cloned into pUC18 and its sequence was confirmed. The resultant plasmid, pUC18-VenusΔ10, was digested with *EcoR* I. The obtained fragment was inserted into the pET28-pRep-NanoLuc-His plasmid, which was previously constructed in our laboratory, and digested with the same restriction enzyme.

The pET28-VenusΔ10-tpRep-NanoLuc-His plasmid for expression of the fusion protein VenusΔ10-tpRep-NanoLuc-His (VtpRNH), in which the tpRep protein was fused to the N-terminus of NanoLuc and the C-terminus of VenusΔ10, was constructed as follows. To destroy an *EcoR* I restriction site located at the beginning of the VenusΔ10 sequence, a single nucleotide base of pET28-VenusΔ10-pRep-NanoLuc-His was changed by site-directed mutagenesis using the primer set 5'-GGCGATTTATGAGTAAAGGAGAAGAA-3' and 5'-ACTCATAAATTCGCCCATGGTATATCT-3'. The resultant plasmid, pET28-delEcoRI-VenusΔ10-pRep-NanoLuc-His, was digested with *EcoR* I and *Sal* I to remove the fragment encoding pRep and the inserted tpRep fragment derived from pET-His-tpRep was digested with the same restriction enzymes. The resultant plasmid was named pET28-VenusΔ10-tpRep-NanoLuc-His.

## 2.2. Protein Expression and Purification

For protein expression, the respective plasmids were introduced into *E. coli* BL21(DE3)-competent cells. Transformed cells were inoculated in LB medium with 50 µg/mL ampicillin for expression of Rep mutants and with 20 µg/mL kanamycin for expression of the fusion proteins VpRNH and VtpRNH. Cells were then cultured at 37 °C until the OD<sub>660</sub> reached 0.6~0.8, followed by addition of 1 mM isopropyl-β-D(-)-thiogalactopyranoside (IPTG) for induction of protein expression. Cells were cultured overnight at 16 °C for expression of fusion proteins and for 6 h at 25 °C for expression of Rep mutants. Cells were then harvested by centrifugation. The collected cells were suspended in phosphate-buffered saline (PBS: 150 mM NaCl, 16 mM Na<sub>2</sub>HPO<sub>4</sub>, 4 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4) and disrupted by sonication, followed by centrifugation to obtain the soluble fraction. The supernatant was added to Profinity<sup>TM</sup> IMAC Ni-Charged Resin (Bio-Rad) equilibrated with PBS followed by rotation at 4 °C for 30 min. After rotation, the samples were washed with Ni-NTA buffer (0.5 M NaCl, 20 mM phosphate buffer, pH 8.0) 5 times, and then washed twice with Ni-NTA buffer containing 10 mM imidazole. Proteins were eluted with Ni-NTA buffer containing 10, 100, 150 mM imidazole. The eluted samples were dialyzed against PBS 3 times using dialysis tubing. The concentrations of purified proteins were evaluated using a BCA assay kit (Pierce).

## 2.3. Evaluation of DNA Binding Ability of pRep

The ssDNA oligonucleotide Rep sub31 (5'-AAGTATTACAAAAACCAGCGCAGTTGGGCAG-3') was used for evaluation of the DNA binding ability of pRep. The underlined sequence denotes the PCV2 Rep recognition sequence. Purified proteins (5 µM) were mixed with Rep sub31 (5 µM) in 10 µL of reaction buffer, PBS with 2.5 mM MgCl<sub>2</sub>. After incubation for 30 min at 37 °C, the samples were run on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and the gels were stained with Coomassie Brilliant Blue (CBB).

#### 2.4. Evaluation of Emission Spectra

BRET of purified fusion proteins were evaluated using the Nano-Glo Luciferase Assay System (Promega). Fusion proteins (1  $\mu$ M) in 50  $\mu$ L of PBS were mixed with the same volume of Nano-Glo Luciferase Assay Reagent. Emission spectra (350–650 nm) was acquired 10 s after reagent addition using a FP6500 spectrofluorophotometer (JASCO, Tokyo, Japan). Emission spectra were normalized to the peak emission of NanoLuc at 450 nm, which was set to an intensity of 1.00.

#### 2.5. Homogeneous Assay with DNA Aptamer

The single-stranded DNA aptamers for lysozyme, namely Lysoapt 42 (5'-AAGTATTACATCTACGAATTCATCAGGGCTAAAGAGTGCAGAGTTACTTAG-3'), and for thrombin, namely TBA 29 (5'-AAGTATTACAGTCCGTGGTAGGGCAGGTTGGGGTGGACT-3'), containing the Rep recognition sequence, were used in the homogeneous assay [19,20]. Human  $\alpha$ -thrombin was purchased from Haematologic Technologies. Equal amounts (1  $\mu$ M each) of VtpRNH and DNA aptamer were mixed in 50  $\mu$ L of PBS containing 0.5 mM of  $MgCl_2$  followed by incubation at 37  $^{\circ}$ C. After incubation for 30 min, 10  $\mu$ L of solutions containing varying concentrations of target molecules was added to the mixture of VtpRNH and DNA aptamer followed by incubation at room temperature for 1 h. For the thrombin-binding DNA aptamer-VtpRNH conjugate, samples containing bovine serum albumin (BSA) (instead of thrombin) were evaluated as a negative control. Samples were mixed with the same volume of Nano-Glo Luciferase Assay Reagent. Emission spectra (350–650 nm) were acquired 10 sec after addition of the reagent using a FP6500 spectrofluorophotometer. Emission spectra were normalized to the peak emission of NanoLuc at 450 nm, which was set to an intensity of 1.00. The change in BRET ratio ( $\Delta$ BRET) was calculated as follows:

$$\Delta\text{BRET} = \left\{ \frac{\text{emission}_{\text{at } 528 \text{ nm}}}{\text{emission}_{\text{at } 450 \text{ nm}}} \right\} - \text{BRET basal ratio}$$

The BRET basal ratio is defined as  $\left( \frac{\text{emission}_{\text{at } 528 \text{ nm}}}{\text{emission}_{\text{at } 450 \text{ nm}}} \right)$  at 0  $\mu$ M of the target molecule.

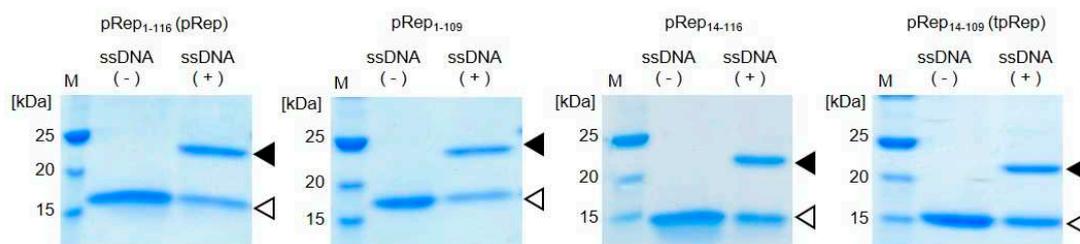
### 3. Results and Discussion

#### 3.1. Truncation of the Catalytic Domain of PCV2 Rep

We previously developed a method for site-specific conjugation of ssDNA to a protein of interest via the fused replication initiator protein (Rep), such as conjugation of Gene A\* from bacteriophage phiX 174 and the catalytic domain of porcine circovirus type 2 Rep comprising residues 1–116 (pRep) [5,13]. Recently, Gordon's group reported a similar strategy for construction of DNA-protein conjugates using several Reps, including pRep, as HUH-tags [21]. pRep is much smaller than Gene A\* and well-expressed in *E. coli*. The structure of pRep (Protein Data Bank (PDB) ID:2HW0) exhibits flexible regions at the N- and C-terminals [18]. In our designed fusion protein, Rep was positioned between Venus and NanoLuc. For construction of BRET-based biosensors, flexible linkers are generally inserted between the sensory domain and the BRET molecule to allow for a certain degree of movement of the BRET molecule [14]. In the present study, to allow for movement of BRET molecules only when the target molecule binds to the DNA aptamer conjugated to the fusion protein, use of truncated variants of pRep was required. Truncated variants of pReps were constructed and their DNA binding abilities were evaluated.

Truncated variants of pRep, namely pRep<sub>1–109</sub> lacking the last 7 amino acids in the sequence, pRep<sub>14–116</sub> lacking the first 13 amino acids in the sequence, and tpRep (pRep<sub>14–109</sub>) lacking the flexible regions at both the N- and C-terminals of pRep, were expressed in *E. coli* and subsequently purified. Truncated pReps were purified from the soluble fraction, similar to pRep. Proteins were mixed with ssDNA containing a Rep recognition sequence for evaluation of the DNA binding abilities of the truncated pReps. After incubation, samples were analyzed by SDS-PAGE (Figure 2). In addition to the bands appearing at the expected sizes of the protein, all samples also showed other bands at

higher molecular weights than pReps without conjugation. Those bands appearing at higher molecular weights were attributed to pReps conjugated with ssDNA. Upon cleavage of specific sequence of ssDNA, PCV2 Rep is conjugate to DNA covalently. In addition of this reaction, PCV2 Rep also catalyze the reaction in the opposite direction [18]. PCV2 Rep has an ability to catalyze joining two ssDNA fragments, a free 3'-OH and the 5'-phosphate covalently linked to Rep, for regeneration of PCV2 Rep recognition sequence. Therefore, there is a possibility to exist unreacted protein. However, this joining activity is not efficient compared to cleavage activity followed by conjugation of DNA with Rep. In this experiment, the ratio of protein and ssDNA is 1:1. By increasing ssDNA concentration, the conjugation efficiencies should increase. These results demonstrate that tpRep retains the ability to covalently bind to ssDNA even after truncation of the flexible regions located at N- and C-terminals of pRep. Therefore, tpRep was used for construction of a BRET-based biosensor with a DNA-protein conjugate.

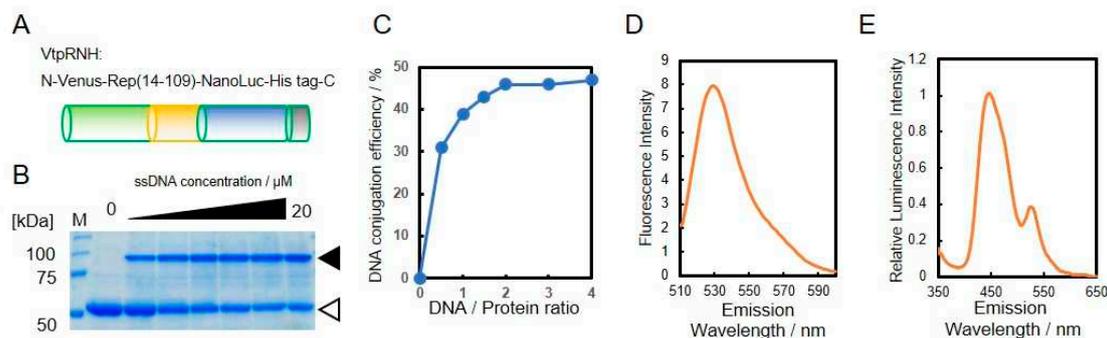


**Figure 2.** DNA binding abilities of truncated pRep proteins. The flexible region of the catalytic domain of PCV2 Rep comprising amino acids 1-116 was truncated from the full protein (pRep<sub>1-116</sub>, abbreviated pRep; MW 15,100); pRep<sub>1-109</sub>: pRep lacking the last 7 amino acids in the sequence (MW 14,300); pRep<sub>14-116</sub>: pRep lacking the first 13 amino acids in the sequence (MW 13,700); and pRep<sub>14-109</sub>, abbreviated tpRep: pRep lacking the flexible regions at both the N- and C-terminals of pRep (MW 12,900); white triangles indicate pRep without DNA and black triangles indicate pRep conjugated with DNA.

### 3.2. Construction of DNA-Protein Conjugates with NanoLuc and Venus for BRET-Based Biosensor

As shown in Figure 1, we designed a BRET-based sensor with DNA-protein conjugates. For construction of the DNA-protein conjugates, the catalytic domain of PCV2 Rep comprising residues 14–109 (tpRep) was fused to the N-terminus of NanoLuc and the C-terminus of VenusΔ10 with a His-tag (Figure 3A). The resultant fusion protein, VtpRNH, was expressed in *E. coli* and subsequently purified from the soluble fraction with a His-tag located at the C-terminus. After purification, the activities of each domain of VtpRNH were evaluated. First, the DNA binding ability of tpRep in the fusion protein was evaluated (Figure 3B). Even after fusion with NanoLuc and VenusΔ10, tpRe retained its DNA binding ability. The conjugation efficiencies were found increase with increasing DNA concentration. The maximum efficiency was around 50% (Figure 3C). We also evaluated the functions of both Venus and NanoLuc. Venus was shown to fluorescence in the fusion protein (Figure 3D). The emission peak was ~530 nm at an excitation wavelength of 500 nm [22].

Bioluminescence of NanoLuc in the fusion protein was also evaluated. The emission spectrum of the fusion protein exhibited a peak at ~450 nm without external excitation, which corresponds to the emission of NanoLuc [23]. These results suggest that VtpRNH retained its NanoLuc activity. Moreover, in addition to the emission peak around 450 nm, there was a small emission peak at ~530 nm. To identify the origin of this peak, the emission spectra were normalized to the peak emission of NanoLuc at 450 nm, which was set to an intensity of 1.00. As shown in Figure 3E, the emission peak of normalized intensity was observed at ~530 nm. The emission peak of Venus occurs at ~530 nm. These results suggest that BRET occurred between NanoLuc and Venus. We measured the emission spectrum after ssDNA binding via Rep to the fusion protein. Even after ssDNA binding to the fusion proteins via Rep, the emission peak at ~530 nm was observed (data not shown).

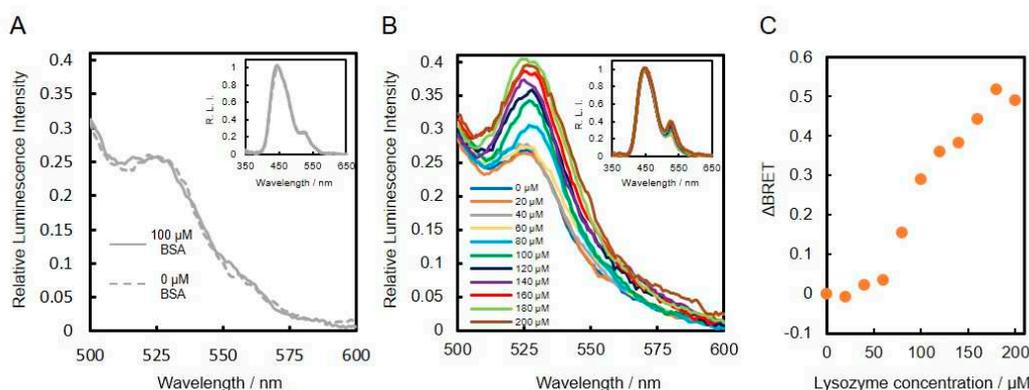


**Figure 3.** Evaluation of DNA binding ability of the fusion protein. Design of the constructed fusion protein, VtpRNH (MW 57,900) (A). DNA binding of tpRep in the fusion protein; the concentration of VtpRNH was 5  $\mu$ M; white triangle indicates pRep without DNA conjugation and black triangle indicates pRep with DNA conjugation (B). DNA-conjugation efficiency of protein was evaluated by ImageJ (C). Emission spectrum of Venus at an excitation wavelength of 500 nm; the concentration of VtpRNH was 3  $\mu$ M (D). NanoLuc activities of fusion proteins. The emission spectrum was normalized to the peak emission of NanoLuc at 450 nm; final concentration of VtpRNH was 0.5  $\mu$ M (E).

### 3.3. Construction and Evaluation of BRET-Based Biosensor with DNA-Protein Conjugates

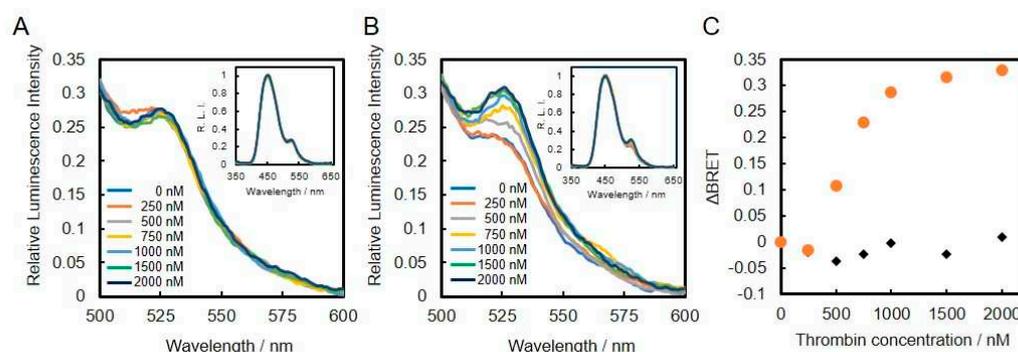
DNA aptamers are well known molecular binders with specificity to a specific molecule. In the present study, thrombin- and lysozyme-binding DNA aptamers were applied for the construction of a BRET-based biosensor with DNA-protein conjugates. For conjugation of DNA aptamer to the fusion protein via Rep, ssDNA comprised of the DNA aptamer and a Rep recognition sequence were reacted with the fusion protein.

Lysozyme-binding DNA aptamer was first conjugated to VtpRNH. After conjugation, the emission spectra of lysozyme binding DNA aptamer–protein conjugates with or without lysozyme were evaluated. As shown in insets of Figure 4A,B, the emission spectra were normalized to the peak emission of NanoLuc at 450 nm, which was set to an intensity of 1.00. In the presence of BSA (100  $\mu$ M), the normalized emission spectra of both DNA-protein conjugates with or without BSA were similar (Figure 4A). On the other hand, presence of lysozyme, the normalized emission peaks at 528 nm were found to increase compared to the control without lysozyme (Figure 4B). These results suggest that BRET efficiencies change upon binding of lysozyme to the DNA aptamers. To confirm this speculation, we evaluated the emission spectra in the presence of different concentrations of lysozyme. As shown in Figure 4C, the changes in BRET ratio ( $\Delta$ BRET) increased with increasing lysozyme concentration. These results confirm that the lysozyme-binding DNA aptamer-VtpRNH conjugate can be used as a BRET-based sensor. The  $K_d$  of the lysozyme-binding DNA aptamer is reported to be 30 nM [19]. Compared to this value, the sensitivity of our BRET-based biosensor with the lysozyme-binding DNA aptamer-VtpRNH conjugate is very low. As shown in Figure 3B, there were unreacted DNA. The sensitivity of sensor should be reduced by those unreacted DNA because unreacted DNA bind to target molecule competitively. However, in this case, it is not main reason for the low sensitivity. It was supposed that lysozyme formed aggregate in the presence of unreacted DNA caused by their net charge. Though monomeric lysozyme could not induce conformational change of fusion protein, BRET ratios would be changed by the aggregated lysozyme. The lysozyme binding DNA aptamer-VpRNH conjugate also showed similar results (Figure S1). We expected  $\Delta$ BRET to be larger for the lysozyme-binding DNA aptamer-VtpRNH conjugate compared with the DNA-VpRNH conjugate because tpRep was truncated to remove the flexible amino acid sequences from the N- and C-terminals. In contrast, the change in BRET ratio for the DNA-VtpRNH conjugate was smaller than that for the DNA-VpRNH conjugate. In case of DNA-VpRNH with high concentration of lysozyme, it was speculated that BRET ratio was changed not only by changes of the orientation but also the distance.



**Figure 4.** Evaluation of lysozyme-binding DNA aptamer–protein conjugates with and without target molecules. Normalized emission spectra of lysozyme-binding DNA aptamer-VtpRNH conjugate in the presence of BSA (A) and lysozyme (B). Change in BRET ratio ( $\Delta$ BRET) as a function of lysozyme concentration (C).

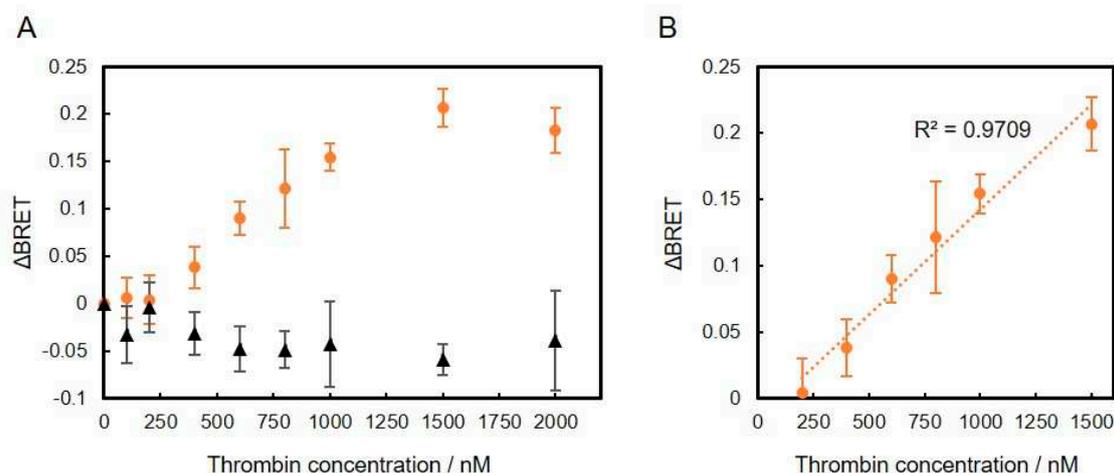
We also evaluated the thrombin-binding DNA aptamer conjugated to VtpRNH for use as a BRET-based biosensor. As shown in insets of Figure 5A,B, the normalized emission spectra of thrombin-binding DNA aptamer- and lysozyme-binding DNA aptamer-VtpRNH conjugates were evaluated in the presence of different thrombin concentrations. As shown in Figure 5A, the emission spectra of the lysozyme-binding DNA aptamer-VtpRNH conjugate were unchanged in the presence of different thrombin concentrations. On the other hand, the emission spectra of thrombin-binding DNA aptamer-VtpRNH conjugate increased with increasing thrombin concentration (Figure 5B). These results were also confirmed by calculation of  $\Delta$ BRET for each of the conjugates (Figure 5C). These results suggest that the BRET ratio changes only occurred with the appropriate combination of DNA aptamer and target molecules.



**Figure 5.** Evaluation of DNA aptamer-VtpRNH conjugate in the presence of different concentrations of thrombin. Conjugation with lysozyme-binding DNA aptamer (A) or thrombin-binding DNA aptamer (B). Change in BRET ratio ( $\Delta$ BRET) as a function of thrombin concentration (C). Conjugation with lysozyme-binding DNA aptamer (black diamonds) or thrombin-binding DNA aptamer (orange circles).

Finally, to demonstrate specificity of the conjugates, we evaluated the BRET ratio of the thrombin-binding DNA aptamer-VtpRNH conjugate in the presence of more different concentrations of thrombin or BSA. As shown in Figure 6,  $\Delta$ BRET of the thrombin-binding DNA aptamer-VtpRNH conjugate increased with increasing thrombin concentration. On the other hand,  $\Delta$ BRET did not increase in the presence of different BSA concentrations. It is suggested that the  $\Delta$ BRET depends on the specific binding between DNA aptamers and target molecules. These results further suggest that the thrombin-binding DNA aptamer-VtpRNH conjugate can be used as a BRET-based thrombin biosensor. The  $K_d$  of the thrombin-binding DNA aptamer is reported to be  $\sim 0.5$  nM, which is different from the value obtained from our BRET-based biosensor with the thrombin-binding DNA aptamer-VtpRNH

conjugate [20]. One of the reasons of that is existence of unconjugated DNA. In this experiment, same amount of DNA aptamer and VtpRNH were mixed. From the result of Figure 3C, around half of DNA aptamer would remain as unreacted DNA. It should decrease the sensitivity of this sensor. To evaluate the property of BRET sensors, separations of DNA aptamer-VtpRNH conjugates are required. In addition, steric effects would reduce the sensitivity of the sensor. Steric effects are required to induce conformational changes. However, steric effects may also reduce the binding ability of the DNA aptamer. Therefore, the  $K_d$  of DNA aptamers should increase in our BRET-based biosensor. The square of the correlation coefficient ( $R^2$ ) of the linear fit equation between concentrations of 200 and 1500 nM is close to 1 ( $R^2 = 0.97$ ). The square of the correlation coefficient was much closer to 1 compared to the results for the thrombin-binding DNA aptamer-VpRNH conjugate ( $R^2 = 0.78$ ) (Figure S2). These differences may be caused by the existence of flexible regions with the pRep protein, which could allow for movement of the BRET molecules even after binding of the target molecule to the DNA aptamer. These results suggest that DNA aptamer-VtpRNH conjugates exhibit the potential for use as BRET-based biosensors. While the sensitivity of these biosensors was too low compared with other DNA aptamer sensors, the sensitivity might be improved by changing the design of the fusion protein and DNA aptamers.



**Figure 6.** Change in BRET ratio ( $\Delta$ BRET) of thrombin-binding DNA aptamer-VtpRNH conjugate as a function of thrombin concentration in the presence of thrombin (orange circles) and BSA (black triangles) (A). Correlation of  $\Delta$ BRET and thrombin concentration between 200 and 1500 nM (B); each value represents the mean of 4 replicates ( $n = 4$ ).

#### 4. Conclusions

We constructed DNA aptamer–protein hybrid molecules for use as BRET-based biosensors. Fusion proteins were enzymatically conjugated to DNA aptamers via the catalytic domain of porcine circovirus type 2 replication initiation protein, which was positioned between NanoLuc luciferase and Venus. The catalytic domain of PCV2 Rep was shown to retain its DNA binding activity, even after elimination of the flexible regions at the N- and C-terminals. The resultant DNA aptamer–protein hybrid molecule exhibited a weak BRET signal, even in the absence of the target molecule of the DNA aptamer. However, the BRET signal was found to depend on the concentration of the target molecule. These results demonstrate the potential for use of the designed fusion protein as a DNA aptamer-based platform for construction of BRET-based biosensors.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/21/7646/s1>, Figure S1. Evaluation of lysozyme-binding DNA aptamer-VpRNH conjugates with and without target molecules. Normalized emission spectra of lysozyme-binding DNA aptamer-VpRNH conjugate in the presence of BSA (A) and lysozyme (B). Change in BRET ratio ( $\Delta$ BRET) as a function of lysozyme concentration (C), Figure S2.  $\Delta$ BRET of thrombin of thrombin-binding DNA aptamer-VpRNH conjugate as a function of thrombin concentration (A).

Correlation of  $\Delta$ BRET and thrombin concentration between 200 and 1500 nM (B); each value represents the mean of 4 replicates ( $n = 4$ ).

**Author Contributions:** Conceptualization, M.M.; methodology, M.M., R.H. and Y.M.; investigation, R.H.; writing—original draft preparation, M.M.; writing—review and editing, Y.M. and E.K.; supervision, M.M. and E.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by JSPS KAKENHI Grant Numbers 16K01388 (M.M.) and 26289310 (E.K.).

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

## References

1. Zhao, D.; Kong, Y.; Zhao, S.; Xing, H. Engineering Functional DNA–Protein Conjugates for Biosensing, Biomedical, and Nanoassembly Applications. *Top. Curr. Chem.* **2020**, *378*, 41. [[CrossRef](#)]
2. Sano, T.; Smith, C.; Cantor, C. Immuno-PCR: Very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science* **1992**, *258*, 120–122. [[CrossRef](#)] [[PubMed](#)]
3. Shimada, J.; Maruyama, T.; Kitaoka, M.; Kamiya, N.; Goto, M. DNA–enzyme conjugate with a weak inhibitor that can specifically detect thrombin in a homogeneous medium. *Anal. Biochem.* **2011**, *414*, 103–108. [[CrossRef](#)] [[PubMed](#)]
4. Takahara, M.; Wakabayashi, R.; Minamihata, K.; Goto, M.; Kamiya, N. Primary Amine-Clustered DNA Aptamer for DNA–Protein Conjugation Catalyzed by Microbial Transglutaminase. *Bioconjug. Chem.* **2017**, *28*, 2954–2961. [[CrossRef](#)]
5. Mie, M.; Niimi, T.; Mashimo, Y.; Kobatake, E. Construction of DNA-NanoLuc luciferase conjugates for DNA aptamer-based sandwich assay using Rep protein. *Biotechnol. Lett.* **2019**, *41*, 357–362. [[CrossRef](#)]
6. Wang, G.; Wang, Y.; Chen, L.; Choo, J. Nanomaterial-assisted aptamers for optical sensing. *Biosens. Bioelectron.* **2010**, *25*, 1859–1868. [[CrossRef](#)]
7. Sassolas, A.; Blum, L.J.; Leca-Bouvier, B.D. Homogeneous assays using aptamers. *Analyst* **2011**, *136*, 257–274. [[CrossRef](#)]
8. Citartan, M.; Gopinath, S.C.B.; Tominaga, J.; Tan, S.-C.; Tang, T.-H. Assays for aptamer-based platforms. *Biosens. Bioelectron.* **2012**, *34*, 1–11. [[CrossRef](#)]
9. Gokulrangan, G.; Unruh, J.R.; Holub, D.F.; Ingram, B.; Johnson, C.K.; Wilson, G.S. DNA Aptamer-Based Bioanalysis of IgE by Fluorescence Anisotropy. *Anal. Chem.* **2005**, *77*, 1963–1970. [[CrossRef](#)]
10. Rupcich, N.; Chiuman, W.; Nutiu, R.; Mei, S.; Flora, K.K.; Li, Y.; Brennan, J.D. Quenching of Fluorophore-Labeled DNA Oligonucleotides by Divalent Metal Ions: Implications for Selection, Design, and Applications of Signaling Aptamers and Signaling Deoxyribozymes. *J. Am. Chem. Soc.* **2006**, *128*, 780–790. [[CrossRef](#)] [[PubMed](#)]
11. Urata, H.; Nomura, K.; Wada, S.; Akagi, M. Fluorescent-labeled single-strand ATP aptamer DNA: Chemo- and enantio-selectivity in sensing adenosine. *Biochem. Biophys. Res. Commun.* **2007**, *360*, 459–463. [[CrossRef](#)]
12. Hao, L.; Zhao, Q. A fluorescein labeled aptamer switch for thrombin with fluorescence decrease response. *Anal. Methods* **2015**, *7*, 3888–3892. [[CrossRef](#)]
13. Mashimo, Y.; Maeda, H.; Mie, M.; Kobatake, E. Construction of Semisynthetic DNA-Protein Conjugates with Phi X174 Gene-A\* Protein. *Bioconjug. Chem.* **2012**, *23*, 1349–1355. [[CrossRef](#)] [[PubMed](#)]
14. Brown, N.E.; Blumer, J.B.; Hepler, J.R. Bioluminescence Resonance Energy Transfer to Detect Protein-Protein Interactions in Live Cells. In *Protein-Protein Interactions*; Meyerkord, C.L., Fu, H., Eds.; Methods in Molecular Biology; Springer: New York, NY, USA, 2015; Volume 1278, pp. 457–465. ISBN 978-1-4939-2424-0.
15. Yang, J.; Cumberbatch, D.; Centanni, S.; Shi, S.; Winder, D.; Webb, D.; Johnson, C.H. Coupling optogenetic stimulation with NanoLuc-based luminescence (BRET) Ca<sup>++</sup> sensing. *Nat. Commun.* **2016**, *7*, 13268. [[CrossRef](#)]
16. Inagaki, S.; Tsutsui, H.; Suzuki, K.; Agetsuma, M.; Arai, Y.; Jinno, Y.; Bai, G.; Daniels, M.J.; Okamura, Y.; Matsuda, T.; et al. Genetically encoded bioluminescent voltage indicator for multi-purpose use in wide range of bioimaging. *Sci. Rep.* **2017**, *7*, 42398. [[CrossRef](#)]
17. Dale, N.C.; Johnstone, E.K.M.; White, C.W.; Pflieger, K.D.G. NanoBRET: The Bright Future of Proximity-Based Assays. *Front. Bioeng. Biotechnol.* **2019**, *7*, 56. [[CrossRef](#)]
18. Vega-Rocha, S.; Byeon, I.-J.L.; Gronenborn, B.; Gronenborn, A.M.; Campos-Olivas, R. Solution Structure, Divalent Metal and DNA Binding of the Endonuclease Domain from the Replication Initiation Protein from Porcine Circovirus 2. *J. Mol. Biol.* **2007**, *367*, 473–487. [[CrossRef](#)] [[PubMed](#)]

19. Kirby, R.; Cho, E.J.; Gehrke, B.; Bayer, T.; Park, Y.S.; Neikirk, D.P.; McDevitt, J.T.; Ellington, A.D. Aptamer-Based Sensor Arrays for the Detection and Quantitation of Proteins. *Anal. Chem.* **2004**, *76*, 4066–4075. [[CrossRef](#)] [[PubMed](#)]
20. Tasset, D.M.; Kubik, M.F.; Steiner, W. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J. Mol. Biol.* **1997**, *272*, 688–698. [[CrossRef](#)] [[PubMed](#)]
21. Lovendahl, K.N.; Hayward, A.N.; Gordon, W.R. Sequence-Directed Covalent Protein–DNA Linkages in a Single Step Using HUH-Tags. *J. Am. Chem. Soc.* **2017**, *139*, 7030–7035. [[CrossRef](#)] [[PubMed](#)]
22. Nagai, T.; Ibata, K.; Park, E.S.; Kubota, M.; Mikoshiba, K.; Miyawaki, A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **2002**, *20*, 87–90. [[CrossRef](#)] [[PubMed](#)]
23. Hall, M.P.; Unch, J.; Binkowski, B.F.; Valley, M.P.; Butler, B.L.; Wood, M.G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; et al. Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem. Biol.* **2012**, *7*, 1848–1857. [[CrossRef](#)] [[PubMed](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# RNA-Peptide Conjugation through an Efficient Covalent Bond Formation

Shun Nakano, Taiki Seko, Zhengxiao Zhang and Takashi Morii \* 

Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan;  
snaka@iae.kyoto-u.ac.jp (S.N.); seko.taiki.56a@st.kyoto-u.ac.jp (T.S.);  
zhang.zhengxiao.38m@st.kyoto-u.ac.jp (Z.Z.)

\* Correspondence: t-morii@iae.kyoto-u.ac.jp

Received: 12 October 2020; Accepted: 10 December 2020; Published: 14 December 2020



**Abstract:** Many methods for modification of an oligonucleotide with a peptide have been developed to apply for the therapeutic and diagnostic applications or for the assembly of nanostructure. We have developed a method for the construction of receptor-based fluorescent sensors and catalysts using the ribonucleopeptide (RNP) as a scaffold. Formation of a covalent linkage between the RNA and the peptide subunit of RNP improved its stability, thereby expanding the application of functional RNPs. A representative method was applied for the formation of Schiff base or dihydroxy-morpholino linkage between a dialdehyde group at the 3'-end of sugar-oxidized RNA and a hydrazide group introduced at the C-terminal of a peptide subunit through a flexible peptide linker. In this report, we investigated effects of the solution pH and contribution of the RNA and peptide subunits to the conjugation reaction by using RNA and peptide mutants. The reaction yield reached 90% at a wide range of solution pH with reaction within 3 h. The efficient reaction was mainly supported by the electrostatic interaction between the RNA subunit and the cationic peptide subunit of the RNP scaffold. Formation of the RNP complex was verified to efficiently promote the reaction for construction of the RNA-peptide conjugate.

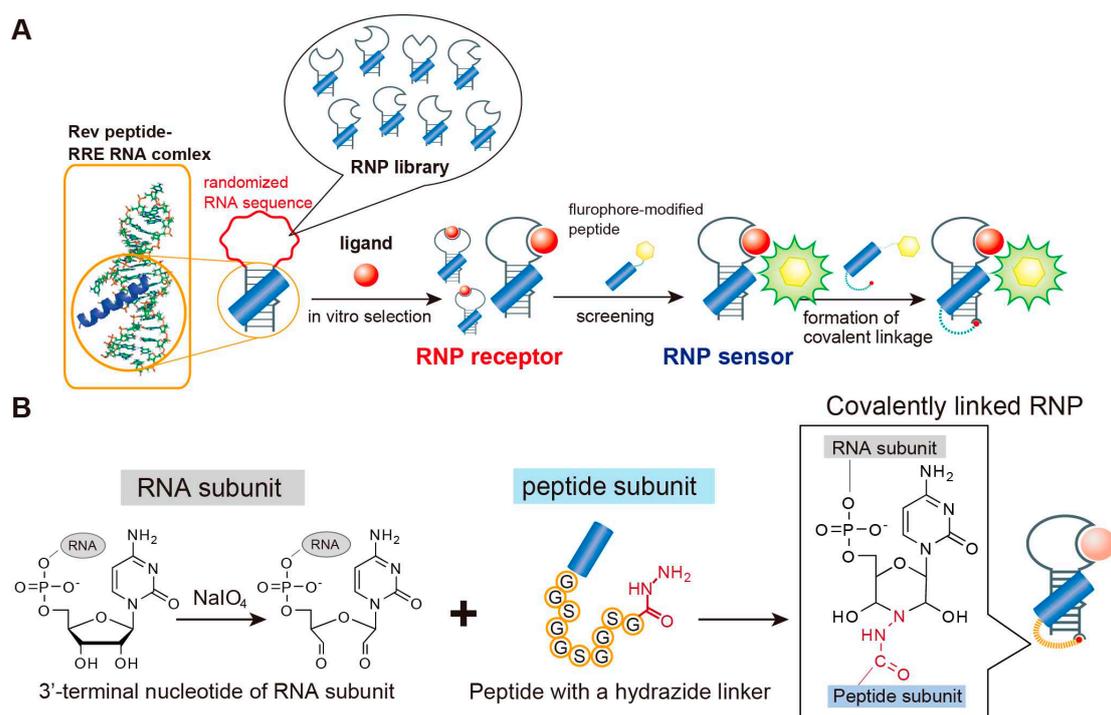
**Keywords:** ribonucleopeptide (RNP); RNA-peptide conjugate; Schiff base; aptamer; fluorescent sensors

## 1. Introduction

RNA and peptide conjugates have been constructed for versatile therapeutic application, such as the delivery of siRNA [1–3] or the screening of a library of peptides in a mRNA display method [4,5]. Efficient reaction for the formation of covalent linkage between RNA and peptide is important for developing practical applications. Many methods have been reported to conjugate an oligonucleotide (RNA or DNA) with an oligopeptide by applying various chemistries [6–9]. Post-synthetic coupling of the respective oligonucleotide and peptide fragments is the common and reliable method to construct oligonucleotide-peptide conjugates because the stepwise solid-phase synthesis of the conjugates encounters difficulties in finding the compatible protecting groups for both nucleobases and amino-acid side chains. Several chemical reactions were applied for coupling the nucleotide and peptide, such as amidation [10,11], disulfide bond formation [12–14], the modification of the thiol group by haloacetyl [15,16], or the maleimide group [12,17,18], native chemical ligation [19–21], and click reaction [22,23]. These reactions are certainly useful for the selective coupling of oligonucleotides and oligopeptides. However, these reactions often associate with drawbacks, such as the instability of the linkage, the necessity for introducing additional chemical groups into both the oligonucleotide and the oligopeptide chains by multistep reactions, limitation for the sequence of oligopeptide due to the solubility or reactivity of the side chains, and/or the potential formation of the side products and the stereoisomers. In some cases, even applying the long reaction time, the yields of the product were

rather low. One of the alternative methods for the modification of nucleotide with functional molecules was based on incorporating the aldehyde group into the nucleotide. Oxidation of the sugar moiety by perchloric acid or periodate is one of the common methods for introducing the aldehyde group on DNA or RNA [24–27]. By coupling the aldehyde group with alkoxyamine, cysteine, and hydrazine, formation of the covalent linkage through oxime, thiazolidine, and hydrazone (or a morpholine-like structure), respectively, has been reported [27–29]. These reactions proceed in mild conditions in a relatively short time to provide high coupling yields in aqueous solution. While this strategy also displays some of the limitations, as mentioned above, it has the advantage of having no need for the introduction of unnatural nucleic acids into oligonucleotide and facile preparation of the reactive peptide.

We have developed a stepwise method for the construction of receptor-based fluorescent sensors by using ribonucleopeptide (RNP) as a scaffold (Figure 1A). In this method, a complex of Rev peptide and Rev Responsive Element (RRE) RNA [30] was utilized to construct the RNP library with a randomized RNA sequence in the RNA subunit. RNP receptors [31] for a target molecule were selected from the RNP library by applying the *in vitro* selection method [32,33]. The peptide subunit of the selected RNP receptors was further modified with a fluorophore to construct a fluorophore-modified RNP receptor (F-RNP) library [34–42]. By screening the F-RNP library, fluorescent RNP sensors showing measurable fluorescent intensity changes upon binding the substrate were selected. Many kinds of RNP receptors and sensors were constructed for various target molecules, such as ATP [34–38,41,42], GTP [34,35,41], dopamine [39], and a tetra peptide containing a phosphorylated tyrosine residue [40]. Furthermore, the RNA subunit and the peptide subunit of fluorescent RNP sensors were covalently linked to improve the chemical and thermal stability [41]. The covalently linked RNP (c-RNP) sensors were applicable for simultaneous detection of multiple target molecules in the solution. Time-course monitoring of the concentration changes of the substrate and product in an enzymatic reaction was demonstrated by simultaneous application of sensors for the substrate and the product [41,42]. As mentioned above, for the covalent bond formation with the aldehyde group, the ribose moiety at the 3'-end of RNA was oxidized by sodium periodate to a dialdehyde group (Figure 1B). A hydrazide group was introduced at the C-terminal of peptide subunit through a ten-amino-acid flexible linker. Coupling of the RNA dialdehyde group and the C-terminal hydrazide group of peptide formed a covalent linkage between the RNA and peptide subunits of the Rev-RRE complex. This reaction proceeded rapidly in a quantitative yield. It is likely that a proximity effect between the reactive groups on the RNA and peptide subunit originated from the complex formation of RNP to assist the efficient coupling reaction. Here, we investigated several conditions for covalent linkage formation between the dialdehyde group of RNA and the hydrazide group of peptide within the ribonucleopeptide scaffold to elucidate the effect of solution pH for the reaction yield. Furthermore, the role of proximity effect in the coupling reaction was investigated by using mutants of RNA and peptide. Our results demonstrated that the efficient coupling reaction within the RNP was indeed the outcome of the proximity effect of reactive groups and provided an optimal condition for the construction of the covalently linked RNA-peptide complex.



**Figure 1.** Schematic illustration for the construction of the covalently linked fluorescent ribonucleopeptide (RNP) sensor. RNA receptors selected from the RNA-derived RNP library were converted into a fluorescent RNP sensor library through modification of the peptide subunit by a fluorophore. (A) A fluorescent RNP sensor, selected from the fluorescent RNP sensor library, was converted to a covalently linked RNP (c-RNP) sensor by the crosslinking reaction between RNA and the peptide subunit [41]. (B) Scheme of covalent linkage formation between RNA and the peptide subunit of RNP.

## 2. Materials and Methods

PrimeSTAR HS DNA polymerase for PCR reactions was obtained from TaKaRa Bio Inc. (Shiga, Japan), and the T7-Scribe Standard RNA IVT Kit was obtained from CELLSRIPT (Madison, WI, USA). *N*- $\alpha$ -Fmoc-protected amino acids, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 1-hydroxybenzotriazole (HOBT), distilled *N,N*-dimethylformamide (DMF), and 4-Hydroxymethylbenzoic acid-polyethylene glycol (PEG) resin (HMBA-PEG resin) were obtained from Watanabe Chemical Industries (Hiroshima, Japan). *N,N*-diisopropylethylamine (DIEA), diisopropylcarbodiimide (DIC), and 2',4',6'-trihydroxyacetophenone monohydrate (THAP) were obtained from Sigma-Aldrich (St. Louis, MO, USA). *N,N*-dimethyl-4-aminopyridine (DMAP), sodium periodate, hydrazine monohydrate, gel electrophoresis grade acrylamide, bisacrylamide, phenol, thioanisole, and 1,2-ethanedithiol were purchased from Wako Chemicals (Tokyo, Japan). Diammonium hydrogen citrate (DAHC) was obtained from Nacalai Tesque (Kyoto, Japan). A reversed-phase C18 column ULTRON VX-ODS (For analysis: 4.6  $\times$  150 mm; For purification: 20  $\times$  250 mm, particle size 5  $\mu$ m) was purchased from Shinwa Chemical Industries (Kyoto, Japan).

### 2.1. Preparation of RNA

Forward and Reverse primers for the construction of the double-stranded template DNA for An16 RNA (5'-TCTAATACGACTCACTATAGGTCTGGGCGCA-3' and 5'-GGCCTGTACCGTC-3') and for scrAn16 RNA (5'-TCTAATACGACTCACTATAGGAGGCTTCAGCTTCG-3' and 5'-CACACAACCGCCCCG-3') were purchased from Sigma-Aldrich, Japan (T7 RNA promoter is underlined). The double-stranded DNA templates and RNAs were prepared as previously described [42]. RNA subunits of RNP receptors were purified by means of denaturing polyacrylamide gel

electrophoresis (8 M urea, 12%). The concentrations of purified RNA were quantified by measuring the absorption at 260 nm ( $An_{16} = 439,500$ ;  $scrAn_{16} = 438,800 \text{ M}^{-1} \text{ cm}^{-1}$ ).

## 2.2. Construction of a Covalently Linked RNP

The peptide subunit for the formation of a covalent linkage (Ac-TRQARRNRRRRWRERQR-GGSGGSGGSG-HZ) was synthesized as follows. A HMBA-PEG resin was placed in a dry flask, and a sufficient amount of DMF was added to soak the resin; this mixture was allowed to swell for 30 min. *N*- $\alpha$ -Fmoc-glycine (10 equivalent relative to resin loading) dissolved in DMF was mixed with a solution of DIC (5 equivalent relative to resin loading) in dry DMF on ice and then incubated for 20 min. The solution of an activated first amino acid was added to the resin prepared above. A DMF solution of DMAP (0.1 equivalent relative to resin loading) was added to the resin/amino acid mixture and incubated at room temperature for 1 h with occasional swirling. This procedure of coupling the first amino acid residue was repeated twice. The subsequent synthesis was performed on an automated peptide synthesizer (PSSM-8; Shimadzu, Kyoto, Japan) according to the Fmoc chemistry protocol using protected Fmoc-amino acids and HBTU. Acetylation of the *N*-terminal of peptide was performed by mixing with 1 M acetic acid anhydride and 1 M 1-methyl imidazole in DMF for 1 h at room temperature. The following procedures, cleavage of the protected peptide from the resin with hydrazine monohydrate, deprotection of the protected peptide, and purification of the peptide were performed as described previously [41]. The synthesized peptides were characterized by MALDI-TOF mass spectrometry (AXIMA-LNR, Shimadzu), as follows: Acetylated Rev peptide hydrazide (Ac-Rev-(GGG)<sub>3</sub>G-HZ),  $m/z$  3154.9 (calcd for  $[M+H]^+$  3155.5); Acetylated flexible peptide linker hydrazide (Ac-(GGG)<sub>3</sub>G-HZ),  $m/z$   $[M+H]^+$  735.3,  $[M+Na]^+$  757.8 (calculated for  $[M+H]^+$  734.3).

Crosslinking reaction of the RNA and the peptide subunits of RNP was carried out as described previously [41,42] with a slight modification of the conditions. Freshly prepared 0.01 M sodium periodate (5  $\mu\text{L}$ ; 50 nmol; 50 equiv) was added to 200  $\mu\text{M}$  RNA (5  $\mu\text{L}$ ; 1 nmol) in 15  $\mu\text{L}$  of 0.03 M sodium acetate (pH 5.2), and the reaction mixture (25  $\mu\text{L}$ ) was incubated for 1 h at 37 °C in the dark. After the reaction, 2.5  $\mu\text{L}$  of 10 M glycerol was added to the reaction mixture to reduce an excess amount of sodium periodate. The resulting oxidized RNA was purified by ethanol precipitation. A coupling reaction between the 3'-modified RNA (40  $\mu\text{M}$ ) and Ac-Rev-(GGG)<sub>3</sub>G-HZ (88  $\mu\text{M}$ ) was performed in 0.03 M sodium acetate (for pH 4 and 5) or 0.03 M sodium phosphate (for pH 6 and 7), containing 0.01 M NaCl (total 25  $\mu\text{L}$ ) at 37 °C in the dark. The reaction mixture was extracted by phenol/chloroform and purified by ethanol precipitation to remove unreacted peptide, then dissolved in TE (25  $\mu\text{L}$ ).

## 2.3. Evaluation of the Reaction Yields of Covalently-Linked RNPs

Denaturing polyacrylamide gel electrophoresis (PAGE) (8 M urea, 12%) was performed to separate the unreacted RNA and covalently linked RNP by loading the same volume of the purified samples. A total of 10 pmol of RNA was loaded on the gel as the control of RNA band mobility. The above stock solution of phenol/chloroform extracted RNP (25  $\mu\text{L}$ ) was further diluted with TE to 200  $\mu\text{L}$ . From this RNP solution, 2  $\mu\text{L}$  was analyzed by PAGE for each reaction. The acrylamide gel was stained with ethidium bromide to detect RNA and RNP. The yield of each reaction was evaluated from the ratio of intensity of bands corresponding to unreacted RNA and RNP. Actual isolation yields of RNP ranged from 30 to 40% after the PAGE purification and successive purification by ethanol precipitation.

## 2.4. MALDI-TOF Mass Analysis of the Reaction Solution

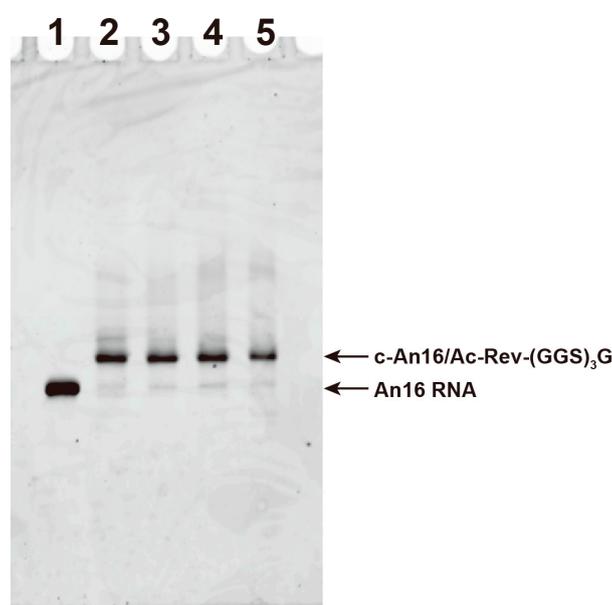
The reaction solution removed the unreacted peptide by phenol/chloroform extraction, and ethanol precipitation was characterized by MALDI-TOF mass spectrometry (AXIMA-Confidence, Shimadzu, Kyoto, Japan). A matrix solution was prepared as a mixture (*v/v*: 9/1) of 10 mg/mL of THAP in acetonitrile/H<sub>2</sub>O (*v/v*: 1/1) and 50 mg/mL of DAHC in pure water. The sample solution was mixed with the same volume of matrix solution (1  $\mu\text{L}$  each). DNA oligonucleotides (molecular weight: 5828.9,

11,397.5, 13,976.2, 23,164.3) were used as the external calibration standard sample. All measurements were performed in linear positive mode.

### 3. Results

#### 3.1. Effect of Solution pH for the Formation of Covalent Linkage

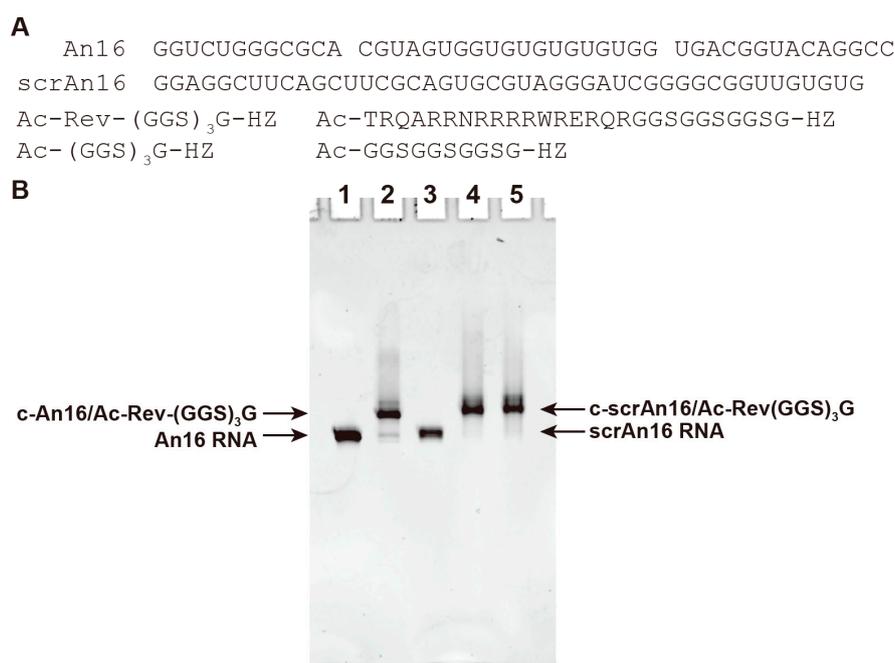
The effect of solution pH for the formation of covalent linkage between the 3'-dialdehyde group of the oxidized RNA subunit and the hydrazide group at the C-terminal of peptide subunit of the Rev-RRE complex was investigated by carrying out the reaction at various pH levels. The RNA subunit of ATP-binding of the RNP receptor, An16, was oxidized by using sodium periodate in a sodium acetate buffer at pH 5. The reaction was performed at 37 °C for 1 h. After isolation of the oxidized RNA by ethanol precipitation, acetylated Rev peptide, modified with a C-terminal hydrazide group through a flexible linker (GGS)<sub>3</sub>G (Ac-Rev-(GGS)<sub>3</sub>G-HZ), was mixed with the oxidized RNA in sodium acetate buffer at pH 4 or 5 or in a sodium phosphate buffer at pH 6 or 7. The reaction mixtures were incubated for 3 h at 37 °C, then the reactions were stopped by the addition of phenol/chloroform and vigorously mixed. The water layer was condensed by ethanol precipitation. The reaction yield of c-RNP in each solution was evaluated by quantitation of the band intensity, corresponding to free RNA and conjugated c-RNP in a denaturing polyacrylamide gel electrophoresis (PAGE) (Figure 2). Over 80% of RNA was reacted with peptide in all conditions. The yields of reaction were 92 ± 4% at pH 4, 92 ± 5% at pH 5, 86 ± 3% at pH 6, and 82 ± 3% at pH 7 (Figure S1A). These results indicated that the solution pH ranging from 4 to 7 did not affect much to the yield of RNA-peptide conjugate. Prolonged incubation over 15 h at pH 5 did not show any difference in the ratio of the band intensity for c-RNP and unreacted RNA (data not shown).



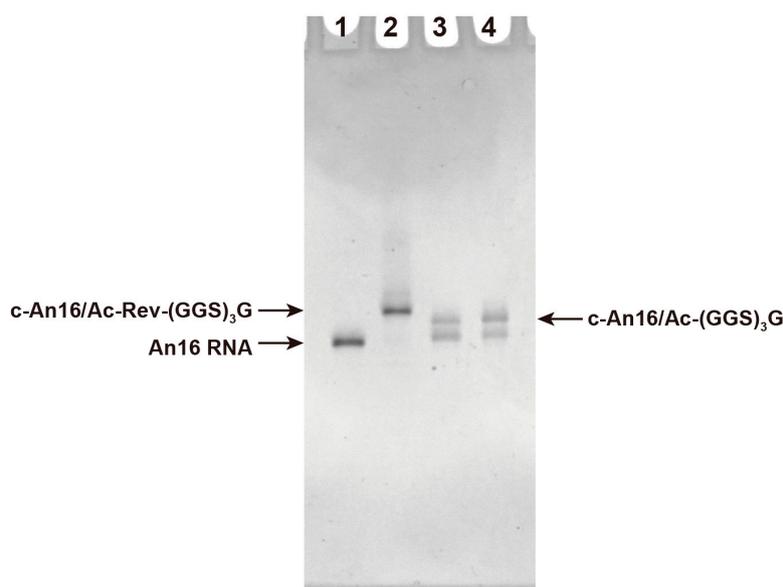
**Figure 2.** A denaturing polyacrylamide gel electrophoresis (PAGE) analysis (8 M Urea) of the conjugation reaction products of c-RNPs at different pH conditions. The yield was calculated from the ratio of intensities of the bands corresponding to c-RNP and RNA. Lane 1: An16 RNA only; Lane 2: An16 and Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted at pH 4; Lane 3: An16 and Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted at pH 5; Lane 4: An16 and Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted at pH 6; Lane 5: An16 and Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted at pH 7.

### 3.2. Contribution of the RNA-Peptide Interaction for the Reaction Yield

RNA sequence dependency of the reaction was next evaluated to address the proximity effect of reactive groups. The specific interaction between the Rev peptide and RRE RNA of the RNP scaffold was expected to have an influence on the efficiency of covalent linkage formation. A sequence-scrambled scrAn16 RNA that contained the same number of each nucleotide to An16 RNA was prepared (Figure 3A). The reaction of the oxidized scrAn16 with Ac-Rev-(GGS)<sub>3</sub>G-HZ was performed at pH 5. The MALDI TOF MS analysis using the reaction solution after removing the unreacted peptide was performed to characterize the products. The production of covalently linked RNP was confirmed by observing respective MS peaks (Figures S2 and S3). In the five trials, scrAn16 showed a similar averaged reaction yield ( $92 \pm 5\%$ ) to that of the parent An16 ( $92 \pm 5\%$ ) when conjugated with the hydrazide group of Ac-Rev-(GGS)<sub>3</sub>G-HZ (Figure 3B and Figure S1B,D). The result showed that the RNA sequence did not affect the efficiency of reaction. The non-specific electrostatic interaction between the RNA and the cationic Rev peptide would relate to the reaction efficiency. To verify this notion, the reaction for the formation of covalently linked RNP was carried out by using a truncated peptide of Ac-Rev-(GGS)<sub>3</sub>G-HZ. A truncated peptide possessing only the GGS peptide linker moiety, Ac-GSGSGSGSG-HZ, was designed by deleting the Rev peptide sequence from Ac-Rev-(GGS)<sub>3</sub>G-HZ. The truncated peptide showed a significant decrease in the yield to  $57 \pm 1\%$ , even after a prolonged reaction time for 14 h (Figure 4 and Figure S1C). This result supported the notion that the electrostatic interaction between RNA and the peptide mainly contributed to increasing the reaction yield. Thus, the RNP scaffold was useful for the efficient production of RNA-peptide conjugates.



**Figure 3.** (A) Nucleotide and peptide sequence of An16 RNA, scrAn16 RNA, Ac-Rev-(GGS)<sub>3</sub>G-HZ peptide, and Ac-(GGS)<sub>3</sub>G-HZ peptide; (B) a denaturing PAGE (8 M Urea) analysis of the conjugation reaction products of c-RNPs constructed from An16 or scrAn16 RNA with Ac-Rev-(GGS)<sub>3</sub>G-HZ. Lane1: An16 RNA; Lane 2: An16 and Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted for 3 h; Lane 3: scrAn16 RNA; Lane 4: scrAn16 and Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted for 3 h; Lane 5: scrAn16 and Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted for 14 h.



**Figure 4.** A denaturing PAGE (8 M Urea) analysis of the conjugation reaction products of c-RNP constructed from An16 and Ac-(GGS)<sub>3</sub>G-HZ. Lane 1: An16 RNA; Lane 2: c-An16/Ac-Rev-(GGS)<sub>3</sub>G-HZ reacted for 3 h; Lane 3: c-An16/Ac-(GGS)<sub>3</sub>G-HZ reacted for 3 h; Lane 4: c-An16/Ac-(GGS)<sub>3</sub>G-HZ reacted for 14 h.

#### 4. Discussion

A Schiff base formation between the hydrazide and aldehyde group rapidly progressed at a mild acidic condition although the stability was less than that at a neutral pH. Because the reaction efficiency of this crosslinking method was based on the balance of such properties, the assays under the wide range of pH helped us to understand a limitation and the applicability of this reaction under the physiological condition for RNA-protein or RNA-peptide conjugation. Formation of the covalent linkage between RNA and the Rev peptide subunit in the Rev-RRE complex quantitatively proceeded in a mild acidic to a neutral pH condition within 3 h. The result indicated the versatility of this reaction in a physiological condition for the crosslink formation of peptide and RNA. Unlike the sequence-specific nature for the stable noncovalent RNP complex formation, formation of the covalent linkage between the 3'-dialdehyde group of RNA and the hydrazide group at the C-terminal of Rev peptide required no specific RNA sequence for the RNA binding of Rev peptide. Even the nonspecific RNA-peptide complex formation driven by the electrostatic interaction was sufficient for exerting the proximity effect of reactive groups, thereby resulting an efficient formation of covalent linkage between the Rev peptide and RNA. The significant decrease of the reaction efficiency by deletion of the cationic moiety in the peptide sequence also showed the contribution of the electrostatic interaction for the efficient proceeding of the crosslinking reaction.

We designed a DNA sequence-specific protein-tag that underwent a proximity-driven intermolecular crosslinking between protein and DNA [43–49]. Selective DNA modification by a self-ligating protein tag conjugated with a DNA-binding domain, termed as a modular adaptor, was achieved by relying on the chemoselectivity of the protein tag [50–52]. By tuning the alkylation kinetics for the protein-tag and its substrate, the sequence-specific crosslinking reaction of the modular adaptor was exclusively driven by the DNA recognition when the dissociation rate of the DNA complex was much larger than the rate constant for the alkylation reaction [45–49]. In connection with these findings, a sequence-specific crosslinking reaction between the 3'-dialdehyde group of RNA and the hydrazide group of RNA-binding peptide would be realized by tuning the reaction conditions, such as the concentration of sodium salt. As we have reported previously, the crosslinking reaction between the 3'-dialdehyde group of RNA and the hydrazide group of peptide was applied for the facile construction

of covalently crosslinked fluorescent RNP sensors [41,42]. The covalently crosslinked fluorescent RNP sensor not only realized an improved stability of RNP sensors but also expand its application, such as the simultaneous detection of multiple targets in the solution. The noncovalent RNP scaffold was effective for the library-based selection and the cooperative functionalization of RNP receptor, such as the fluorescent sensors and catalysts [53]. Once a functional noncovalent RNP was obtained, formation of a covalent linkage within the RNA-peptide complex provided stable RNP for a wide range of applications. A natural RNA oligonucleotide and facile preparation of the reactive peptide was only needed to perform this reaction. Our investigations regarding the effect of the pH, RNA, and a peptide sequence for the reaction efficiencies in this reaction would help us to understand the limitation and applicability of this method for construction of RNA-peptide or protein conjugate. It would also be helpful for the development of an emerging application of oligonucleotide-peptide conjugates, such as a self-assembled scaffold of nanostructure or an efficient delivery system of functional RNA into the cell.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/24/8920/s1>, Figure S1: The results of band intensity analysis of denaturing PAGE, Figure S2: MALDI TOF MS analysis of the reaction solution, Figure S3: Summary of the results of MALDI TOF MS analysis.

**Author Contributions:** S.N. and T.M. conceived and designed the experiments; S.N., T.S. and Z.Z. performed the experiments; T.M. supervised the project. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by JSPS KAKENHI Grant Numbers 18K14335 (S.N.) and 17H01213 (T.M.), Japan and by JST CREST Grant Number JPMJCR18H5 (T.M.), Japan.

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

## References

1. Roberts, T.C.; Langer, R.; Wood, M.J.A. Advances in oligonucleotide drug delivery. *Nat. Rev. Drug Discov.* **2020**, *19*, 673–694. [CrossRef] [PubMed]
2. Tai, W. Current aspects of siRNA bioconjugate for in vitro and in vivo delivery. *Molecules* **2019**, *24*, 2211. [CrossRef] [PubMed]
3. Tai, W.; Gao, X. Functional peptides for siRNA delivery. *Adv. Drug Deliv. Rev.* **2017**, *110–111*, 157–168. [CrossRef] [PubMed]
4. Roberts, R.W.; Szostak, J.W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12297–12302. [CrossRef]
5. Nemoto, N.; Miyamoto-Sato, E.; Husimi, Y.; Yanagawa, H. In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **1997**, *414*, 405–408.
6. Tung, C.H.; Stein, S. Preparation and applications of peptide-oligonucleotide conjugates. *Bioconjug. Chem.* **2000**, *11*, 605–618. [CrossRef]
7. Venkatesan, N.; Kim, B.H. Peptide conjugates of oligonucleotides: synthesis and applications. *Chem. Rev.* **2006**, *106*, 3712–3761. [CrossRef]
8. Lu, K.; Duan, Q.-P.; Ma, L.; Zhao, D.-X. Chemical strategies for the synthesis of peptide-oligonucleotide conjugates. *Bioconjug. Chem.* **2010**, *21*, 187–202. [CrossRef]
9. Singh, Y.; Murat, P.; Defrancq, E. Recent developments in oligonucleotide conjugation. *Chem. Soc. Rev.* **2010**, *39*, 2054–2070. [CrossRef]
10. Haralambidis, J.; Duncan, L.; Angus, K.; Tregear, G.W. The synthesis of polyamide-oligonucleotide conjugate molecules. *Nucleic Acids Res.* **1990**, *18*, 493–499. [CrossRef]
11. Kachalova, A.; Zubin, E.; Stetsenko, D.; Gait, M.; Oretskaya, T. Oligonucleotides with 2'-O-carboxymethyl group: Synthesis and 2'-conjugation via amide bond formation on solid phase. *Org. Biomol. Chem.* **2004**, *2*, 2793–2797. [CrossRef] [PubMed]
12. Eritja, R.; Pons, A.; Escarceller, M.; Giralt, E.; Albericio, F. Synthesis of defined peptide-oligonucleotide hybrids containing a nuclear transport signal sequence. *Tetrahedron* **1991**, *47*, 4113–4120. [CrossRef]

13. Vivès, E.; Lebleu, B. Selective coupling of a highly basic peptide to an oligonucleotide. *Tetrahedron Lett.* **1997**, *38*, 1183–1186. [[CrossRef](#)]
14. Turner, J.J.; Arzumanov, A.A.; Gait, M.J. Synthesis, cellular uptake and HIV-1 Tat-dependent trans-activation inhibition activity of oligonucleotide analogues disulphide-conjugated to cell-penetrating peptides. *Nucleic Acids Res.* **2005**, *33*, 27–42. [[CrossRef](#)] [[PubMed](#)]
15. Zhu, T.; Tung, C.-H.; Breslauer, K.J.; Dickerhof, W.A.; Stein, S. Preparation and physical properties of conjugates of oligodeoxynucleotides with poly( $\delta$ )ornithine peptides. *Antisense Res. Dev.* **1993**, *3*, 349–356. [[CrossRef](#)] [[PubMed](#)]
16. Arar, K.; Aubertin, A.-M.; Roche, A.-C.; Monsigny, M.; Mayer, R. Synthesis and antiviral activity of peptide oligonucleotide conjugates prepared by using NR-(Bromoacetyl) peptides. *Bioconjug. Chem.* **1995**, *6*, 573–577. [[CrossRef](#)]
17. Ghosh, S.S.; Kao, P.M.; McCue, A.W.; Chappelle, H.L. Use of maleimide-thiol coupling chemistry for efficient syntheses of oligonucleotide-enzyme conjugate hybridization probes. *Bioconjug. Chem.* **1990**, *1*, 71–76. [[CrossRef](#)]
18. Ede, N.J.; Tregear, G.W.; Haralambidis, J. Routine preparation of thiol oligonucleotides: Application to the synthesis of oligonucleotide-peptide hybrids. *Bioconjug. Chem.* **1994**, *5*, 373–378. [[CrossRef](#)]
19. McPherson, M.; Wright, M.C.; Lohse, P.A. Synthesis of an RNA-peptide conjugate by orthogonal ligation. *Synlett* **1999**, *51*, 978–980. [[CrossRef](#)]
20. Stetsenko, D.A.; Gait, M.J. Efficient conjugation of peptides to oligonucleotides by “native ligation”. *J. Org. Chem.* **2000**, *65*, 4900–4908. [[CrossRef](#)]
21. Takeda, S.; Tsukiji, S.; Nagamune, T. A cysteine-appended deoxyuridine for the postsynthetic DNA modification using native chemical ligation. *Tetrahedron Lett.* **2005**, *46*, 2235–2238. [[CrossRef](#)]
22. Gogoi, K.; Mane, M.V.; Kunte, S.S.; Kumar, V.A. A versatile method for the preparation of conjugates of peptides with DNA/PNA/analog by employing chemo-selective click reaction in water. *Nucleic Acids Res.* **2007**, *35*, e139. [[CrossRef](#)] [[PubMed](#)]
23. Brown, S.D.; Graham, D. Conjugation of an oligonucleotide to Tat, a cell-penetrating peptide, via click chemistry. *Tetrahedron Lett.* **2010**, *51*, 5032–5034. [[CrossRef](#)]
24. Proudnikov, D.; Timofeev, E.; Mirzabekov, A. Immobilization of DNA in polyacrylamide gel for the manufacture of DNA and DNA-oligonucleotide microchips. *Anal. Biochem.* **1998**, *259*, 34–41. [[CrossRef](#)] [[PubMed](#)]
25. Khym, J.X. The reaction of methylamine with periodate-oxidized adenosine 5'-phosphate. *Biochemistry* **1963**, *2*, 344–350. [[CrossRef](#)] [[PubMed](#)]
26. Brown, D.M.; Read, A.P. 932. Nucleotides. Part XLIX. The reduction of the adduct of periodate-oxidised adenosine-5' phosphate and methylamine. *J. Chem. Soc. Perkin 1.* **1965**, 5072–5074. [[CrossRef](#)] [[PubMed](#)]
27. Zatsepin, T.S.; Stetsenko, D.A.; Gait, M.J.; Oretskaya, T.S. Use of carbonyl group addition–elimination reactions for synthesis of nucleic acid conjugates. *Bioconjug. Chem.* **2005**, *16*, 471–489. [[CrossRef](#)] [[PubMed](#)]
28. Zatsepin, T.S.; Stetsenko, D.A.; Arzumanov, A.A.; Romanova, E.A.; Gait, M.J.; Oretskaya, T.S. Synthesis of peptide-oligonucleotide conjugates with single and multiple peptides attached to 2'-aldehydes through thiazolidine, oxime, and hydrazine linkages. *Bioconjug. Chem.* **2002**, *13*, 822–830. [[CrossRef](#)]
29. Forget, D.; Boturyn, D.; Defrancq, E.; Lhomme, J.; Dumy, P. Highly efficient synthesis of peptide-oligonucleotide conjugates: Chemoselective oxime and thiazolidine formation. *Chem. Eur. J.* **2001**, *7*, 3976–3984. [[CrossRef](#)]
30. Battiste, J.L.; Mao, H.; Rao, N.S.; Tan, R.; Muhandiram, D.R.; Kay, L.E.; Frankel, A.D.; Williamson, J.R. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **1996**, *273*, 1547–1551. [[CrossRef](#)]
31. Morii, T.; Hagihara, M.; Sato, S.; Makino, K. In vitro selection of ATP-binding receptors using a ribonucleopeptide complex. *J. Am. Chem. Soc.* **2002**, *124*, 4617–4622. [[CrossRef](#)] [[PubMed](#)]
32. Ellington, A.D.; Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, *346*, 818–822. [[CrossRef](#)] [[PubMed](#)]
33. Tuerk, C.; Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **1990**, *249*, 505–510. [[CrossRef](#)]
34. Hagihara, M.; Fukuda, M.; Hasegawa, T.; Morii, T. A modular strategy for tailoring fluorescent biosensors from ribonucleopeptide complexes. *J. Am. Chem. Soc.* **2006**, *128*, 12932–12940. [[CrossRef](#)] [[PubMed](#)]

35. Nakano, S.; Nakata, E.; Morii, T. Facile conversion of RNA aptamers to modular fluorescent sensors with tunable detection wavelengths. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4503–4506. [[CrossRef](#)] [[PubMed](#)]
36. Nakano, S.; Tamura, T.; Das, R.K.; Nakata, E.; Chang, Y.-T.; Morii, T. A diversity-oriented library of fluorophore-modified receptors constructed from a chemical library of synthetic fluorophores. *ChemBioChem* **2017**, *18*, 2212–2216. [[CrossRef](#)] [[PubMed](#)]
37. Hasegawa, T.; Hagihara, M.; Fukuda, M.; Morii, T. Stepwise functionalization of ribonucleopeptides: Optimization of the response of fluorescein ribonucleopeptide sensors for ATP. *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 1277–1281. [[CrossRef](#)]
38. Nakano, S.; Mashima, T.; Matsugami, A.; Inoue, M.; Katahira, M.; Morii, T. Structural aspects for the recognition of ATP by ribonucleopeptide receptors. *J. Am. Chem. Soc.* **2011**, *133*, 4567–4579. [[CrossRef](#)]
39. Liew, F.F.; Hasegawa, T.; Fukuda, M.; Nakata, E.; Morii, T. Construction of dopamine sensors by using fluorescent ribonucleopeptide complexes. *Bioorg. Med. Chem.* **2011**, *19*, 4473–4481. [[CrossRef](#)]
40. Hasegawa, T.; Hagihara, M.; Fukuda, M.; Nakano, S.; Fujieda, N.; Morii, T. Context-dependent fluorescence detection of a phosphorylated tyrosine residue by a ribonucleopeptide. *J. Am. Chem. Soc.* **2008**, *130*, 8804–8812. [[CrossRef](#)]
41. Nakano, S.; Fukuda, M.; Tamura, T.; Sakaguchi, R.; Nakata, E.; Morii, T. Simultaneous detection of ATP and GTP by covalently linked fluorescent ribonucleopeptide sensors. *J. Am. Chem. Soc.* **2013**, *135*, 3465–3473. [[CrossRef](#)] [[PubMed](#)]
42. Nakano, S.; Shimizu, M.; Dinh, H.; Morii, T. Highly selective dual sensing of ATP and ADP using fluorescent ribonucleopeptide sensors. *Chem. Commun.* **2019**, *55*, 1611–1614. [[CrossRef](#)] [[PubMed](#)]
43. Nakata, E.; Liew, F.F.; Uwatoko, C.; Kiyonaka, S.; Mori, Y.; Katsuda, Y.; Endo, M.; Sugiyama, H.; Morii, T. Zinc-finger proteins for site-specific protein positioning on DNA-origami structures. *Angew. Chem. Int. Ed.* **2012**, *51*, 2421–2424. [[CrossRef](#)] [[PubMed](#)]
44. Ngo, T.A.; Nakata, E.; Saimura, M.; Kodaki, T.; Morii, T. A protein adaptor to locate a functional protein dimer on molecular switchboard. *Methods* **2014**, *67*, 142–150. [[CrossRef](#)]
45. Nakata, E.; Dinh, H.; Ngo, T.A.; Saimura, M.; Morii, T. A modular zinc finger adaptor accelerates the covalent linkage of proteins at specific locations on DNA nanoscaffolds. *Chem. Commun.* **2015**, *51*, 1016–1019. [[CrossRef](#)]
46. Ngo, T.A.; Nakata, E.; Saimura, M.; Morii, T. Spatially organized enzymes drive cofactor-coupled cascade reactions. *J. Am. Chem. Soc.* **2016**, *138*, 3012–3021. [[CrossRef](#)]
47. Nguyen, T.M.; Nakata, E.; Saimura, M.; Dinh, H.; Morii, T. Design of modular protein-tags for the orthogonal covalent bond formation at specific DNA sequences. *J. Am. Chem. Soc.* **2017**, *139*, 8487–8496. [[CrossRef](#)]
48. Nakata, E.; Nakano, S.; Rajendran, A.; Morii, T. Covalent bond formation by modular adaptors to locate multiple enzymes on a DNA scaffold. In *Kinetic Control in Synthesis and Self-Assembly*; Numata, M., Yagai, S., Hamura, T., Eds.; Academic Press: Cambridge, MA, USA, 2019; pp. 163–183.
49. Nguyen, T.M.; Nakata, E.; Zhang, Z.; Saimura, M.; Dinh, H.; Morii, T. Rational design of a DNA sequence-specific modular protein tag by tuning the alkylation kinetics. *Chem. Sci.* **2019**, *10*, 9315–9325. [[CrossRef](#)]
50. Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **2003**, *21*, 86–89. [[CrossRef](#)]
51. Gautier, A.; Juillerat, A.; Heinis, C.; Correa, I.R.; Kindermann, M.; Beaufils, F.; Johnsson, K. An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **2008**, *15*, 128–136. [[CrossRef](#)]
52. Los, G.V.; Encell, L.P.; McDougall, M.G.; Hartzell, D.D.; Karassina, N.; Zimprich, C.; Wood, M.G.; Learish, R.; Ohana, R.F.; Urh, M.; et al. HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **2008**, *3*, 373–382. [[CrossRef](#)] [[PubMed](#)]
53. Tamura, T.; Nakano, S.; Nakata, E.; Morii, T. Construction of a library of structurally diverse ribonucleopeptides with catalytic groups. *Bioorg. Med. Chem.* **2017**, *25*, 1881–1888. [[CrossRef](#)] [[PubMed](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Design of the Crosslinking Reactions for Nucleic Acids-Binding Protein and Evaluation of the Reactivity

Kenta Odaira, Ken Yamada, Shogo Ishiyama, Hidenori Okamura and Fumi Nagatsugi \* 

Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai 980-8577, Japan; k.odaira@kobayashi.co.jp (K.O.); Ken.Yamada@umassmed.edu (K.Y.); rare.earth.144@gmail.com (S.I.); hidenori.okamura.b8@tohoku.ac.jp (H.O.)

\* Correspondence: nagatugi@tohoku.ac.jp; Tel.: +81-22-217-5633

Received: 2 October 2020; Accepted: 29 October 2020; Published: 30 October 2020



**Featured Application:** Alkylation for nucleic acids-binding protein.

**Abstract:** Selective chemical reactions of biomolecules are some of the important tools for investigations by biological studies. We have developed the selective crosslinking reactions to form covalent bonds to DNA or RNA using crosslinking oligonucleotides (CFO) bearing reactive bases. In this study, we designed the cross-linkable 4-amino-6-oxo-2-vinyltriazine derivative with an acyclic linker (*acyAOVT*) to react with the nucleic acids-binding protein based on our previous results. We hypothesized that the *acyAOVT* base would form a stable base pair with guanine by three hydrogen bonds at the positions of the vinyl group in the duplex DNA major groove, and the vinyl group can react with the nucleophilic species in the proximity, for example, the cysteine or lysine residue in the nucleic acids-binding protein. The synthesized oligonucleotides bearing the *acyAOVT* derivative showed a higher reactivity than that of the corresponding pyrimidine derivative without one nitrogen. The duplex containing *acyAOVT*-guanine (G) formed complexes with Hha1 DNMT even in the presence of 2-mercaptoethanol. We expect that our system will provide a useful tool for the molecular study of nucleic acids-binding proteins.

**Keywords:** oligonucleotide; crosslink; nucleic acid binding protein; cytosine methyltransferase

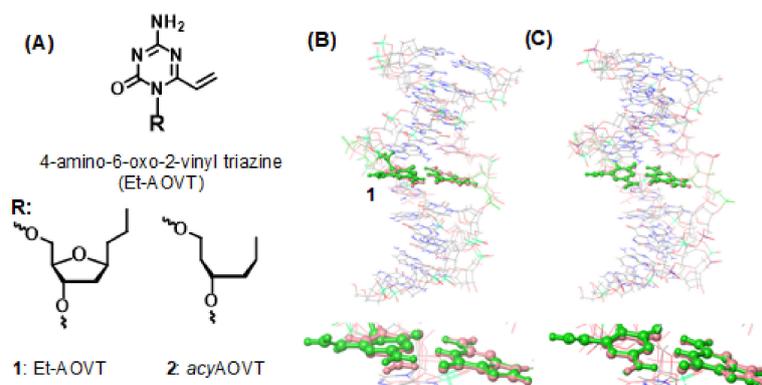
## 1. Introduction

The binding of proteins to nucleic acids (DNA and RNA) is central to all aspects of gene expression regulation. DNA-binding proteins include the transcription factor to bind to specific DNA sequences and control the transcription rate of genetic information from the DNA to RNA [1,2]. Many DNA modifying enzymes are also included in the DNA binding protein, for example, repair enzymes [3,4] and epigenetic modification enzymes. DNA methyl transferase is one of the epigenetic modification enzymes and catalyses the transfer of the methyl group to DNA modifying the function of genes and affecting gene expression [5,6]. RNA-binding proteins have important functions in the post-transcriptional process, such as splicing regulation [7] and modulation of the mRNA translation. Recent studies have revealed that post-transcriptional gene regulation by non-coding RNAs is involved in most biological activities. The association of the RNA binding protein with ncRNA plays a crucial role in these biological functions [8–10]. The chemical tools for the control of the interaction between the nucleic acids-binding protein and nucleic acids have the potential for the development as the new strategy for artificial control of the gene expression.

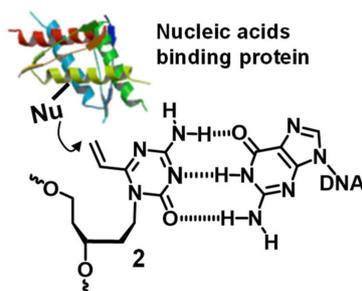
Crosslinking reactions between nucleic acids and binding proteins are powerful tools for analyzing these interactions. Photoreactive functional groups, such as diazirines and benzophenones,

were utilized for DNA or RNA-protein crosslinking [11–14], but lack selectivity for the target amino acid. Furthermore, due to their intrinsic reactivity, these groups react with water and other reactive chemical species resulting in low yields. Oligodeoxynucleotides (ODN) with reactive groups are exploited for their effective DNA-protein crosslinking reactions by the proximity effect of the specific DNA recognition with the binding protein. The reactive groups activated by oxidation, such as a diol [15] and furan [16], react with a proximal lysine or arginine in the protein or peptides. Hocek et al. reported that the ODN bearing a vinylsulfone amide [17] or chloroacetamide group [18] efficiently cross-linked with the p53 protein through alkylation of the cysteine in the proximal position with the reactive groups. Recently, they demonstrated that ODN with 2-vinylhypoxantine reacted with a thiol-containing minor groove binding peptide by proximity effect [19].

In our previous study, we reported the synthesis of the ODN bearing 4-amino-6-oxo-2-vinyltriazine with an ethyl linker (Et-AOVT: (1)) as a cross linkable derivative and an evaluation of its properties [20]. ODN showed a relatively high reactivity with the pyrimidine nucleobases of the complementary DNAs and the lowest reactivity to guanine. We hypothesized that the AOVT base would form a stable base pair with guanine (G) by three hydrogen bonds to be positioned at the vinyl group in the DNA major groove. In this study, we have designed an AOVT derivative with an acyclic linker (*acy*AOVT: (2)) as a cross-linkable probe of the nucleic acids-binding protein. The distance between the reactive base and sugar moiety of 2 is shorter than that of 1, and consequently, the *acy*AOVT (2) should form a base pair with guanine (G) similar to the natural cytosine (C)-G pair. This assumption is supported by the molecular modeling shown in Figure 1. The molecular modeling revealed that the structure of the base pair of *acy*AOVT (2)-G and the natural C-G significantly overlap each other (Figure 1C). On the other hand, the structure of Et-AOVT (1)-G is subtly shifted from the natural base pair one (Figure 1B). The duplex DNA containing the *acy*AOVT (2)-G is expected to bind to the nucleic acid binding proteins similar to the natural duplex and form covalent linkages with the nucleophilic residues of the protein in the proximal position of the vinyl group (Figure 2).



**Figure 1.** (A) Structures of cross-linkable derivatives; (B,C) Superimposed structures between the natural duplex (pink) and the duplex contained 1 (B) or 2 (C) (green).



**Figure 2.** Design of the cross-linkable probe to nucleic acids-binding protein.

## 2. Materials and Methods

### 2.1. General

The  $^1\text{H}$ -NMR spectra were obtained by 400 or 600 MHz spectrometers (Bruker, Billerica, MA, USA). The  $^1\text{H}$  chemical shifts are described as  $\delta$  values in ppm relative to acetone- $d_6$  (2.05 ppm), DMSO- $d_6$  (2.50 ppm),  $\text{CDCl}_3$  (7.26 ppm), and tetramethylsilane (1H, 0.00 ppm). The  $^{13}\text{C}$  NMR spectra were obtained by a Bruker 600 MHz spectrometer. The  $^{13}\text{C}$  chemical shifts are described as the  $\delta$  values in ppm relative to acetone- $d_6$  (29.84 ppm), DMSO- $d_6$  (39.52 ppm), and  $\text{CDCl}_3$  (77.16 ppm). The  $^{31}\text{P}$ -NMR spectra were recorded by a Bruker 500 MHz (202 MHz for  $^{31}\text{P}$ ). The multiplicity and qualifier abbreviations are as follows: s = singlet, d = doublet, t = triplet, q = quartet, quin. = quintet, sept. = septet, m = multiplet, br = broad. The electrospray ionization (ESI) mass spectra were recorded using a BioTOF II mass spectrometer or APEX III (Bruker Daltonics, Bruker, Billerica, MA, USA). The matrix-assisted laser desorption/ionization (MALDI-TOF) mass spectra were recorded Autoflex speed mass spectrometer and a laser at 337 nm in the negative mode using 3-hydroxypicolinic acid as the matrix or in the positive mode using 2,5-dihydroxybenzoic acid as the matrix. Thin-layer chromatography (TLC) was performed using silica gel 60 F<sub>254</sub> pre-coated plates. Column chromatography was performed using silica gel 60 N (spherical, neutral, 100–210  $\mu\text{m}$ , Kanto Chemical, Tokyo, Japan). Flash chromatography was performed using Kanto Chemical silica gel 60 N (spherical, neutral, 40–50  $\mu\text{m}$ ). The ultraviolet-visible (UV-vis) absorption spectra were recorded using a DU800 spectrometer (Beckman Coulter, Brea, CA, USA). The ODN synthesis was carried out using an automated DNA synthesizer (392 DNA/RNA synthesizer, ABI, Foster City, CA, USA) following the standard phosphoramidite chemistry. High performance liquid chromatography (HPLC) was performed using a Cosmosil 5C18MSII column (4.6 or 10  $\times$  250 mm, Nacalai Tesque, Kyoto, Japan), a PU-986 pump (JASCO, Tokyo, Japan), JASCO 2075 UV detector and a JASCO 2067 column oven. pH measurements were measured by a Seven Easy pH meter (Mettler Toledo, Columbus, OH) using an 8220BNWP electrode. The denaturing polyacrylamide gel plates were visualized and quantified using a FLA-5100 Fluor Imager (Fujifilm, Tokyo, Japan). Anhydrous methanol, DMF, THF,  $\text{CH}_2\text{Cl}_2$ , dioxane, pyridine, DMSO,  $\text{CH}_3\text{CN}$ , toluene and THF were purchased from Wako Pure Chemical Industries, Ltd. (Tokyo, Japan). Unless otherwise noted, all the synthetic reactions were carried out at ambient temperature. The reactions requiring anhydrous conditions were achieved under an argon atmosphere in flasks dried under 1–3 mmHg. Commercially available reagents were obtained from Wako Pure Chemical Industries, Ltd., TCI, Inc., (Tokyo, Japan), Sigma-Aldrich Co. LLC (St. Louis, MO) and Kanto Chemical Co., Inc (Tokyo, Japan) and used without further purification. Human DNA (cytosine-5) methyltransferase (DNMT1) was obtained from New England Bio Labs Japan, Inc. (Tokyo, Japan)

### 2.2. Synthesis of the Nucleoside Derivatives

#### 2.2.1. Synthesis of 4-amino-6-oxo-1-(3',5'-O-t-butylidimethylsilyl-2',4'-dideoxy-D-ribose) triazine (4)

The linker **3** (1.3 g, 3.0 mmol), which was prepared by a modified literature procedure [21], was added to 5-aza cytosine Na salt (1.0 g, 7.5 mmol) in DMSO (30 mL). The mixture was stirred at 60  $^\circ\text{C}$  for 18 h. The resulting mixture was diluted with EtOAc and quenched with  $\text{NH}_4\text{Cl}$ . The product was extracted with EtOAc and the combined organic layer was washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered, and evaporated under reduced pressure. The resulting oil was purified by silica gel column chromatography ( $\text{CH}_3\text{OH}/\text{CHCl}_3$ , 1:100, then  $\text{CH}_3\text{OH}/\text{CHCl}_3$ , 1:50) to afford the corresponding N1-alkylation (772 mg, 1.74 mmol, 58%) and N3-alkylation (126 mg, 0.29 mmol, 10%) products.

N1-alkylation product:  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.90 (s, 1H), 6.98 (brs, 1H), 5.65 (brs, 1H), 3.97 (ddd, 2H,  $J = 12.0, 12.0, 6.0$  Hz), 3.90–3.77 (m, 2H), 3.64 (dd, 2H,  $J = 6.4, 6.4$  Hz), 1.99–1.60 (m, 6H), 1.37–1.23 (m, 1 H), 0.89 (s, 9H), 0.87 (s, 9H), 0.07 (s, 3H), 0.06 (s, 3H), 0.03 (s, 6H);  $^{13}\text{C}$ -NMR (125 MHz,  $\text{CDCl}_3$ )  $\delta$  166.7, 158.7, 154.5, 67.4, 59.6, 45.1, 40.2, 36.2, 26.12, 26.09, 18.4, 18.3, −4.2, −4.3, −5.14, −5.16; HRMS (ESI) calcd. for  $\text{C}_{20}\text{H}_{43}\text{N}_4\text{O}_3\text{Si}_2^+$  [ $\text{M} + \text{H}$ ] $^+$   $m/z$  443.2868, found  $m/z$  443.2879.

N3-alkylation product:  $^1\text{H-NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$  8.34 (s, 1H), 5.90 (brs, 1H), 5.61 (brs, 1H), 4.42–4.35 (m, 1H), 4.05–4.01 (m, 1H), 3.70–3.64 (m, 2H), 1.99–1.95 (m, 1H), 1.89–1.84 (m, 2H), 1.7–1.67 (m, 2H), 0.871 (s, 9H), 0.867 (s, 9H), 0.046 (s, 3H), 0.027 (s, 3H), 0.025 (s, 3H), 0.018 (s, 3H);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ )  $\delta$  170.4, 167.93, 167.88, 66.3, 64.5, 59.6, 40.3, 35.9, 25.9, 25.8, 18.2, 18.0, –4.6, –4.8, –5.36, –5.38; HRMS (ESI) calcd. for  $\text{C}_{20}\text{H}_{43}\text{N}_4\text{O}_3\text{Si}_2^+$   $[\text{M} + \text{H}]^+$   $m/z$  443.2868, found  $m/z$  443.2879

#### 2.2.2. Synthesis of 4-amino-2-(2-octylthioethyl)-6-oxo-1-(3',5'-O-t-butyl dimethylsilyl)-2',4'-dideoxy-D-ribose triazine (6)

A 28% aq  $\text{NH}_4\text{OH}$  (10 mL) solution was added to a solution of the N1-alkylation product **4** (172 mg, 390  $\mu\text{mol}$ ) in dioxane-methanol (10 mL, 1/1, *v/v*), then sealed with a glass stopper. The mixture was stirred at 55 °C for 24 h. The resulting solution was evaporated under vacuum at 43 °C. The resulting moist solid was co-evaporated with ethanol  $\times$  3 to afford guanylylurea **5**. This product was directly used for the next step without further purification. The orthoester (597 mg, 1.95 mmol, purity = 91%) was added to a solution of guanylylurea (169 mg, 390  $\mu\text{mol}$ ) in DMF (0.2 M, 1.95 mL). The mixture was stirred at 120 °C for 2 h. The resulting yellow solution was diluted with  $\text{Et}_2\text{O}$ . The product was extracted with  $\text{Et}_2\text{O}$  and the combined organic layer was washed with  $\text{NH}_4\text{Cl}$  aq.,  $\text{NaHCO}_3$  aq., and brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and evaporated. The crude product was purified by silica gel column chromatography ( $\text{CHCl}_3$  then  $\text{CHCl}_3$  with 1%  $\text{Et}_3\text{N}$ , then  $\text{CH}_3\text{OH}/\text{CHCl}_3$ , 1:300 with 1%  $\text{Et}_3\text{N}$ ) to afford **6** (198.3  $\mu\text{mol}$ , 51%) as a yellow oil:  $^1\text{H-NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$  5.70 (brs, 1H), 5.17 (brs, 1H), 4.02–3.89 (m, 3H), 3.65 (dd, 2H,  $J = 6.6, 6.6$  Hz), 2.95–2.88 (m, 3H), 2.54 (t, 2H,  $J = 7.8$  Hz), 1.93–1.88 (m, 1 H), 1.79–1.64 (m, 4H), 1.59 (quin, 2H,  $J = 7.2$  Hz), 1.39–1.25 (m, 10H), 0.91 (s, 9H), 0.89 (t, 3H,  $J = 7.2$  Hz), 0.88 (s, 9H), 0.089 (s, 3H), 0.086 (s, 3H), 0.04 (s, 6H);  $^{13}\text{C-NMR}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  168.8, 165.3, 155.8, 67.6, 59.6, 41.6, 40.4, 35.8, 34.3, 33.0, 32.1, 29.9, 29.5, 29.2, 28.2, 26.19, 26.18, 22.9, 18.5, 18.3, 14.4, 0.26, –4.18, –4.26, –4.30, –5.1; HRMS (ESI) calcd. for  $\text{C}_{30}\text{H}_{63}\text{N}_4\text{O}_3\text{SSi}_2^+$   $[\text{M} + \text{H}]^+$   $m/z$  615.4154, found  $m/z$  615.4158.

#### 2.2.3. Synthesis of 2-(2-octylthioethyl)-6-oxo-4-acetylamino-1-(3',5'-O-t-butyl dimethylsilyl)-2',4'-dideoxy-D-ribose triazine (7)

$\text{AcCl}$  (5.3  $\mu\text{L}$ , 74.7  $\mu\text{mol}$ ) was added to a solution of **6** (15 mg, 24.9  $\mu\text{mol}$ ) in pyridine (250  $\mu\text{L}$ ). The mixture was stirred at rt for 19 h. The resulting solution was diluted with  $\text{Et}_2\text{O}$  and quenched with  $\text{NH}_4\text{Cl}$ . The product was extracted with  $\text{Et}_2\text{O}$  and the combined organic layer was washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and evaporated. The crude product was purified by silica gel column chromatography ( $\text{CHCl}_3$ , then  $\text{CHCl}_3$  with 1%  $\text{Et}_3\text{N}$ ) to afford **7** (13.4  $\mu\text{mol}$ , 54%) as a colorless oil:  $^1\text{H-NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.71 (br, 1H), 4.08–3.98 (m, 3H), 3.66 (t, 2H,  $J = 6.0$  Hz), 3.65 (t, 2H,  $J = 6.0$  Hz), 3.00 (t, 2H,  $J = 6.6$  Hz), 2.91 (t, 2H,  $J = 6.6$  Hz), 2.61 (s, 3H), 2.54 (t, 2H,  $J = 7.2$  Hz), 1.95 (m, 1H), 1.81–1.64 (m, 3H), 1.58 (qt, 2H,  $J = 7.2, 7.2$  Hz), 1.39–1.26 (m, 10 H), 0.92 (s, 9H), 0.91 (t, 3H,  $J = 7.2$  Hz), 0.88 (s, 9H), 0.095 (s, 3H), 0.093 (s, 3H), 0.041 (s, 3H), 0.03 (s, 3H);  $^{13}\text{C-NMR}$  (150 MHz,  $\text{CDCl}_3$ )  $\delta$  172.0, 170.2, 161.5, 154.8, 41.9, 40.0, 34.9, 34.1, 32.8, 29.7, 29.5, 29.2, 28.8, 27.7, 25.87, 25.86, 25.85, 22.6, 18.2, 18.0, 14.1, –4.52, –4.55, –5.38, –5.41; HRMS (ESI) calcd. for  $\text{C}_{32}\text{H}_{64}\text{N}_4\text{O}_4\text{SSi}_2^+$   $[\text{M} + \text{H}]^+$   $m/z$  657.4246, found  $m/z$  657.4246.

#### 2.2.4. Synthesis of 2-(2-octylthioethyl)-6-oxo-4-acetylamino-1-(2',4'-dideoxy-D-ribose) triazine (8)

Boron trifluoride-diethylether complex (51  $\mu\text{L}$ , 398  $\mu\text{mol}$ ) was slowly added to a solution of **7** (131 mg, 199  $\mu\text{mol}$ ) in  $\text{CH}_3\text{CN}$  (800  $\mu\text{L}$ ) at 0 °C. After the addition, the mixture was stirred at 0 °C for 2 h. The resulting solution was diluted with  $\text{EtOAc}$ , then quenched with phosphate buffer (pH = 6.0, 1 M). The product was extracted with  $\text{EtOAc}$  and the combined organic layer was washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and evaporated. The crude product was purified by silica gel chromatography, eluting with  $\text{CH}_3\text{OH}/\text{CHCl}_3$ , 1:50, then 1:10 to give **8** (103  $\mu\text{mol}$ , 52%) as a white solid:  $^1\text{H-NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.80 (br, 1H), 4.38–4.33 (m, 1H), 4.00–3.96 (m, 1H), 3.91–3.88 (m, 1H), 3.85–3.81 (m, 2H), 3.15 (dt,  $J = 16.8, 1\text{H}, 7.2$  Hz), 3.05 (dt, 1H,  $J = 16.8, 7.2$  Hz), 2.93 (t, 2H,  $J = 7.2$  Hz), 2.61 (s, 3H), 2.56 (t, 2H,  $J = 7.2$  Hz), 1.95–1.90 (m, 1H), 1.82–1.71 (m, 3H), 1.59 (qt, 2H,  $J = 7.2, 7.2$  Hz), 1.38–1.25 (m, 10H),

0.88 (t, 3H,  $J = 7.2$  Hz);  $^{13}\text{C}$ -NMR (150 MHz,  $\text{CDCl}_3$ )  $\delta$  171.9, 170.9, 161.7, 156.0, 68.1, 61.63, 61.59, 41.7, 38.0, 36.2, 34.1, 32.8, 31.8, 29.6, 29.2, 28.8, 27.9, 25.9, 22.6, 14.1; HRMS (ESI) calcd. for  $\text{C}_{20}\text{H}_{36}\text{N}_4\text{O}_4\text{S}^+$   $[\text{M} + \text{H}]^+$   $m/z$  429.2528, found  $m/z$  429.2530.

### 2.2.5. Synthesis of 2-(2-octylthioethyl)-6-oxo-4-acetylamino-1-(5'-O-(4,4'-dimethoxytrityl)-2',4'-dideoxy-D-ribose) triazine (9)

DMTrCl (58 mg, 170  $\mu\text{mol}$ ) was added to a solution of **8** (497 mg, 114  $\mu\text{mol}$ ) in pyridine (487  $\mu\text{L}$ ) at 0 °C. After the addition, the mixture was stirred at 0 °C for 1 h. The resulting mixture was diluted with  $\text{CH}_2\text{Cl}_2$ , then quenched with  $\text{NaHCO}_3$ . The product was extracted with  $\text{CH}_2\text{Cl}_2$  and the combined organic layer was washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and evaporated. The crude product was purified by chromatography on silica gel, eluting with  $\text{CHCl}_3$  and 1%  $\text{Et}_3\text{N}$ , then  $\text{CH}_3\text{OH}/\text{CHCl}_3$ , 1:200 with 1%  $\text{Et}_3\text{N}$  to give **9** (80  $\mu\text{mol}$ , 70%) as a colorless oil:  $^1\text{H}$ -NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.72 (br, 1H), 7.39–7.38 (m, 2H), 7.30–7.27 (m, 6H), 7.21–7.19 (m, 1H), 6.83–6.80 (m, 2H), 4.20 (dt, 1H,  $J = 13.2, 7.2$  Hz), 4.04 (dt, 1H,  $J = 13.2, 7.2$  Hz), 3.15 (dt, 1H,  $J = 16.8, 7.2$  Hz), 3.06 (dt, 1H,  $J = 16.8, 7.2$  Hz), 2.96 (t, 2H,  $J = 7.2$  Hz), 2.60 (s, 3H), 2.54 (t, 2H,  $J = 7.2$  Hz), 1.96–1.91 (m, 1H), 1.83–1.77 (m, 1H), 1.69–1.66 (m, 2H), 1.57 (qt, 2H,  $J = 7.2, 7.2$  Hz), 1.37–1.26 (m, 10H), 0.87 (t, 3H,  $J = 7.2$  Hz);  $^{13}\text{C}$ -NMR (150 MHz,  $\text{CDCl}_3$ )  $\delta$  172.0, 170.8, 161.6, 158.5, 155.3, 144.5, 135.7, 135.6, 129.88, 129.86, 128.0, 127.9, 126.9, 113.2, 86.9, 68.4, 62.4, 55.2, 42.1, 36.5, 35.5, 34.1, 32.6, 31.8, 29.5, 29.19, 29.17, 28.8, 25.9, 22.6, 14.1; HRMS (ESI) calcd. for  $\text{C}_{41}\text{H}_{54}\text{N}_4\text{O}_6\text{S}^+$   $[\text{M} + \text{H}]^+$   $m/z$  731.3837, found  $m/z$  731.3834.

### 2.2.6. Synthesis of 2-(2-octylthioethyl)-6-oxo-4-acetylamino-1-(3'-N,N-diisopropylcyanoethyl-phosphoramidyl-5'-O-(4,4'-dimethoxytrityl)-2',4'-dideoxy-D-ribose) triazine (10)

DIPEA (83.2  $\mu\text{L}$ , 14.3  $\mu\text{mol}$ ) was added to a solution of **9** (15.2 mg, 20.8  $\mu\text{mol}$ ) in  $\text{CH}_2\text{Cl}_2$  (416  $\mu\text{L}$ ) and the mixture was cooled to 0 °C.  $\text{NCCCH}_2\text{CH}_2\text{OP}[(\text{N}(i\text{-Pr})_2)]\text{Cl}$  (9.3  $\mu\text{L}$ , 41.6  $\mu\text{mol}$ ) was then added to the mixture. After the addition, the mixture was stirred at 0 °C for 1 h, diluted with  $\text{CH}_2\text{Cl}_2$ , then quenched with  $\text{NaHCO}_3$ . The product was extracted with  $\text{CH}_2\text{Cl}_2$  and the combined organic layer was washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and evaporated. The crude product was co-evaporated with toluene, then purified by chromatography on silica gel, eluting with  $\text{EtOAc}/\text{hexane}$ , 1:1 with 1%  $\text{Et}_3\text{N}$ , then 2:3 with 1%  $\text{Et}_3\text{N}$  to give **10** (15.0  $\mu\text{mol}$ , 72%) as a colorless oil mixture of two diastereomers:  $^1\text{H}$ -NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.65 (br, 1H), 7.41 (d, 2H,  $J = 7.2$  Hz), 7.31–7.27 (m, 6H), 7.21–7.18 (m, 1H), 6.83–6.81 (m, 4H), 4.22–4.04 (m, 3H), 3.82–3.49 (m, 4H), 3.79 (s, 6H), 3.22–3.15 (m, 2H), 3.10–2.95 (m, 2H), 2.91–2.86 (m, 2H), 2.64–2.61 (m, 1H), 2.62–2.61 (m, 3H), 2.55–2.50 (m, 2H), 2.48–2.46 (m, 1H), 2.01–1.77 (m, 4H), 1.59–1.54 (m, 2H), 1.39–1.25 (m, 10H), 1.16–1.03 (m, 12H), 0.88 (t, 3H,  $J = 7.2$  Hz);  $^{13}\text{C}$ -NMR (150 MHz,  $\text{CDCl}_3$ )  $\delta$  172.04, 172.00, 170.4, 170.2, 161.44, 161.40, 158.3, 154.9, 154.7, 145.10, 145.05, 136.32, 136.30, 136.29, 136.27, 130.0, 128.1, 128.0, 127.7, 126.7, 126.6, 118.0, 117.6, 113.0, 86.0, 77.2, 77.0, 76.8, 70.4, 70.3, 70.0, 69.9, 60.4, 60.0, 58.0, 57.8, 57.4, 57.3, 55.18, 55.15, 43.08, 43.05, 43.00, 42.97, 42.0, 41.9, 36.82, 36.80, 36.25, 36.23, 34.4, 34.3, 34.2, 34.04, 33.95, 32.82, 32.79, 31.8, 29.53, 29.52, 29.18, 29.16, 28.84, 28.83, 27.7, 25.9, 24.80, 24.75, 24.60, 24.56, 24.52, 24.50, 22.6, 21.1, 20.50, 20.45;  $^{31}\text{P}$ -NMR (202 MHz,  $\text{CDCl}_3$ )  $\delta$  147.70, 146.72; HRMS (ESI) calcd. for  $\text{C}_{50}\text{H}_{71}\text{N}_6\text{O}_7\text{PS}^+$   $[\text{M} + \text{H}]^+$   $m/z$  931.4915, found  $m/z$  931.4919.

## 2.3. Synthesis of the Oligodeoxynucleotides Containing acyAOVT Derivatives

**ODN1** and **2** were synthesized on a 1  $\mu\text{mol}$  scale by an ABI 392 DNA/RNA synthesizer with standard  $\beta$ -cyanoethyl chemistry. 5'-Terminal dimethoxytrityl-bearing **ODN1** and **2** was removed from the solid support by treatment with 45 mM  $\text{K}_2\text{CO}_3$ -MeOH containing 10 mM 1-octanethiol (0.5 mL) and the residue was evaporated under reduced pressure. The crude product was purified by reverse phase HPLC with a C-18 column (Nacalai Tesque: Cosmosil 5C18-MS-II, 10  $\times$  250 mm) by a linear gradient of 5–40%/25 min of acetonitrile in 0.1% TEAA buffer at the flow rate of 4 mL/min. The dimethoxytrityl group of the purified **ODN** was removed with 10% AcOH for 30 min and the mixture was additionally purified by reverse phase HPLC to afford **ODN1**; MALDI-TOF MS ( $m/z$ ):

calcd. for [M-H]<sup>-</sup> *m/z* 4045.8746; found 4045.780; **ODN2**; MALDI-TOF MS (*m/z*): calcd. for [M-H]<sup>-</sup> *m/z* 6960.8115; found 6958.589.

To a solution of **ODN1** or **2** (70 μM) in ddH<sub>2</sub>O was added a solution of MMPP (1 mM) in carbonate buffer (pH = 10) at room temperature. After 1 h, 5% AcOH was added to the mixture and the mixture was left for an additional 2 h to give **ODN3**. MALDI-TOF MS **ODN** (SOOct): calcd. for [M-H]<sup>-</sup> *m/z* 4061.8736, found 4061.755, **ODN5** (vinyl): calcd. for [M-H]<sup>-</sup> *m/z* 3899.5826, found 3899.304; **ODN6** (vinyl): calcd. for [M-H]<sup>-</sup> *m/z* 6814.5195, found 6811.219

**ODN7** was prepared by using 1.0 μM **ODN5** and 100 mM sodium thiomethoxide in 50 mM MES buffer (pH = 7.0) at 37 °C for 19 h. The reaction mixture was purified by reverse phase HPLC to afford **ODN7**: calcd. for [M-H]<sup>-</sup> *m/z* 3947.6856, found 3947.309.

#### 2.4. General Procedure for the Crosslinking Reactions

The reaction was performed using 5.0 μM **ODN5** and 2.0 μM of the target DNA or RNA labelled by fluorescein at the 5'-end in a buffer of 100 mM NaCl and 50 mM MES buffer (pH = 7.0). The reaction was incubated at 37 °C for 1–24 h. The reaction was quenched with the addition of loading dye (95% formamide, 20 mM EDTA, 0.05% xylene cyanol, and 0.05% bromophenol blue). The cross-linked products were analyzed by a denaturing 20% polyacrylamide gel electrophoresis containing urea (7 M) with TBE buffer at 200 V for 1 h. The labelled bands were visualized and quantified using a FLA-5100 Fluor Imager.

#### 2.5. *T<sub>m</sub>* Measurement

All samples for the *T<sub>m</sub>* measurements consisted of 100 mM NaCl, 25 mM MES buffer (pH = 7.0) and 1 μM Duplex. The *T<sub>m</sub>* measurements were performed using a temperature controller. Both the heating and cooling curves were measured three times over the temperature range of 25 °C to 80 °C at 0.5 °C/min. The absorbance at 260 nm was recorded every 0.5 °C.

#### 2.6. Alkylation with ODN Probe to HhaI DNMT-1

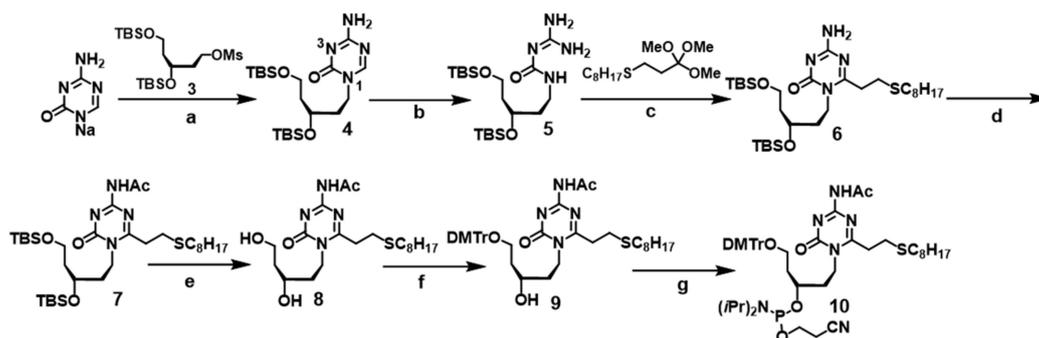
The duplex oligo (final 1 μM) in pH 7.5 reaction buffer (50 mM Tris-HCl, 10 mM EDTA, 5 mM 2-mercaptoethanol) was pre-incubated at 37 °C for 5 min. HhaI Methyltransferase was then added to the reaction mixture and the mixture was pre-incubated at 37 °C (or 4 °C) for another 5 min. SAM was added to the mixture and the mixture was incubated at 37 °C (or 4 °C) for 0.5–24 h. A 4.5 μL aliquot of the reaction mixture (4.5 μL: for silver stain, 12 μL: for Flamingo gel stain) was used every time. Loading buffer (6×Tris.HCl, SDS, glycerol, 0.05% bromophenol blue, 200 mM DTT (final 20 mM)) was added and the mixture was heated at 65 °C for 15 min. The samples were separated by 10% SDS-PAGE (0.025 M Tris, 0.192 M glycine, 0.1% SDS) at room temperature (10 mA/1.5 h then 20 mA/1.0 h). Visualization of the 5'- FAM oligonucleotide was performed by fluorescence imaging using an FLA-5100 Fluor Imager. Visualization of the protein was performed by silver stain or Flamingo gel stain.

### 3. Results

#### 3.1. Synthesis of the acyAOVT Nucleoside Phosphoramidite

We planned to synthesize the C6-substituted 5-azacytosine by ring opening of the 5-azacytosine moiety forming guanylylurea and subsequent ring closing reaction that follows a previous report (Scheme 1) [20]. In our previous study, the acyclic linkers-protected MOM or Bn group produced the coupling products in low yields. The acyclic side chain 3-protected TBS group was synthesized from 2'-deoxy-D-ribose by a previous modified procedure [21]. The coupling reaction between the sodium salt of 5-azacytosine and the acyclic side chain compound **3** was done to form the N1- and N3- alkylated products. The reaction conditions were screened in order to optimize the formation of the N1-alkylated product **4** and a considerable amount of the N1-alkylation (58%) in DMSO at 60 °C

was achieved with a N1:N3 ratio = 6:1. The formation of the glycosidic bond at the N1 position of 5-azacytosine was confirmed by HMQC and HMBC analyses.

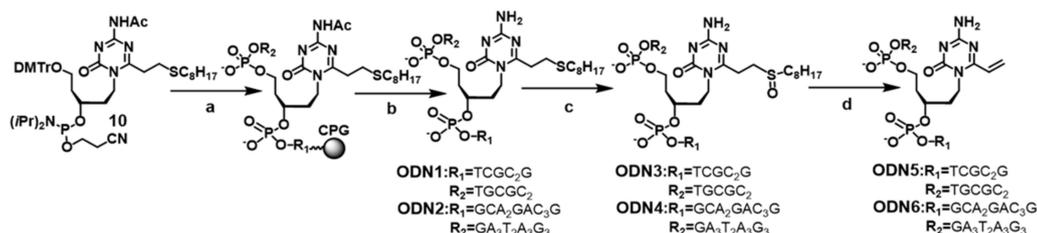


**Scheme 1.** Reagents and conditions: (a) 5-azacytosine sodium salt, DMSO, 60 °C, 18 h, 58%; (b) 7M NH<sub>4</sub>OH, MeOH-1,4-dioxane (1:1), 55 °C, 24 h; (c) orthoester, DMF, 120 °C, 2 h, 51% (2 steps); (d) AcCl, pyridine, rt, 19 h, 54%; (e) BF<sub>3</sub>·OEt<sub>2</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C, 2 h, 52%; (f) DMTrCl, pyridine, 0 °C, 1 h, 70%; (g) 2-Cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C, 1 h, 72%.

The 5-azacytidine derivative **4** was treated with NH<sub>3</sub> to give the guanylylurea intermediate **5**, which was condensed with the orthoester to afford the desired C6-octylthioethyl-5-azacytidine derivative **6** in 45% yield (2 steps). After the 4-*N*-acetylation and the deprotection of the TBS groups by BF<sub>3</sub>OEt<sub>2</sub>, the 5'-hydroxyl group was selectively protected with the DMTr group, then the 3'-hydroxyl group was phosphorylated to yield the *acy*AOVT nucleoside phosphoramidite **10**.

### 3.2. Synthesis of the Oligonucleotides Containing *acy*AOVT and Evaluation of the Crosslinking Reactivity

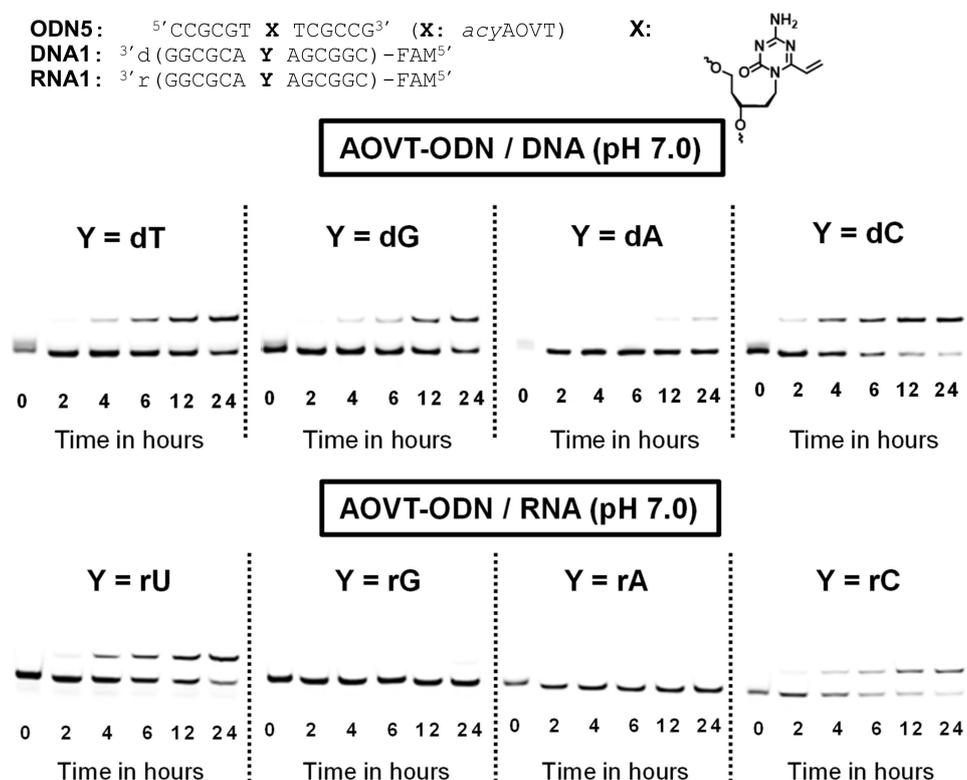
We synthesized two kinds of ODNs **1**, **2** by a DNA synthesizer using phosphoramidite **10** (Scheme 2). The synthesized ODNs were treated with 45 mM K<sub>2</sub>CO<sub>3</sub>/MeOH containing 10 mM 1-octanethiol to cleave them from the resin. After the DMTr-ON purification by reverse-phase (RP) HPLC, detritylation was carried out in an aqueous 10% AcOH solution at room temperature for 30 min. The obtained ODNs were further purified by RP-HPLC and were characterized by a MALDI-TOF MS analysis. The sulfide group of ODNs **1**, **2** was oxidized with magnesium monoperoxyphthalate (MMPP) and treated with aqueous 10% AcOH to give ODNs **5** and **6** according to the reported procedure [20]. The characterization of the ODNs was performed by MALDI-TOF MS.



**Scheme 2.** Reagents and conditions: (a) DNA synthesizer; (b) (i) 45 mM K<sub>2</sub>CO<sub>3</sub>-MeOH, 10 mM 1-octanethiol, rt, 4 h; (ii) 10% AcOH, rt, 30 min.; (c) 1 mM MMPP (2 eq) in carbonate buffer (pH 10), rt, 1 h; (d) 5% aq. AcOH, rt, 2 h.

The crosslinking reactions using the reactive ODN5 to initially the complementary target DNA1 or RNA1 labelled with fluorescein at the 5' end were investigated under neutral conditions. The reaction mixture was analyzed by 20% polyacrylamide gel electrophoresis (PAGE) containing 7 M urea and the yields of the cross-linked product were calculated based on the fluorescent intensity of each band observed on the gels. Figure 3 shows a comparison of the reactivity of ODN5 toward the different bases at the target site of the DNA1 or RNA1. ODN5 showed relatively high crosslinking reactivity to

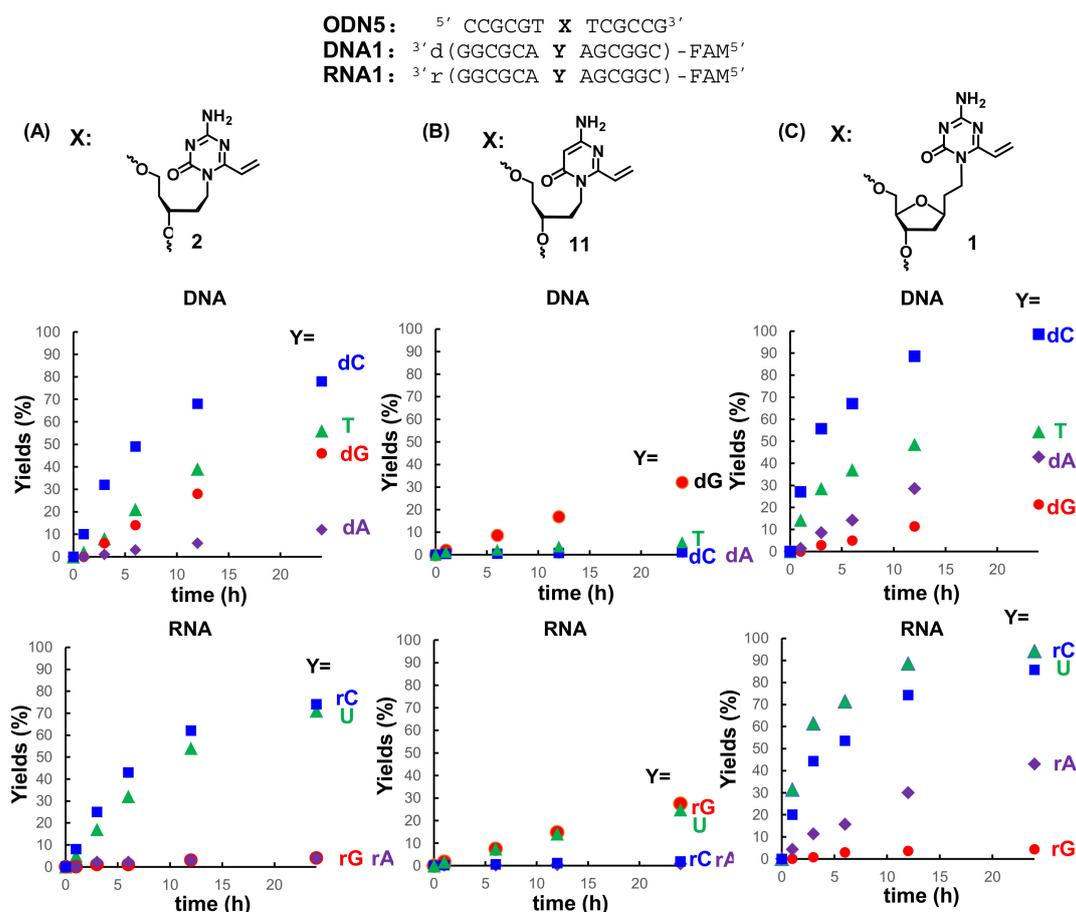
dC, dT, and dG, except for dA in the DNA. On the other hand, the crosslinked product was observed in high yields with rC and rU, and no significant products were observed with rG and rA in the RNA. The comparison of the reactivity to DNA or RNA with acyAOVT 2, 4-amino-6-oxo-2-vinyl pyrimidine having an acyclic linker (11: acyAOVP) and Et-AOVT 1, is shown in Figure 4. By comparison, acyAOVT 2 and acyAOVP 11 showed a significant difference in the reaction rate and the base selectivity to the target DNA and RNA (Figure 4A,B).



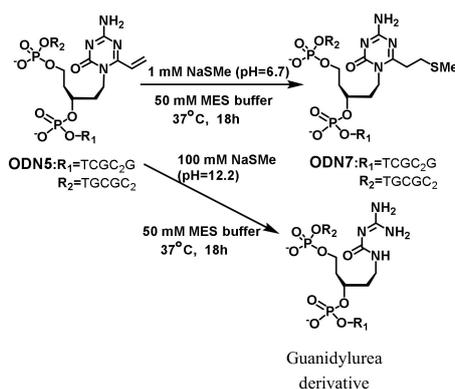
**Figure 3.** Denaturing gel electrophoresis of crosslink reaction to target DNA1 and RNA1 with ODN5. The reaction was performed in 50 mM MES-buffer (pH 7.0) containing 100 mM NaCl, ODN5 (5  $\mu$ M) and cDNAs or cRNAs (2  $\mu$ M) at 37  $^{\circ}$ C.

The reaction rate with acyAOVT 2 was faster than that of acyAOVP 11. The higher reaction rate with acyAOVT 2 might be attributed to the electron-withdrawing effect of the triazine increasing the reactivity of the 6-vinyl group of acyAOVT 2 than that of the pyrimidine-type acyAOVP 11. Compared to the target base selectivity in DNA1 with previously reported the acyAOVP 11 [21], AOVT 1 and 2 showed a drastically increased reactivity with the cytosine. On the other hand, the crosslinking yields to RNA1 with AOVT 1 and 2 were significant higher to cytosine and uracil, and relatively lower to guanine than that of AOVP 11. The reaction rate and selectivity to the pyrimidine bases (C, T and U) with acyAOVT 2 was comparable to that of Et-AOVT 1. The reactivity to adenine (A) in DNA and RNA with 2 was significantly lower than that of Et-AOVT 1 (Figure 4A,C).

Next, we carried out a thermal denaturing study of the duplex containing the non-reactive acyAOVT. The ODN7 bearing the non-reactive acyAOVT (SMe) was prepared by the addition of NaSMe to the vinyl derivative under weak acidic conditions for avoiding the decomposition of acyAOVT to guanylurea (Scheme 3).

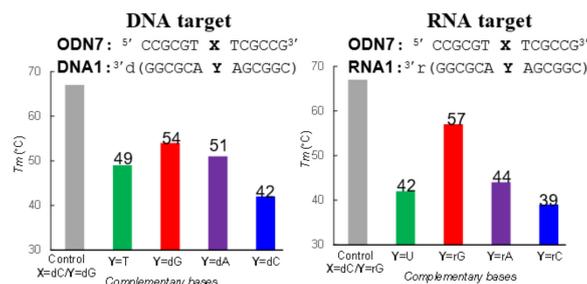


**Figure 4.** Comparison of the reaction yields calculated from the gel electrophoresis analysis of the cross-linking to target DNA1 (Y = dT, dG, dC, dA) or RNA1 (N = U, G, C, A) with *acyAOVT* (2) (A), *acyAOVP* (11) (B) and Et-AOVT (1) (C).



**Scheme 3.** Synthesis of the ODN bearing a stable precursor.

The melting temperature ( $T_m$ ) of the ODN7/DNA1 or RNA1 is summarized in Figure 5. The  $T_m$  values of the ODN7/DNA1 or RNA1 were observed to be lower than that of the unmodified natural duplex ( $T_m = 67^\circ\text{C}$ ). It should be noted that the  $T_m$  value for the duplex ODN7/DNA1 (Y = dG) or RNA1 (Y = rG) was  $54^\circ\text{C}$  and  $57^\circ\text{C}$  respectively, higher than the other duplexes. Taken together, our results of the  $T_m$  measurement suggest that AOV might form stable triple hydrogen bonds with guanine to orient the C6-vinyl group of AOV toward the opposite side of the complementary bases. Next, we attempted the crosslinking reactions to the nucleic acids-binding protein using the duplex DNA bearing AOV derivative 2.

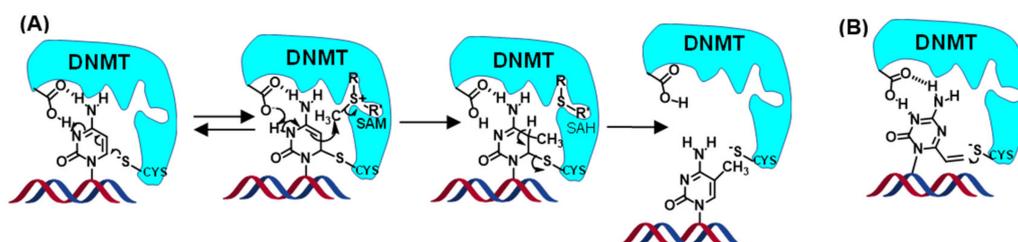


**Figure 5.** Comparative thermal denaturing analysis of duplexes ODN7/DNA1 or RNA1.  $T_m$  values were measured in 50 mM MES buffer (pH 7.0) containing 100 mM NaCl.

### 3.3. Crosslinking Reactions to DNA Methyl Transferase

We chose cytosine-5-DNA methyl transferase (DNMT) as a model DNA binding protein. DNMT performs the methylation of cytosines for epigenetic modification of the genome that is involved in regulating many cellular processes [6]. In addition, aberrant methylation can cause many diseases, such as cancer [22–24]. The efficient inhibition of DNMT offers the possibility of interfering with the methylation process, which may allow control of the epigenetic regulation of cells and treat various cancers [25].

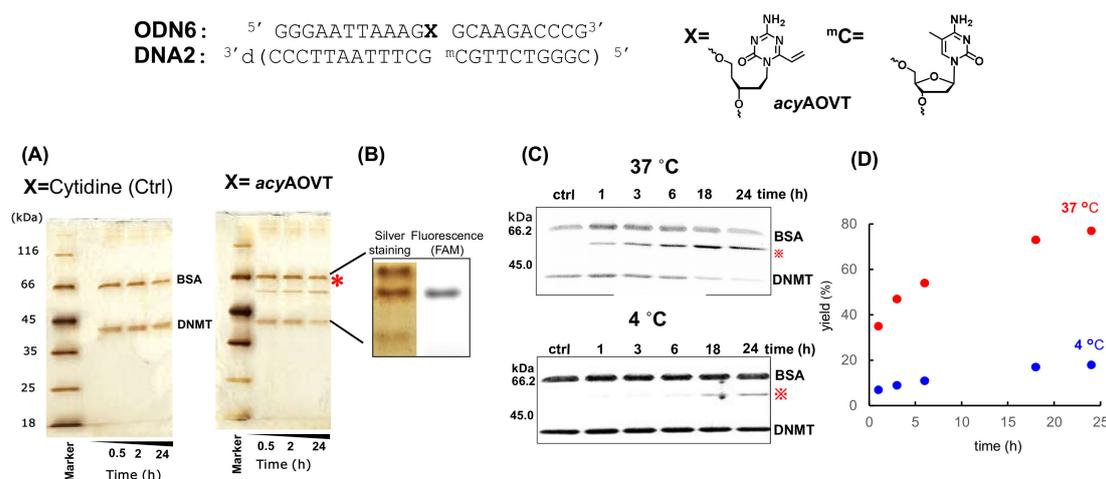
The mechanism of the cytosine-5 methylation with DNMT is illustrated in Figure 6A. A thiol of the Cys residue in the active site of DNMT acts as a nucleophile and attacks the C6 position of the cytosine to form an enzyme-linked intermediate. The resulting nucleophilic C5 position of the cytosine is then methylated by S-adenosyl-L-methionine (SAM). Subsequent abstraction of the proton at the 5 position and β-elimination provides the 5-substituted pyrimidine and active enzyme (Figure 6A) [26]. Based on this mechanism, the modified bases forming a covalent bond to DNMT were reported as irreversible inhibitors [27–31]. We expected that *acyAOVT* (2) replacement of the cytosine methylated by DNMT can react with the Cys residue in this enzyme (Figure 6B). In our investigation, we used the bacterial DNMT, i.e., the commercially-available Hha1 DNMT-1 from *Haemophilus haemolyticus*. Hha1 DNMT-1 methylated the internal 2'-deoxycytidine in the target sequence 5'-GCGC-3', in which AOV<sub>T</sub> was substituted in place of the internal 2'-dC. We synthesized ODN6, which contained AOV<sub>T</sub> in the Hha1 DNMT-1 recognition sequence and evaluated the crosslink reactivity to DNA. The base selectivity was similar, but the reactivity to the target DNA was significantly lower compared to that of the previous sequence (Figure S12).



**Figure 6.** (A) Plausible mechanism for methylation by cytosine-5 methyltransferases (DNMTs) with SAM as a cofactor (B) Possible mechanism for covalent trapping with AOV<sub>T</sub> (2) derivative.

We selected DNA2 containing methyl cytosine as a target DNA because Dnmt1 methylates hemimethylated CpG sites on one strand of double-stranded DNA. The methylation conditions were determined using the duplex DNA between DNA2 and DNA3 contained cytosine instead of acyclic AOV<sub>T</sub> and the mono-methylated product in DNA3 was confirmed by MALDI-TOF MS. The duplex DNA (ODN6 and DNA2) was incubated with Hha1 DNMT-1 in the reaction buffer containing 5 mM 2-mercaptoethanol at 37 °C in the determined conditions. The reaction mixture was analyzed by sodium dodecyl sulfate –15% polyacrylamide gel electrophoresis (SDS–PAGE) and visualized with

silver staining. The slower mobility band (\*) was observed around 51kDa and not observed with the non-reactive duplex DNA and Hha1 DNMT-1 (Figure 7A). We performed the reaction using the 5' FAM-labelled ODN6 and analyzed by SDS-PAGE. The slower mobility band stained silver matched the fluorescent band in the reaction using the 5' FAM-labelled ODN6 (Figure 7B) and this band was derived from the complex between Hha1 DNMT-1 and the duplex DNA. The quantification of the products was next performed by using the Flamingo™ Gel staining in-gel fluorescence analysis. We confirmed the linear correlation between the protein amount and the fluorescence intensity. Thus, the complex between the protein and ODN can be quantified by measuring the fluorescence intensity on the gel (Figure 7C). The reactions were performed at 37 °C and 4 °C. The yields were obtained by quantification of the fluorescence intensity for each band and reached 77% at 37 °C after 24 h.



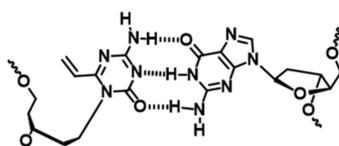
**Figure 7.** Evaluation of the crosslinking reaction between duplex DNA and Hha1 DNMT-1 (A) SDS-Page gel stained silver for analysis of reactions. Duplex DNA (ODN6 (X = C or *acyAOVT*) and DNA2: 3 μM) and Hha1 DNMT-1 (1 μM) was incubated in the presence of SAM at 4 °C. (B) 5'-FAM labelled ODN6 was used in the same reaction of (A) and the gel was visualized by silver staining and fluorescence. (C) SDS-Page gel stained Flamingo™ for analysis of reactions. Duplex DNA (ODN6 (X = C for control or *acyAOVT*) and DNA2: 1 μM) and Hha1 DNMT-1 (1 μM) was incubated in the presence of SAM at 4 °C or 37 °C. (D) Time course of the yields at 4 °C or 37 °C.

The yields significantly decreased at 4 °C (Figure 7D), suggesting that the reaction rate depends on the reaction temperature. The reactions were performed using ODN6 and DNA4 containing 4 kinds of bases (G, A, C and T) at the complementary site for *acyAOVT*. The crosslinking yields to these targets were similar to that observed with G (Figure S13). These results suggested that DNMT can bind to the mismatch sequences and make a transition to flip out of the target *acyAOVT* to induce the crosslinking reactions. Taken together, the results of the evaluation for the reaction to Hha1 DNMT-1 indicated that the *acyAOVT* in the duplex would react with Hha1 DNMT-1 even in the presence of 2-mercaptoethanol.

#### 4. Discussion

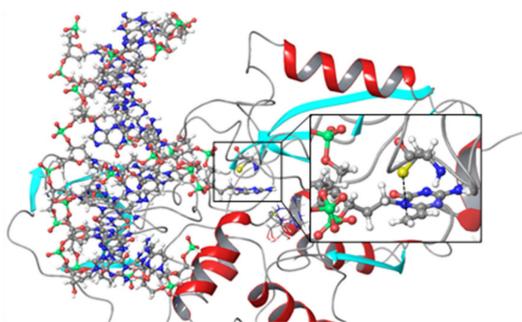
We have synthesized ODN containing the *acyAOVT* base 2 that is expected to react with the nucleophilic residue of the nucleic acid-binding protein in the proximity of the vinyl group. The reactivity with *acyAOVT* to DNA or RNA was significantly higher than that of *acyAOVP*. The interesting result that only one nitrogen substitution on AOVV drastically increased the reactivity could be attributed to the electron withdrawing effect of the nitrogen atom. The *acyAOVT* (2) showed relatively high yields to dC, dT and dG in DNA and rC, rU in RNA. The highest reactivity to cytosine with the AOVV derivatives 1 and 2 might be due to the high nucleophilicity of the amino group in cytosine and to the flexible linker, which makes it possible to access the vinyl group in the amino group of the cytosine. The cross-linking yields to adenine in the DNA or RNA with *acyAOVT* 2 were

significantly lower than that of Et-AOVT **1**. The vinyl group in *acy*AOVT **2** might be located away from the reactive site of adenine because of being positioned at similar natural base-pair by the short linker with **2**, resulting in the low crosslinking yields to adenine (Figure S14). The melting temperature ( $T_m$ ) values of the duplex for ODN7/DNA1(G) or RNA1(G) were the highest in all the duplexes containing the other AOVT-base pairing. These results suggested that the duplex containing *acy*AOVT-G would be stabilized by forming three hydrogen bonding, resulting in orientation of the vinyl group in *acy*AOVT to the opposite side of the complementary bases (Figure 8).



**Figure 8.** Hypothetical structure for *acy*AOVT and guanine in duplex DNA.

We attempted the reaction to the commercially-available DNMT using the duplex containing *acy*AOVT-G as a model reaction with the nucleic acids-binding protein. The results indicated that *acy*AOVT might react with the thiol group of the cysteine residue in DNMT in close proximity as shown in Figure 9.



**Figure 9.** Predictive structure for reaction intermediate between *acy*AOVT and cysteine in DNMT (PDB-5MHT).

We expect that this *acy*AOVT-G base pairing system would provide a “cross-linkable duplex material” to react with the nucleic acid-binding proteins.

## 5. Conclusions

We have designed *acy*AOVT **2** as a reactive base to the nucleic acid-binding protein. The ODN containing **2** exhibited a higher reactivity to the pyridine bases (C, T and U) at the complementary site of **2** in duplexes DNA-DNA and DNA-RNA than that of *acy*AOVP. The results of the  $T_m$  measurements suggested that *acy*AOVT **2** can form a stable base pair with G, and consequently, the vinyl group of **2** might be located on the opposite side of the complementary bases. The preliminary results for the reaction to DNMT with the duplex DNA containing *acy*AOVT demonstrated the potential of *acy*AOVT-G base pairing system in the reaction to the nucleic acids-binding protein. This system can be used for the crosslinking reaction to the RNA binding protein. We expect that our system will provide a useful tool for the molecular study of the RNA binding protein in control of the RNA biological functions.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/21/7709/s1>.

**Author Contributions:** K.O. and K.Y. conceived and designed the experiments. K.O. and S.I. performed the experiments. F.N. and H.O. wrote the study. F.N. supervised the project. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by a Grant-in-Aid for Scientific Research on Innovative Areas “Middle Molecular Strategy” (No. JP15H05838) from the Japan Society for the Promotion of Science (JSPS). This work was supported in part by “Dynamic Alliance for Open Innovation Bridging Human, Environment and Materials” from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

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

## References

1. Bouhlel, M.A.; Lambert, M.; David-Cordonnier, M.H. Targeting transcription factor binding to DNA by competing with DNA binders as an approach for controlling gene expression. *Curr. Top. Med. Chem.* **2015**, *15*, 1323–1358. [[CrossRef](#)] [[PubMed](#)]
2. Todeschini, A.L.; Georges, A.; Veitia, R.A. Transcription factors: Specific DNA binding and specific gene regulation. *Trends Genet.* **2014**, *30*, 211–219. [[CrossRef](#)] [[PubMed](#)]
3. O'Brien, P.J. Catalytic promiscuity and the divergent evolution of DNA repair enzymes. *Chem. Rev.* **2006**, *106*, 720–752. [[CrossRef](#)] [[PubMed](#)]
4. Nevinsky, G.A. Structural, thermodynamic, and kinetic basis for the activities of some nucleic acid repair enzymes. *J. Mol. Recognit.* **2011**, *24*, 656–677. [[CrossRef](#)]
5. Gujar, H.; Weisenberger, D.J.; Liang, G.N. The roles of human DNA methyltransferases and their isoforms in shaping the epigenome. *Genes* **2019**, *10*, 171. [[CrossRef](#)]
6. Lyko, F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* **2018**, *19*, 81–92. [[CrossRef](#)]
7. Taylor, K.; Sobczak, K. Intrinsic regulatory role of RNA structural arrangement in alternative splicing control. *Int. J. Mol. Sci.* **2020**, *21*, 5161. [[CrossRef](#)]
8. Janakiraman, H.; House, R.P.; Gangaraju, V.K.; Diehl, J.A.; Howe, P.H.; Palanisamy, V. The long (lncRNA) and short (miRNA) of it: TGF beta-mediated control of RNA-binding proteins and noncoding RNAs. *Mol. Cancer Res.* **2018**, *16*, 567–579. [[CrossRef](#)]
9. Ito, K.K.; Watanabe, K.; Kitagawa, D. The emerging role of ncRNAs and RNA-binding proteins in Mmitotic apparatus formation. *Non Coding RNA* **2020**, *6*, 13. [[CrossRef](#)]
10. Sternburg, E.L.; Karginov, F.V. Global approaches in studying RNA-binding protein interaction networks. *Trends Biochem. Sci.* **2020**, *45*, 593–603. [[CrossRef](#)]
11. Lercher, L.; McGouran, J.F.; Kessler, B.M.; Schofield, C.J.; Davis, B.G. DNA modification under mild conditions by Suzuki-Miyaura cross-coupling for the generation of functional probes. *Angew. Chem. Int. Ed.* **2013**, *52*, 10553–10558. [[CrossRef](#)] [[PubMed](#)]
12. Gerard-Hirne, T.; Thiebaut, F.; Sachon, E.; Desert, A.; Drujon, T.; Guerineau, V.; Michel, B.Y.; Benhida, R.; Coulon, S.; Saintome, C.; et al. Photoactivatable oligonucleotide probes to trap single-stranded DNA binding proteins: Updating the potential of 4-thiothymidine from a comparative study. *Biochimie* **2018**, *154*, 164–175. [[CrossRef](#)] [[PubMed](#)]
13. Smith, C.C.; Hollenstein, M.; Leumann, C.J. The synthesis and application of a diazirine-modified uridine analogue for investigating RNA-protein interactions. *RSC Adv.* **2014**, *4*, 48228–48235. [[CrossRef](#)]
14. Dziuba, D.; Hoffmann, J.E.; Hentze, M.W.; Schultz, C.A. Genetically Encoded Diazirine Analogue for RNA-Protein Photo-crosslinking. *Chembiochem* **2020**, *21*, 88–93. [[CrossRef](#)] [[PubMed](#)]
15. Wickramaratne, S.; Mukherjee, S.; Villalta, P.W.; Scharer, O.D.; Tretyakova, N.Y. Synthesis of sequence-specific DNA-protein conjugates via a reductive amination strategy. *Bioconj. Chem.* **2013**, *24*, 1496–1506. [[CrossRef](#)]
16. Carrette, L.L.G.; Morii, T.; Madder, A. Toxicity inspired cross-linking for probing DNA-peptide interactions. *Bioconj. Chem.* **2013**, *24*, 2008–2014. [[CrossRef](#)]
17. Dadova, J.; Orsag, P.; Pohl, R.; Brazdova, M.; Fojta, M.; Hocek, M. Vinylsulfonamide and acrylamide modification of DNA for cross-linking with proteins. *Angew. Chem. Int. Ed.* **2013**, *52*, 10515–10518. [[CrossRef](#)]
18. Olszewska, A.; Pohl, R.; Brazdova, M.; Fojta, M.; Hocek, M. Chloroacetamide-linked nucleotides and DNA for cross-linking with peptides and proteins. *Bioconj. Chem.* **2016**, *27*, 2089–2094. [[CrossRef](#)]
19. Matyasovsky, J.; Hocek, M. 2-Substituted 2'-deoxyinosine 5'-triphosphates as substrates for polymerase synthesis of minor-groove-modified DNA and effects on restriction endonuclease cleavage. *Org. Biomol. Chem.* **2020**, *18*, 255–262. [[CrossRef](#)]

20. Yamada, K.; Ishiyama, S.; Onizuka, K.; Nagatsugi, F. Synthesis and properties of cross-linkable DNA duplex using 4-amino-2-oxo-6-vinyl-1,3,5-triazine. *Tetrahedron* **2017**, *73*, 1424–1435. [[CrossRef](#)]
21. Kusano, S.; Ishiyama, S.; Lam, S.L.; Mashima, T.; Katahira, M.; Miyamoto, K.; Aida, M.; Nagatsugi, F. Crosslinking reactions of 4-amino-6-oxo-2-vinylpyrimidine with guanine derivatives and structural analysis of the adducts. *Nucl. Acids Res.* **2015**, *43*, 7717–7730. [[CrossRef](#)] [[PubMed](#)]
22. Robertson, K.D. DNA methylation and human disease. *Nat. Rev. Genet.* **2005**, *6*, 597–610. [[CrossRef](#)]
23. Jones, P.A.; Baylin, S.B. The epigenomics of cancer. *Cell* **2007**, *128*, 683–692. [[CrossRef](#)] [[PubMed](#)]
24. Pfeifer, G.P. Defining driver DNA methylation changes in human cancer. *Int. J. Mol. Sci.* **2018**, *19*, 1166. [[CrossRef](#)]
25. Zhou, Z.H.; Li, H.Q.; Liu, F. DNA Methyltransferase inhibitors and their therapeutic potential. *Curr. Top. Med. Chem.* **2018**, *18*, 2448–2457. [[CrossRef](#)] [[PubMed](#)]
26. Klimasauskas, S.; Kumar Roberts, R.J.; Cheng, X. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **1994**, *76*, 357–369. [[CrossRef](#)]
27. Shigdel, U.K.; He, C. A New 1'-methylenedisulfide deoxyribose that forms an efficient cross-link to DNA cytosine-5-methyltransferase (DNMT). *J. Am. Chem. Soc.* **2008**, *130*, 17634–17635. [[CrossRef](#)]
28. Sato, K.; Kawamoto, K.; Shimamura, S.; Ichikawa, S.; Matsuda, A. An oligodeoxyribonucleotide containing 5-formyl-2'-deoxycytidine (fC) at the CpG site forms a covalent complex with DNA cytosine-5 methyltransferases (DNMTs). *Bioorg. Med. Chem. Lett.* **2016**, *26*, 5395–5398. [[CrossRef](#)]
29. Sato, K.; Kunitomo, Y.; Kasai, Y.; Utsumi, S.; Suetake, I.; Tajima, S.; Ichikawa, S.; Matsuda, A. Mechanism-based inhibitor of DNA cytosine-5-methyltransferase by a SNAr reaction with an oligodeoxyribonucleotide containing a 2-amino-4-halopyridine-C-nucleoside. *Chembiochem* **2018**, *19*, 865–872. [[CrossRef](#)]
30. Utsumi, S.; Sato, K.; Ichikawa, S. Insight into the recognition mechanism of DNA cytosine-5 methyltransferases (DNMTs) by incorporation of acyclic 5-fluorocytosine (C-F) nucleosides into DNA. *Bioorg. Med. Chem. Lett.* **2018**, *28*, 2189–2194. [[CrossRef](#)]
31. Wildenhof, T.M.; Schiffers, S.; Traube, F.R.; Mayer, P.; Carell, T. Influencing epigenetic information with a hydrolytically stable carbocyclic 5-aza-2'-deoxycytidine. *Angew. Chem. Int. Ed.* **2019**, *58*, 12984–12987. [[CrossRef](#)]

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

# Encapsulation of mRNA into Artificial Viral Capsids via Hybridization of a $\beta$ -Annulus-dT<sub>20</sub> Conjugate and the Poly(A) Tail of mRNA

Yoko Nakamura <sup>1</sup>, Yuki Sato <sup>1</sup>, Hiroshi Inaba <sup>1,2</sup>, Takashi Iwasaki <sup>3</sup> and Kazunori Matsuura <sup>1,2,\*</sup> 

<sup>1</sup> Department of Chemistry and Biotechnology, Graduate School of Engineering, Tottori University, Tottori 680-8552, Japan; yotasa0926@gmail.com (Y.N.); yuki.s09hr@gmail.com (Y.S.); hinaba@tottori-u.ac.jp (H.I.)

<sup>2</sup> Centre for Research on Green Sustainable Chemistry, Tottori University, Tottori 680-8552, Japan

<sup>3</sup> Department of Bioresources Science, Graduate School of Agricultural Sciences, Tottori University, Tottori 680-8553, Japan; itaka@tottori-u.ac.jp

\* Correspondence: ma2ra-k@tottori-u.ac.jp; Tel.: +81-857-31-5262

Received: 18 September 2020; Accepted: 9 November 2020; Published: 12 November 2020



**Abstract:** Messenger RNA (mRNA) drugs have attracted considerable attention as promising tools with many therapeutic applications. The efficient delivery of mRNA drugs using non-viral materials is currently being explored. We demonstrate a novel concept where mCherry mRNA bearing a poly(A) tail is encapsulated into capsids co-assembled from viral  $\beta$ -annulus peptides bearing a 20-mer oligothymine (dT<sub>20</sub>) at the N-terminus and unmodified peptides via hybridization of dT<sub>20</sub> and poly(A). Dynamic light scattering measurements and transmission electron microscopy images of the mRNA-encapsulated capsids show the formation of spherical assemblies of approximately 50 nm. The encapsulated mRNA shows remarkable ribonuclease resistance. Further, modification by a cell-penetrating peptide (His16) on the capsid enables the intracellular expression of mCherry of encapsulated mRNA.

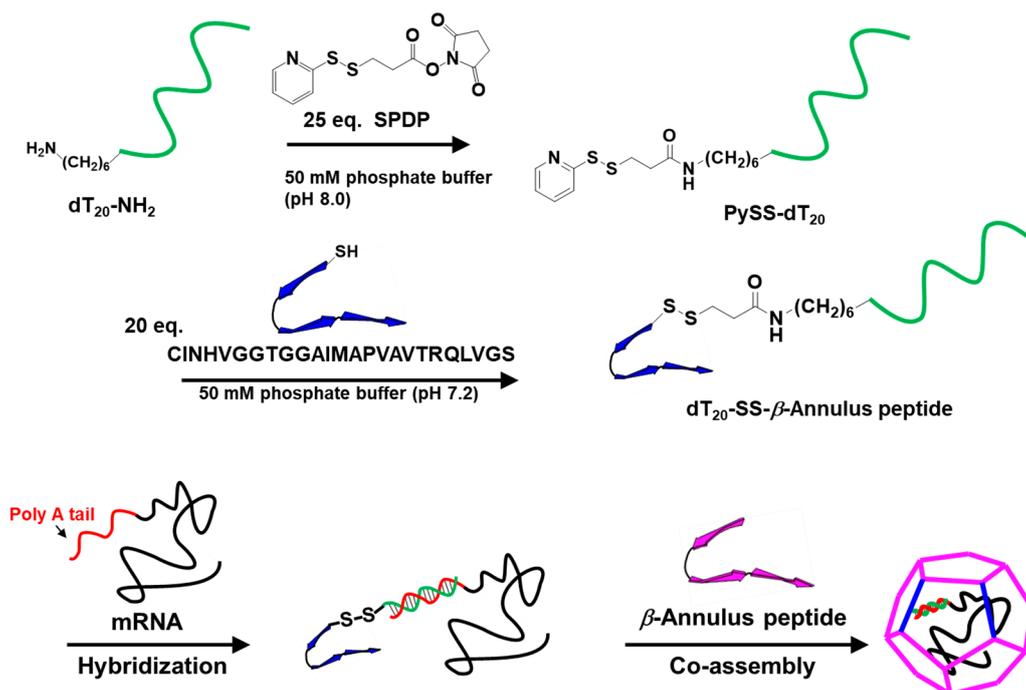
**Keywords:** mRNA; poly(A) tail; artificial viral capsid; encapsulation; nanocapsule; self-assembly;  $\beta$ -annulus peptide; peptide-DNA conjugate

## 1. Introduction

Therapeutic mRNA have attracted attention in recent years as a new type of nucleic acid drugs. Such mRNAs show potential for use in protein replacement therapy and vaccination [1–6]. mRNA drugs have the advantage of inducing direct protein synthesis in the cytoplasm, thus differing from DNA therapy. Naked mRNA is, however, unstable outside of cells and is unable to effectively penetrate cell membranes due to electrostatic repulsion. Thus, various materials for mRNA delivery, such as liposomes, dendrimers, and polyion complex micelles, have been developed and are attractive for their flexibility in molecular design [2–11]. Conversely, RNA viruses package specific genome molecules inside their outer protein shell, the viral capsid [12,13]. Viral capsids are also attractive for precise mRNA packaging and delivery but the use of natural viruses for mRNA delivery has safety concerns. Virus-like artificial protein cages for RNA packaging are, therefore, attractive. Recently, Hilvert and coworkers demonstrated that a non-viral protein cage formed by *Aquifex aeolicus* lumazine synthase selectively packages mRNA via cationic peptide tags [14,15].

Eukaryotic mRNA possesses a long adenine nucleotide (poly(A)) tail of about 100–200 nt at the 3'-end to regulate translation [16–21]. We propose a novel strategy for the encapsulation of mRNA into virus-like peptide nanocapsules via hybridization between dT<sub>20</sub> and the poly(A) tails. We developed an “artificial viral capsid” with a size of 30–50 nm self-assembled from the  $\beta$ -annulus

peptide (INHVGTTGGAIMAPVAVTRQLVGS) that participates in the formation of the dodecahedral internal skeleton of the tomato bushy stunt virus [22–35]. Artificial viral capsids encapsulate anionic guest molecules and His-tagged green fluorescence protein (GFP) into the cationic and Ni-NTA (Ni-nitrilotriacetic acid complex) modified interior [25–28]. Modification of the C-terminal, expected to be directed to the exterior of capsids, enabled the surface modification of artificial viral capsids with gold nanoparticles, coiled-coil peptides, single-stranded DNAs, and proteins [29–34]. Recently, we demonstrated that  $\beta$ -annulus peptide possessing ssDNA at the N-terminal self-assembled into ssDNA-encapsulated artificial viral capsids [35]. Thus, we designed a  $\beta$ -annulus peptide with dT<sub>20</sub> at the N-terminal to direct mRNA to the interior of capsids via hybridization with poly(A) tails (Figure 1).



**Figure 1.** Schematic of synthesis of dT<sub>20</sub> modified  $\beta$ -annulus peptide (dT<sub>20</sub>-SS- $\beta$ -annulus) and formation of an artificial viral capsid by co-assembly of  $\beta$ -annulus peptide and dT<sub>20</sub>-SS- $\beta$ -annulus hybridized with mRNA.

## 2. Materials and Methods

### 2.1. General

Reverse-phase HPLC was performed at ambient temperature with a Shimadzu LC-6AD liquid chromatograph equipped with a UV-Vis detector (220 nm and 260 nm, Shimadzu SPD-10AVvp, Kyoto, Japan) using an Inertsil WP300 C18 column (GL Science, 250 mm  $\times$  4.6 mm or 250  $\times$  20 mm). MALDI-TOF mass spectra were obtained on an Autoflex T2 (Bruker Daltonics, Billerica, USA) in linear/positive mode with  $\alpha$ -cyano-4-hydroxycinnamic acid ( $\alpha$ -CHCA) or 3-hydroxypicolinic acid (3-HPA) with diammonium hydrogen citrate as the matrix. The UV-Vis spectra of DNA-conjugated peptides were measured at 260 nm using a Jasco V-630 (JASCO Corporation, Tokyo, Japan) with a quartz cell (S10-UV-1, GL Science). An mRNA coding mCherry fluorescent protein (mCherry mRNA, ~1 kb) with a 3'-polyadenylic acid (poly(A)) tail was purchased from OZ Biosciences (Marseille, France). RNase A was purchased from Nacalai Tesque (Kyoto, Japan). All other reagents were obtained from a commercial source and used without further purification. Deionized water of high resistivity (>18 M $\Omega$  cm) was purified using a Millipore Purification System (Milli-Q water, Merck Millipore, Burlington, USA) and used as a solvent.

## 2.2. Preparation of dT<sub>20</sub>-SS-β-Annulus and β-Annulus Peptides

β-annulus (INHVGTTGGAIMAPVAVTRQLVGS) and Cys-β-annulus (CINHVGTTGGAIMA PVAVTRQLVGS) peptides were synthesized with a Biotage Initiator<sup>+</sup> (Biotage, Uppsala, Kingdom of Sweden) using standard Fmoc-based coupling chemistry as previously described [22,28]. MALDI-TOF-MS of the β-annulus peptide (matrix: α-CHCA):  $m/z = 2306$  (exact mass: 2306) and of the Cys-β-annulus peptide (matrix: α-CHCA):  $m/z = 2409$  (exact mass: 2409).

An amine-modified dT<sub>20</sub> (5′NH<sub>2</sub>-(CH<sub>2</sub>)<sub>6</sub>-TTTTTTTTTTTTTTTTTTTT-3′, Gene Design Inc., Osaka, Japan) in 50 mM sodium phosphate buffer (pH 8.0) was added to 20-fold molar excess of *N*-succinimidyl 3-(2-pyridyldithio)propionate (SPDP) in acetonitrile and the mixture was incubated for 1 h at 25 °C. A Spectra/por7 membrane with a cutoff Mw 1000 (Spectrum Laboratories, Inc., Rancho Dominguez, USA) was used for dialysis against water for 24 h. The internal solution was lyophilized to afford a flocculent solid. This product was dissolved in water and the concentration was defined by UV-Vis spectroscopy. Product (PySS-dT<sub>20</sub>) yield was 43.6 nmol (94.8%)—MALDI-TOF-MS (matrix: 3-HPA):  $m/z = 6409$  (exact mass: 6398).

An aqueous solution of PySS-dT<sub>20</sub> (0.1 mM) in a 50 mM sodium phosphate buffer (pH 7.2) was mixed with the Cys-β-annulus peptide (2 mM) in a 50 mM sodium phosphate buffer (pH 7.2) and the mixture was incubated for 24 h at 25 °C. The solution was purified by reverse-phase HPLC eluted with a linear gradient of CH<sub>3</sub>CN/0.1 M ammonium formate aqueous solution (10/90 to 100/0 over 95 min). The elution fraction was collected, concentrated in a centrifugal evaporator, and dialyzed as above against water for 20 h. The internal solution was lyophilized to afford a flocculent solid. This product was dissolved in water and the concentration was defined by UV-Vis spectroscopy. Product (dT<sub>20</sub>-SS-β-annulus peptide) yield was 12.6 nmol (27.4%)—MALDI-TOF-MS (matrix: 3-HPA):  $m/z = 8710$  (exact mass: 8698).

## 2.3. Preparation and Characterization of Artificial Viral Capsids

Lyophilized dT<sub>20</sub>-SS-β-annulus and β-annulus peptides were dissolved in water, respectively. Peptide solutions were mixed at molar ratios of dT<sub>20</sub>-SS-β-annulus:β-annulus = 1:9, 1:4.5, 1:2, and 1:1, sonicated for 5 min, and lyophilized. Stock solutions of peptides were prepared by dissolving peptide powders in 1× phosphate buffered saline (PBS, pH 7.4) and sonicating for 3 min.

Dynamic light scattering (DLS) was measured with a Zetasizer NanoZS (Malvern, Kobe, Japan) instrument at 25 °C using an incident He-Ne laser (633 nm). Correlation times of scattered light intensities  $G(\tau)$  were measured several times and the means were fitted to Equation (1), where  $B$  is the baseline,  $A$  is amplitude,  $q$  is the scattering vector,  $\tau$  is the delay time, and  $D$  is the diffusion coefficient.

$$G(\tau) = B + A \exp(-2q^2D\tau) \quad (1)$$

Hydrodynamic radii ( $R_H$ ) of scattering particles were calculated using the Stokes-Einstein Equation (2), where  $\eta$  is the solvent viscosity,  $k_B$  is Boltzmann's constant, and  $T$  denotes absolute temperature.

$$R_H = k_B T / 6\pi\eta D \quad (2)$$

Transmission electron microscopy (TEM) images were obtained with a JEOL JEM 1400 Plus (JEOL Ltd., Tokyo, Japan), using an accelerating voltage of 80 kV. Aliquots (5 μL) of DLS samples were applied to hydrophilized carbon-coated Cu-grids (C-SMART Hydrophilic TEM grids, Alliance Biosystems, Osaka, Japan) for 60 s and then removed. Subsequently, the TEM grid was instilled in the aqueous solution (5 μL) of 2% phosphotungstic acid, Na<sub>3</sub>(PW<sub>12</sub>O<sub>40</sub>)(H<sub>2</sub>O)<sub>n</sub>. The staining solution was removed after 60 s and the sample-loaded grids were dried *in vacuo*.

#### 2.4. Preparation of Complex of dT<sub>20</sub>-SS-β-Annulus:β-Annulus Peptide Hybridized with mCherry mRNA

An aqueous solution of mCherry mRNA in 1× PBS (1 mg/mL, (nucleotide) ≈ 3 mM, and pH 7.4) was mixed with dT<sub>20</sub>-SS-β-annulus:β-annulus peptide powders by gentle pipetting without sonication. The peptide solutions were diluted in 1× PBS (pH 7.4) so that the nucleotide concentration of mCherry mRNA was equal to dT<sub>20</sub> ((mRNA nt) = (T) = 1 mM, (dT<sub>20</sub>) = 50 μM, (RNA) = 1 μM). In a typical experiment, the dT<sub>20</sub>-SS-β-annulus:β-annulus/mRNA solution was incubated for 30 min at 25 °C.

#### 2.5. Electrophoretic Mobility Shift Assay and Nuclease Resistance Assay

Electrophoretic mobility shift assay (EMSA) was employed for the detection of peptide complexes with nucleic acids. The dT<sub>20</sub>-SS-β-annulus:β-annulus/mRNA solutions were loaded onto 3% *w/v* agarose gels in a TAE buffer, pre-cast with GelRed (Wako Pure Chemical Industries, Osaka, Japan) for nucleic acid detection. Three microliters of the dT<sub>20</sub>-SS-β-annulus:β-annulus/mRNA solution ((mRNA nt) = (T) = 1 mM) was loaded with 5 μL of Bluejuice Gel loading buffer (Thermo Fisher, Waltham, USA). Electrophoresis used 210 V for 30 min on an Atto AE-6100 with Atto mypower II300 AE-8130 (Atto, Tokyo, Japan). The bands were visualized with a UV illuminator (TP-15 MP, Atto, Tokyo, Japan) and the images were recorded with a digital camera.

Ribonuclease A from bovine pancreas (Nacalai Tesque, Kyoto, Japan) is an endoribonuclease which specifically hydrolyzes the 3' end of pyrimidine residues in single-stranded RNA. Each peptide/mRNA solution was incubated with a solution of RNase A (3 μL, 1 U/μL) in PBS for 10–60 min at 37 °C. Enzymatic degradation of mRNA was evaluated by EMSA.

#### 2.6. Confocal Laser Scanning Microscopy (CLSM) Measurements of In-Cell Expression of mCherry mRNA

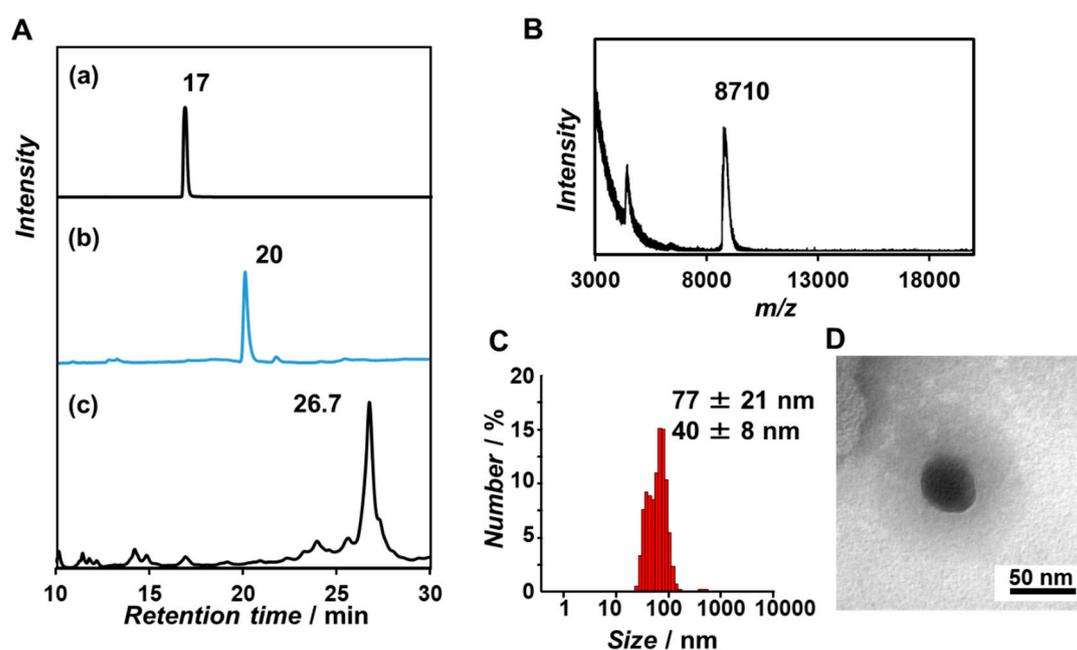
Confocal laser scanning microscopy (CLSM) used a Fluo View FV10i (Olympus, Tokyo, Japan). Human hepatoma HepG2 cells were cultured in DMEM (10 *v/v*% FBS, 100 μg/mL streptomycin, 100 units/mL penicillin, 1 mM sodium pyruvate, and 1 *v/v*% MEM nonessential amino acids) for 24 h at 37 °C in a 5% CO<sub>2</sub> atmosphere. The cells were seeded onto single-well bottom dishes with 2.0 × 10<sup>4</sup> cells/well in a final volume of 100 μL and incubated for 24 h at 37 °C and 5% CO<sub>2</sub>. The medium was removed and 50 μM dT<sub>20</sub>-SS-β-annulus, 450 μM β-annulus, and 1 μM mCherry mRNA in the fresh medium were added to the cells, then incubated for 48 h under 5% CO<sub>2</sub>. The co-assembly mixture of 50 μM dT<sub>20</sub>-SS-β-annulus, 450 μM β-annulus-His<sub>16</sub>, and 1 μM mCherry mRNA was added to the cells for evaluation of expression of mCherry fluorescent protein. A complex of 2 μM TransIT-mRNA (Transfection kit, Mirus Bio LLC, Madison, USA) with 1 μM mCherry mRNA was added to cells as a positive control. CLSM images of mCherry fluorescence were obtained with excitation at 587 nm and an mCherry band-pass filter (Red). Fluorescence intensity was measured by cell from CLSM images by subtracting background intensity using ImageJ 1.51 software.

### 3. Results and Discussion

#### 3.1. Synthesis and Self-Assembling Behavior of β-Annulus Peptide Bearing dT<sub>20</sub> at the N-Terminus

We designed a β-annulus peptide bearing dT<sub>20</sub> at the N-terminus that was directed to the interior of the capsid (Figure 1) to encapsulate mRNA via hybridization between dT<sub>20</sub> and the poly(A) tail at the 3' end of mRNA. A 5'-terminal aminated 20-mer thymidine nucleotide (dT<sub>20</sub>-NH<sub>2</sub>) was reacted with *N*-succinimidyl 3-(2-pyridyldithio)propionate (SPDP) to obtain a dT<sub>20</sub> bearing pyridyl disulfide group at the 5'-end (PySS-dT<sub>20</sub>). A Cys-β-annulus peptide (CINHVGTTGGAIMAPVAVTRQLVGS) containing Cys at the N-terminus was synthesized by standard Fmoc-based solid-phase methods according to our reported procedure [28]. PySS-dT<sub>20</sub> was attached to the Cys-β-annulus peptide with a disulfide-exchange reaction. A reversed-phase HPLC chart of the reaction mixture showed one peak at 26.7 min, which was different from the retention time of dT<sub>20</sub>-NH<sub>2</sub> and PySS-dT<sub>20</sub> (Figure 2A). The purified product was assigned by MALDI-TOF-MS as dT<sub>20</sub>-modified β-annulus peptide (dT<sub>20</sub>-SS-β-Annulus) (Figure 2B). DLS of a 50 μM solution of dT<sub>20</sub>-SS-β-annulus in PBS (pH 7.4) exhibited

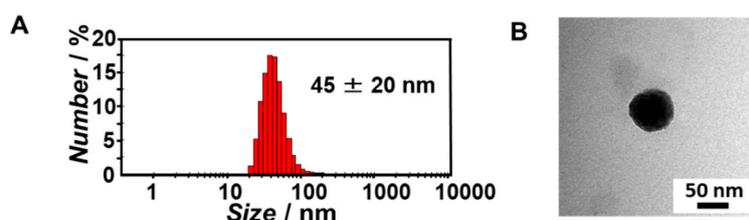
formation of  $77 \pm 21$  and  $40 \pm 8$  nm assemblies (Figure 2C). TEM images of the aqueous solution stained with phosphotungstic acid showed the formation of spherical assemblies of approximately 40 nm in diameter (Figure 2D). The concentration dependence of the size distribution in DLS measurement indicated that dT<sub>20</sub>-SS- $\beta$ -annulus formed assemblies with sizes ranging from 30 to 100 nm in a concentration range of 5–125  $\mu$ M (Figure S1). This tendency was similar to the findings in our previous report [35]. Notably,  $\beta$ -annulus peptide modified with highly charged dT<sub>20</sub> formed relatively stable artificial viral capsids.



**Figure 2.** (A) Reverse-phase HPLC chart of (a) dT<sub>20</sub>-NH<sub>2</sub>, (b) PySS-dT<sub>20</sub> detected at 260 nm, eluted with a linear gradient of CH<sub>3</sub>CN/0.1 M NH<sub>4</sub>HCO<sub>2</sub> aq (0/100 to 100/0 over 95 min), and (c) dT<sub>20</sub>-SS- $\beta$ -annulus detected at 260 nm, eluted with a linear gradient of CH<sub>3</sub>CN/0.1 M NH<sub>4</sub>HCO<sub>2</sub> aq (10/90 to 100/0 over 95 min). (B) MALDI-TOF-MS of the purified dT<sub>20</sub>-SS- $\beta$ -annulus (matrix: 3-HPA). (C,D) TEM image and size distribution obtained from dynamic light scattering (DLS) for an aqueous solution of 50  $\mu$ M dT<sub>20</sub>-SS- $\beta$ -annulus in PBS (pH 7.4) at 25 °C.

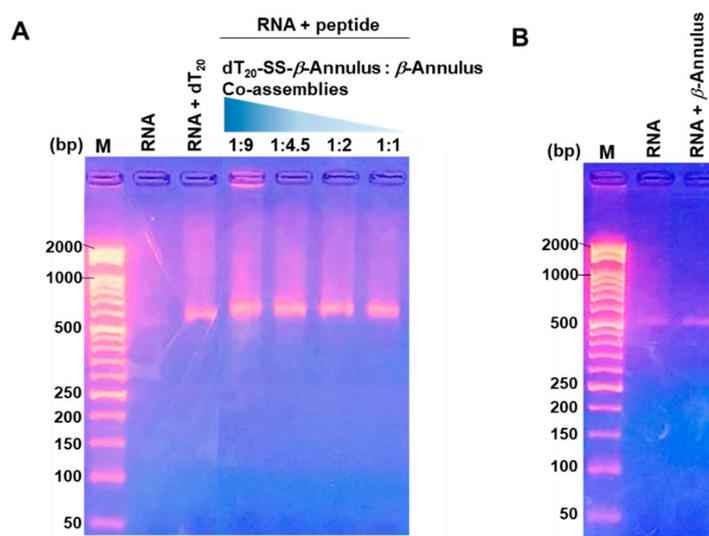
### 3.2. Complexation of mCherry mRNA and Artificial Viral Capsid Bearing dT<sub>20</sub>

We initially attempted encapsulation of mRNA within the artificial viral capsid consisting of dT<sub>20</sub>-SS- $\beta$ -annulus via hybridization between dT<sub>20</sub> and mRNA. However, when poly(A) and dT<sub>20</sub>-SS- $\beta$ -annulus were mixed at equimolar base concentrations, capsid formation was insufficient (data are not shown). It is probably difficult to self-assemble among dT<sub>20</sub>-SS- $\beta$ -annulus peptides hybridized on a large poly(A) molecule due to the excluded volume effect. Therefore, we developed an alternate strategy to encapsulate mRNA with dT<sub>20</sub>-SS- $\beta$ -annulus and unmodified  $\beta$ -annulus peptide (INHVGTTGGAIMAPVAVTRQLVGS) (Figure 1). DLS and TEM images of the mixture of dT<sub>20</sub>-SS- $\beta$ -annulus/unmodified  $\beta$ -annulus peptide at 1:9 molar ratio in PBS showed the formation of spherical co-assemblies of size  $45 \pm 20$  nm (Figure 3). The peptides mixture at other molar ratios (dT<sub>20</sub>-SS- $\beta$ -annulus: $\beta$ -annulus = 1:4.5, 1:2, 1:1) also formed spherical co-assemblies of similar size (Figure S2). Next, we encapsulated mRNA coding in the mCherry fluorescent protein (mCherry mRNA) by co-assembly at different molar ratios of dT<sub>20</sub>-modified and unmodified  $\beta$ -annulus peptide.

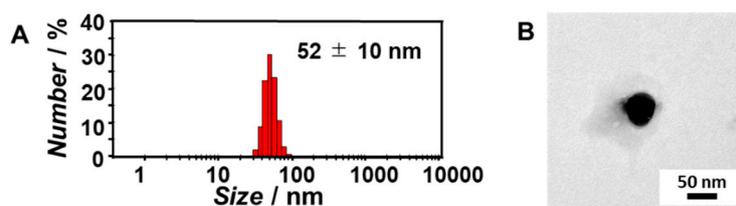


**Figure 3.** (A) Size-distributions obtained from DLS and (B) TEM images for co-assembly peptides ( $dT_{20}$ -SS- $\beta$ -annulus: $\beta$ -annulus = 1:9) in water at 25 °C.

An aqueous solution of mCherry mRNA in PBS was added to the lyophilized powder of  $dT_{20}$ -SS- $\beta$ -annulus: $\beta$ -annulus peptide followed by dilution with PBS and incubation for 30 min at 25 °C to prepare capsids containing mCherry mRNA. The electrophoretic mobility shift assay (EMSA) showed a band of mCherry mRNA significantly shifted by complexation with a co-assembly mixture of  $dT_{20}$ -SS- $\beta$ -annulus: $\beta$ -annulus peptide at a 1:9 molar ratio (Figure 4A, lane 4). In contrast, the other co-assembly mixtures showed only a slight retardation of mRNA (Figure 4A, lane 5–7), which reflect almost neutral net charge and relatively small size of the  $\beta$ -annulus peptide. The migration position of a mixed solution of unmodified  $\beta$ -annulus peptide and mCherry mRNA was not shifted from mCherry mRNA alone (Figure 4B). DLS and TEM images of the 1:9 co-assembly complexes with mCherry mRNA show the formation of spherical assemblies of size  $52 \pm 10$  nm (Figure 5A,B). We previously reported that the N-terminus of the  $\beta$ -annulus peptide is directed toward the interior of the capsid and the surface  $\zeta$ -potential of the capsid is almost zero at neutral pH [25]. Thus, when negatively charged RNA is encapsulated at neutral pH, the artificial viral capsid should minimally migrate from the applied position due to the capsid charge shield. Therefore, EMSA results indicate that the  $dT_{20}$ -SS- $\beta$ -annulus: $\beta$ -annulus peptide 1:9 co-assembly encapsulated mCherry mRNA via hybridization with  $dT_{20}$  directed to the interior of capsids. EMSA shows that mCherry mRNA could be complexed with the 1:9 co-assembly at 4 °C, 25 °C, 37 °C, and 60 °C (Figure S3). However, complexation at 4 °C and 60 °C afforded large aggregates with bimodal size distribution (Figure S4). Thus, mCherry mRNA-encapsulated artificial viral capsid with unimodal size distribution can be obtained by incubation with the 1:9 co-assembly mixture at temperatures of 25 °C–37 °C.



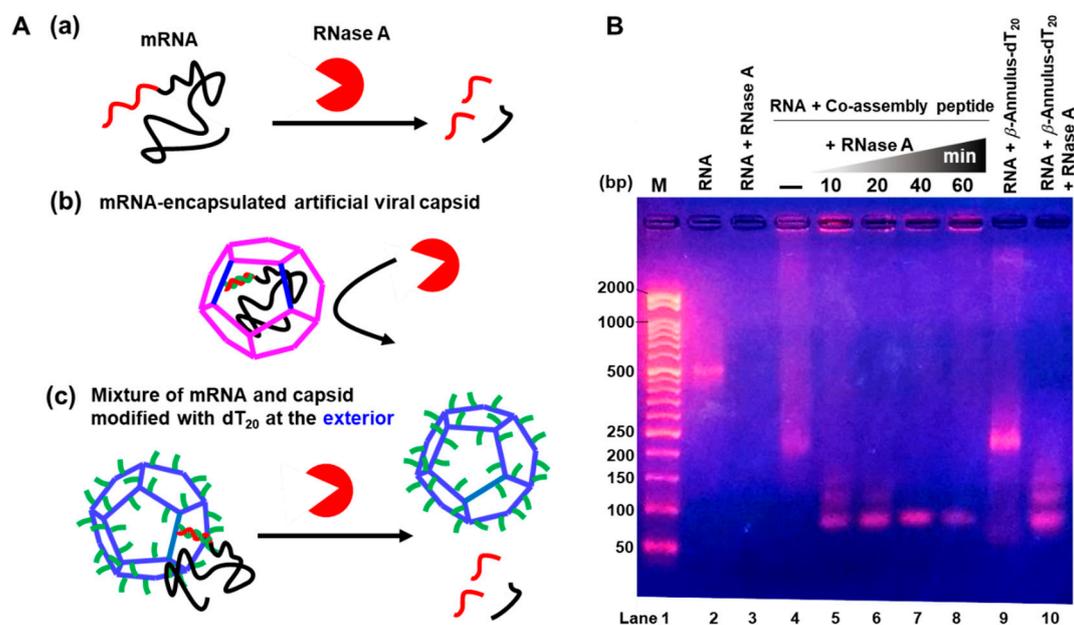
**Figure 4.** Gel shift assay of (A) co-assembly peptides hybridized with mCherry mRNA, (B) mixture of  $\beta$ -annulus peptide and mCherry mRNA in phosphate buffered saline (PBS) (pH 7.4) incubated at 25 °C for 30 min. M: DNA marker. All DNA markers used in the gel shift assay are the same. The gels were stained with GelRed.



**Figure 5.** (A) Size-distributions obtained from DLS and (B) TEM images for co-assembly peptides (dT<sub>20</sub>-SS-β-annulus:β-annulus = 1:9) and mCherry mRNA in PBS (pH 7.4) obtained after incubation at 25 °C.

### 3.3. Nuclease Resistance of mRNA-Encapsulated Artificial Viral Capsid

Resistance of mCherry mRNA-encapsulated artificial viral capsid to endoribonuclease (RNase A) was confirmed by EMSA. Naked mCherry mRNA was digested by RNase A within 10 min (Figure 6(Aa),B lane 3). Conversely, mRNA in artificial viral capsids (band near the loading well) was minimally digested after 60 min (Figure 6(Ac),B lane 5–8) due to protection by the capsid. For the artificial viral capsids modified with dT<sub>20</sub>, the exteriors were constructed as a control by self-assembly of β-annulus peptide modified with dT<sub>20</sub> at the C-terminal that directed dT<sub>20</sub> to the exterior of the capsid. The mixture of mCherry mRNA and the dT<sub>20</sub>-modified capsid at the exterior did not cause a significant mobility shift (Figure 6B lane 9), indicating that mRNA was not encapsulated. mRNA was rapidly digested by RNase A within 10 min (Figure 6(Ac),B lane 10). Thus, mRNA is protected from degradation by artificial viral capsids.

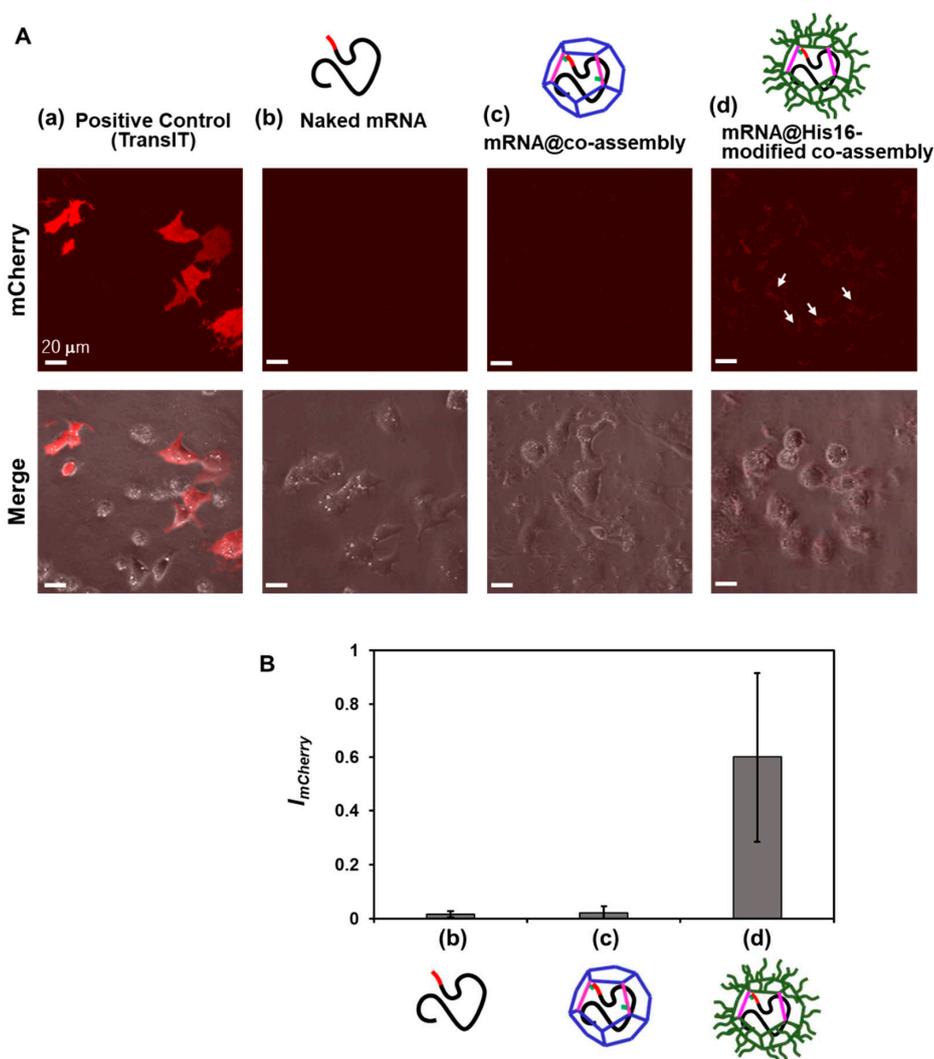


**Figure 6.** (A) Illustration of digestion of mRNA by RNase A: (a) mRNA alone, (b) mRNA-encapsulated artificial viral capsid, and (c) Mixture of mRNA and capsid modified with dT<sub>20</sub> at the exterior. (B) Gel electrophoresis assay of digestions mRNA by RNase A at 37 °C for 10 min (mRNA alone and β-annulus-dT<sub>20</sub>) and 10–60 min (dT<sub>20</sub>-SS-β-annulus:β-annulus = 1:9 co-assembly). The gel was stained with GelRed.

### 3.4. In-Cell Expression of mCherry mRNA Encapsulated in Artificial Viral Capsid

Uptake of mCherry mRNA-encapsulated capsids into human hepatoma HepG2 cells were analyzed by CLSM to evaluate mCherry mRNA expression inside cells. mCherry mRNA and a strong transfection reagent (TransIT-mRNA) were incubated with the cells for 48 h at 37 °C as a positive control; significant red fluorescence of mCherry was observed (Figure 7(Aa)). Unfortunately,

minimal expression of encapsulated mCherry mRNA was observed as similar to the naked mCherry mRNA (Figure 7A(b,c)). We employed a histidine 16-mer (His16), a cell-penetrating peptide uncharged under physiological conditions [36,37] to enhance cell permeability of capsids. We constructed artificial viral capsids modified with His16 on the exterior and using the 1:9 co-assembly of dT<sub>20</sub>-SS- $\beta$ -annulus peptide and  $\beta$ -annulus peptide modified with His16 at the C-terminal ( $\beta$ -annulus-His16). Significant red fluorescence of mCherry was observed in cells incubated with His16-modified capsids (Figure 7(Ad)). The expression level of mCherry was quantified by fluorescence intensity in the cell areas of the CLSM image using the software ImageJ. The expression level using His16-modified capsids was 26.5 times the expression observed with unmodified capsids (Figure 7B). This indicates that the intracellular expression of mCherry was significantly enhanced by the modification of His16 on the surface of the artificial viral capsid. We have demonstrated in our previous work that the reductive cleavage of the disulfide bond of the artificial viral capsid by dithiothreitol (DTT) caused the controlled-release of ssDNA with the destruction of the capsid [35]. mCherry mRNA was likely released from capsids in the reducing environment in cells.



**Figure 7.** (A) Confocal laser scanning microscopy (CLSM) images of HepG2 Cells transfected with mCherry mRNA encapsulated in artificial viral capsids: (a) positive control (2  $\mu$ M TransIT-mRNA with 1  $\mu$ M mCherry mRNA), (b) 1  $\mu$ M Naked mCherry mRNA, (c) mRNA-encapsulated artificial viral capsid (50  $\mu$ M dT<sub>20</sub>-SS- $\beta$ -annulus, 450  $\mu$ M  $\beta$ -annulus, and 1  $\mu$ M mCherry mRNA), and (d) mRNA-encapsulated histidine 16-mer (His16)-artificial viral capsid (50  $\mu$ M dT<sub>20</sub>-SS- $\beta$ -annulus, 450  $\mu$ M  $\beta$ -annulus-His16, and 1  $\mu$ M mCherry mRNA). (B) Relative fluorescence intensity of mCherry from HepG2 Cells.

#### 4. Conclusions

We constructed artificial viral capsids co-assembled from viral  $\beta$ -annulus peptides bearing dT<sub>20</sub> at the N-terminus and unmodified peptides. Capsids bearing interior dT<sub>20</sub> enable encapsulation of mCherry mRNA bearing poly(A) tails via hybridization to form ribonuclease resistant spherical assemblies of approximately 50 nm in diameter. Further, we demonstrate novel material for mRNA delivery by His16-modification on the exterior of capsids. Further research will demonstrate potential applications for artificial viral capsids as delivery systems for nucleic acid drugs.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/22/8004/s1>, Figure S1: Concentration dependence of size distribution obtained from DLS for the aqueous solution of dT<sub>20</sub>-SS- $\beta$ -annulus in PBS (pH7.4) at 25 °C, Figure S2: Size-distributions obtained from DLS and TEM images for co-assembly peptides (dT<sub>20</sub>-SS- $\beta$ -annulus: $\beta$ -annulus = 1:4.5 (a), 1:2 (b), and 1:1 (c) in water at 25 °C), Figure S3: Gel shift assay obtained after hybridization of co-assembly peptides and mCherry mRNA in PBS (pH 7.4) incubated at various temperatures for 10–30 min, and Figure S4: Size-distributions obtained from DLS and TEM images for co-assembly peptides (dT<sub>20</sub>-SS- $\beta$ -annulus: $\beta$ -annulus = 1:9) and mCherry mRNA in PBS buffer (pH 7.4) obtained after incubation at 4 °C (a), 37 °C (b), and 60 °C (c) for 10–30 min.

**Author Contributions:** Conceptualization, K.M.; methodology, T.I. and H.I.; investigation, Y.N. and Y.S.; writing, Y.N. and K.M.; visualization, Y.S. and H.I.; supervision, K.M.; funding acquisition, K.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by a Grant-in-Aid for Scientific Research on Innovative Areas “Chemistry for Multimolecular Crowding Biosystems” (grant number: JP18H04558) and a Grant-in-Aid for Scientific Research (B) (grant number: JP18H02089).

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

#### References

- Sahin, U.; Karikó, K.; Türeci, Ö. mRNA-based therapeutics—Developing a new class of drugs. *Nat. Rev. Drug Discov.* **2014**, *13*, 759–780. [[CrossRef](#)] [[PubMed](#)]
- Yin, H.; Kanasty, R.L.; Eltoukhy, A.A.; Vegas, A.J.; Dorkin, J.R.; Anderson, D.G. Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* **2014**, *15*, 541–555. [[CrossRef](#)] [[PubMed](#)]
- Kauffman, K.J.; Webber, M.J.; Anderson, D.G. Materials for non-viral intracellular delivery of messenger RNA therapeutics. *J. Control. Release* **2016**, *240*, 227–234. [[CrossRef](#)] [[PubMed](#)]
- Guan, S.; Rosenecker, J. Nanotechnologies in delivery of mRNA therapeutics using non-viral vector-based delivery systems. *Gene Ther.* **2017**, *24*, 133–143. [[CrossRef](#)] [[PubMed](#)]
- Kowalski, P.S.; Rudra, A.; Miao, L.; Anderson, D.G. Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery. *Mol. Ther.* **2019**, *27*, 710–728. [[CrossRef](#)]
- Meng, C.; Chen, Z.; Li, G.; Welte, T.; Shen, H. Nanoplatfoms for mRNA Therapeutics. *Adv. Ther.* **2020**. [[CrossRef](#)]
- Uchida, H.; Itaka, K.; Nomoto, T.; Ishii, T.; Suma, T.; Ikegami, M.; Miyata, K.; Oba, M.; Nishiyama, N.; Kataoka, K. Modulated Protonation of Side Chain Aminoethylene Repeats in N-Substituted Polyaspartamides Promotes mRNA Transfection. *J. Am. Chem. Soc.* **2014**, *136*, 12396–12405. [[CrossRef](#)]
- Baba, M.; Itaka, K.; Kondo, K.; Yamasoba, T.; Kataoka, K. Treatment of neurological disorders by introducing mRNA In Vivo using polyplex nanomicelles. *J. Control. Release* **2015**, *201*, 41–48. [[CrossRef](#)]
- Matsui, A.; Uchida, S.; Ishii, T.; Itaka, K.; Kataoka, K. Messenger RNA-based therapeutics for the treatment of apoptosis-associated diseases. *Sci. Rep.* **2015**, *5*, 15810. [[CrossRef](#)]
- Li, J.; Wang, W.; He, Y.; Li, Y.; Yan, E.Z.; Zhang, K.; Irvine, D.J.; Hammond, P.T. Structurally Programmed Assembly of Translation Initiation Nanoplex for Superior mRNA Delivery. *ACS Nano* **2017**, *11*, 2531–2544. [[CrossRef](#)]
- Benner, N.L.; McClellan, R.L.; Turlington, C.R.; Haabeth, O.A.W.; Waymouth, R.M.; Wender, P.A. Oligo(serine ester) Charge-Altering Releasable Transporters: Organocatalytic Ring-Opening Polymerization and their Use for In Vitro and In Vivo mRNA Delivery. *J. Am. Chem. Soc.* **2019**, *141*, 8416–8421. [[CrossRef](#)] [[PubMed](#)]
- Rein, A. Retroviral RNA packaging: A review. In *Positive-Strand RNA Viruses*; Brinton, M.A., Calisher, C.H., Rueckert, R., Eds.; Springer: Vienna, Austria, 1994; Volume 9, pp. 513–522.
- Perlmutter, J.D.; Hagan, M.F. Mechanisms of Virus Assembly. *Annu. Rev. Phys. Chem.* **2015**, *66*, 217–239. [[CrossRef](#)] [[PubMed](#)]
- Azuma, Y.; Edwardson, T.G.W.; Terasaka, N.; Hilvert, D. Modular Protein Cages for Size-Selective RNA Packaging in Vivo. *J. Am. Chem. Soc.* **2018**, *140*, 566–569. [[CrossRef](#)] [[PubMed](#)]

15. Terasaka, N.; Azuma, Y.; Hilvert, D. Laboratory evolution of virus-like nucleocapsids from non-viral protein cages. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 5432–5437. [[CrossRef](#)]
16. Munroe, D.; Jacobson, A. Tales of poly(A): A review. *Gene* **1990**, *91*, 151–158. [[CrossRef](#)]
17. Sachs, A. The role of poly(A) in the translation and stability of mRNA. *Curr. Opin. Cell Biol.* **1990**, *2*, 1092–1098. [[CrossRef](#)]
18. Eckmann, C.R.; Rammelt, C.; Wahle, E. Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 348–361. [[CrossRef](#)]
19. Weill, L.; Belloc, E.; Bava, F.-A.; Méndez, R. Translational control by changes in poly(A) tail length: Recycling mRNAs. *Nat. Struct. Mol. Biol.* **2012**, *19*, 577–585. [[CrossRef](#)]
20. Gruber, A.J.; Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* **2019**, *20*, 599–614. [[CrossRef](#)]
21. Nicholson, A.L.; Pasquinelli, A.E. Tales of Detailed Poly(A) Tails. *Trends Cell Biol.* **2019**, *29*, 191–200. [[CrossRef](#)]
22. Matsuura, K.; Watanabe, K.; Matsuzaki, T.; Sakurai, K.; Kimizuka, N. Self-Assembled Synthetic Viral Capsids from a 24-mer Viral Peptide Fragment. *Angew. Chem. Int. Ed.* **2010**, *49*, 9662–9665. [[CrossRef](#)]
23. Matsuura, K. Rational design of self-assembled proteins and peptides for nano and micro-sized architectures. *RSC Adv.* **2014**, *4*, 2942–2953. [[CrossRef](#)]
24. Matsuura, K. Synthetic approaches to construct viral capsid-like spherical nanomaterials. *Chem. Commun.* **2018**, *54*, 8944–8959. [[CrossRef](#)]
25. Matsuura, K.; Watanabe, K.; Matsushita, Y.; Kimizuka, N. Guest-binding behavior of peptide nanocapsules self-assembled from viral peptide fragments. *Polym. J.* **2013**, *45*, 529–534. [[CrossRef](#)]
26. Fujita, S.; Matsuura, K. Inclusion of zinc oxide nanoparticles into virus-like peptide nanocapsules self-assembled from viral  $\beta$ -annulus peptide. *Nanomaterials* **2014**, *4*, 778–791. [[CrossRef](#)]
27. Fujita, S.; Matsuura, K. Encapsulation of CdTe Quantum Dots into Synthetic Viral Capsids. *Chem. Lett.* **2016**, *45*, 922–924. [[CrossRef](#)]
28. Matsuura, K.; Nakamura, T.; Watanabe, K.; Noguchi, T.; Minamihata, K.; Kamiya, N.; Kimizuka, N. Self-assembly of Ni-NTA-modified  $\beta$ -annulus peptides into artificial viral capsids and encapsulation of His-tagged proteins. *Org. Biomol. Chem.* **2016**, *14*, 7869–7874. [[CrossRef](#)]
29. Matsuura, K.; Ueno, G.; Fujita, S. Self-assembled artificial viral capsid decorated with gold nanoparticles. *Polym. J.* **2014**, *47*, 146–151. [[CrossRef](#)]
30. Nakamura, Y.; Yamada, S.; Nishikawa, S.; Matsuura, K. DNA-modified artificial viral capsids self-assembled from DNA-conjugated  $\beta$ -annulus peptide. *J. Pept. Sci.* **2017**, *23*, 636–643. [[CrossRef](#)]
31. Matsuura, K.; Matsuura, K. Self-assembled artificial viral capsids bearing coiled-coils at the surface. *Org. Biomol. Chem.* **2017**, *15*, 5070–5077. [[CrossRef](#)]
32. Matsuura, K.; Honjo, T. Artificial Viral Capsid Dressed Up with Human Serum Albumin. *Bioconjug. Chem.* **2019**, *30*, 1636–1641. [[CrossRef](#)] [[PubMed](#)]
33. Matsuura, K.; Ota, J.; Fujita, S.; Shiomi, Y.; Inaba, H. Construction of Ribonuclease-Decorated Artificial Virus-like Capsid by Peptide Self-assembly. *J. Org. Chem.* **2019**, *85*, 1668–1673. [[CrossRef](#)] [[PubMed](#)]
34. Matsuura, K. Dressing up artificial viral capsids self-assembled from C-terminal-modified  $\beta$ -annulus peptides. *Polym. J.* **2020**, *52*, 1035–1041. [[CrossRef](#)]
35. Nakamura, Y.; Inaba, H.; Matsuura, K. Construction of Artificial Viral Capsids Encapsulating Short DNAs via Disulfide Bonds and Controlled Release of DNAs by Reduction. *Chem. Lett.* **2019**, *48*, 544–546. [[CrossRef](#)]
36. Iwasaki, T.; Tokuda, Y.; Kotake, A.; Okada, H.; Takeda, S.; Kawano, T.; Nakayama, Y. Cellular uptake and In Vivo distribution of polyhistidine peptides. *J. Control. Release* **2015**, *210*, 115–124. [[CrossRef](#)] [[PubMed](#)]
37. Hayashi, T.; Shinagawa, M.; Kawano, T.; Iwasaki, T. Drug delivery using polyhistidine peptide-modified liposomes that target endogenous lysosome. *Biochem. Biophys. Res. Commun.* **2018**, *501*, 648–653. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Communication

# Bifunctional Aptamer Drug Carrier Enabling Selective and Efficient Incorporation of an Approved Anticancer Drug Irinotecan to Fibrin Gels

Hiroto Fujita, Yuka Kataoka and Masayasu Kuwahara \* 

Graduate School of Integrated Basic Sciences, Nihon University, 3-25-40 Sakurajosui, Setagaya-ku, Tokyo 156-8550, Japan; fujita.hiroto@nihon-u.ac.jp (H.F.); kataoka.yuka@nihon-u.ac.jp (Y.K.)

\* Correspondence: mkuwa@chs.nihon-u.ac.jp; Tel.: +81-03-5317-9398

Received: 10 September 2020; Accepted: 4 December 2020; Published: 7 December 2020



**Abstract:** We have previously developed a bifunctional aptamer (bApt) binding to both human thrombin and camptothecin derivative (CPT1), and showed that bApt acts as a drug carrier under the phenomenon named selective oligonucleotide entrapment in fibrin polymers (SOEF), which enables efficient enrichment of CPT1 into fibrin gels, resulting in significant inhibition of tumor cell growth. However, although the derivative CPT1 exhibits anticancer activity, it is not an approved drug. In this study, we evaluated the binding properties of bApt to irinotecan, a camptothecin analog commonly used for anticancer drug therapy, in addition to unmodified camptothecin (CPT). Furthermore, we have revealed that irinotecan binds to bApt like CPT1 and is selectively concentrated on fibrin gels formed around the tumor cells under the SOEF phenomenon to suppress cell proliferation.

**Keywords:** drug delivery system; anticancer drug; camptothecin derivative; irinotecan

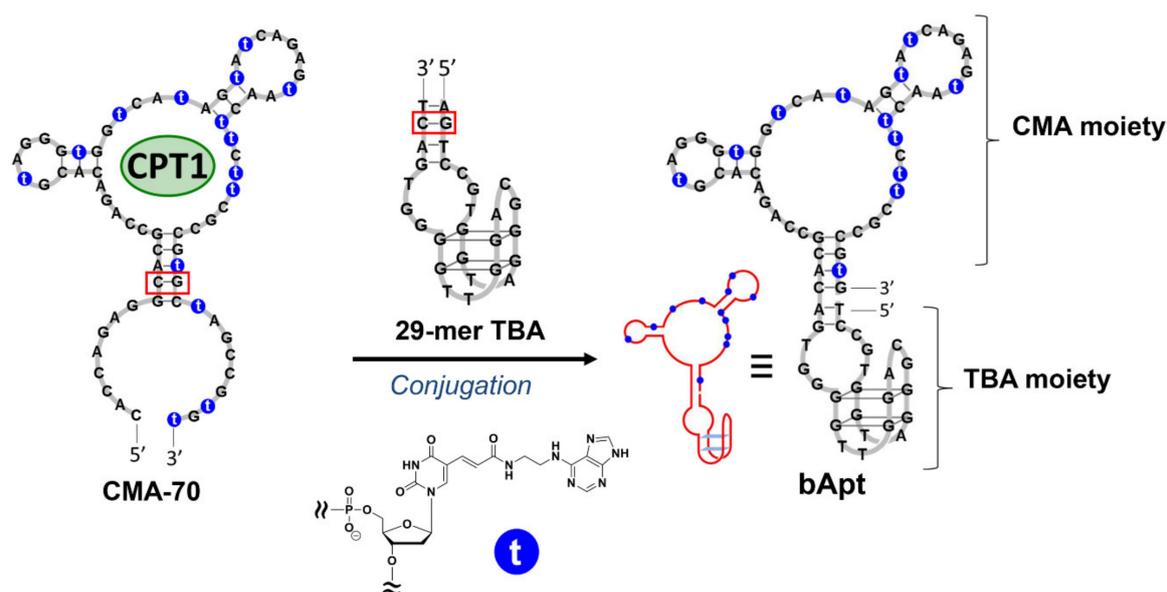
## 1. Introduction

Highly water-soluble derivatives of camptothecin, a type of quinoline alkaloid (irinotecan and topotecan), are currently known to have a broad spectrum of anticancer activity such as in lung cancer, colorectal cancer, ovarian cancer, and malignant lymphoma [1–7]. Its derivatives are the most clinically used because they bind to type I topoisomerase and inhibit recombination with DNA, inhibiting DNA synthesis, causing cancer cell apoptosis, and resulting in an extremely potent anticancer activity [4]. Irinotecan is biotransformed because of the action of carboxylesterase in SN-38 (7-ethyl-10-hydroxy-20(S)-camptothecin), which lacks a substituent at the 10-position on the piperidine ring. Additionally, the cytotoxicity of SN-38 is 100–1000 times greater than that of irinotecan (Figure S1) [8–10]. However, it causes serious side effects, and there is demand for developing a mechanism such as a drug delivery system (DDS) to deliver specific drugs and reduce dosages [11–16]. For example, a liposomal irinotecan called Onivyde has been reported and approved for medical use [17,18].

Previously, we have reported that thrombin-binding aptamer (TBA) is selectively and efficiently incorporated into the gel through thrombin during the fibrin gel formation known as the blood coagulation reaction [19–24]. In the fibrin gel formation process, first, thrombin cleaves FpA and FpB from the  $A\alpha$  and  $B\beta$  chains located at the N-terminus of fibrinogen to produce fibrin monomers [25]. The resulting fibrin monomer engulfs thrombin and polymerizes to form a gel. Therefore, it was suggested that by using this phenomenon named selective oligonucleotide entrapment in fibrin polymers (SOEF), the desired substance or functional group could be introduced into the fibrin gel with TBA [26].

Recently, it has been shown that fibrin plays an important role when cancer cells metastasize or invade other organs, and that fibrinogen is abundant in the cancer stroma [11]. When considering SOEF,

the bifunctional aptamer (bApt), which is a combined 29-mer TBA and camptothecin binding aptamer, can deliver and condense camptothecin as a potent anticancer drug around cancer cells and efficiently suppress their cell growth (Figure 1). Previously, we developed a camptothecin (CPT)-binding modified DNA aptamer (CMA-70), which is a base-modified DNA aptamer obtained by the systematic evolution of ligands via the exponential enrichment (SELEX) method targeting the camptothecin derivative (CPT1) (Figure 1) [27–29]. Therefore, the ability to selectively introduce CPT1 into fibrin gel using bApt as a drug carrier under the SOEF phenomenon provides a new strategy for cancer chemotherapy [30]. However, as a shortcut to the practical application of the said DDS, it is desirable to use an approved anticancer drug such as irinotecan, and thus, this study was conducted.



**Figure 1.** Design of bifunctional aptamer (bApt) conjugated with 29-mer thrombin-binding aptamer (TBA) and a camptothecin (CPT)-binding modified DNA aptamer (CMA-70) containing modified nucleic acids (2'-deoxyuridine-5'-phosphate (dU<sup>ad</sup>TP)) represented with 't'. Red boxes indicate the conjugation site.

## 2. Materials and Methods

### 2.1. Materials

To synthesize CMA-70 and bApt, oligonucleotides, nucleoside triphosphates (dATP, dGTP, dCTP), and *KOD dash* DNA polymerase were, respectively, purchased from Japan Bio Services Co. Ltd. (Saitama, Japan), Roche Diagnostics K. K. (Tokyo, Japan), and Toyobo Co. Ltd. (Tokyo, Japan). For the fluorescence polarization assay, human fibrinogen and thrombin were purchased from Merck K. K. (Tokyo, Japan) and irinotecan and CPT were supplied by Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan). For the cell cultivation and cell growth inhibition assay, Dulbecco's Modified Eagle's Medium (DMEM) low glucose were purchased from Wako Pure Chemical Industries, Ltd., fetal bovine serum (Gibco™) was purchased from Thermo Fischer Scientific K. K. (Tokyo, Japan), CellTracker™ Green CMFDA (5-chloromethylfluorescein diacetate) Dye was purchased from Life Technologies Japan, Ltd. (Osaka, Japan), and HeLa cells (JCRB9004) were purchased from the National Institutes of Biomedical Innovation, Health and Nutrition (Osaka, Japan). All other reagents were of research grade.

### 2.2. Enzymatic Synthesis of CMA-70 and bApt

The CMA-70 aptamer was synthesized by one-primer PCR using a primer (CMA-70\_P1), a template (CMA-70\_Temp), three 2'-deoxyribonucleoside triphosphates (dATP, dGTP, and dCTP), and a

modified 2'-deoxyuridine-5'-triphosphate (dU<sup>ad</sup>TP) with *KOD Dash* DNA polymerase (Table S1) [27,29]. The antisense strand of 5'-monophosphate-labeled ODNs (CMA-70\_Temp) was selectively degraded through  $\lambda$ -exonuclease treatment. Then, the resulting CMA-70 was purified via polyacrylamide gel electrophoresis.

The aptamer bApt was synthesized by one-primer PCR using a primer (TBA\_P1), a template (T1), three 2'-deoxyribonucleoside triphosphates (dATP, dGTP, and dCTP), and a modified 2'-deoxyuridine-5'-triphosphate (dU<sup>ad</sup>TP) with *KOD Dash* DNA polymerase (Table S1) [30]. The synthesized bApt was purified by polyacrylamide gel electrophoresis (Figure S2).

### 2.3. Fluorescence Polarization Assay for CMA-70 Versus CPT Derivatives

To analyze target binding specificity, a fluorescence polarization assay for a target (CPT1, irinotecan, and CPT; final concentration of 0.10  $\mu$ M) with increasing concentrations of CMA-70 (final concentrations of 0, 0.010, 0.025, 0.050, 0.075, 0.10, 0.50, and 1.0  $\mu$ M) were performed at 25 °C using an LS-55 fluorescence spectrometer.

First, CMA-70 was dissolved in 1  $\times$  phosphate-buffered saline (PBS) (11.8 mM HPO<sub>4</sub><sup>2-</sup>, 140 mM Cl<sup>-</sup>, 157 mM Na<sup>+</sup>, 4.5 mM K<sup>+</sup>; pH 7.4) at appropriate concentrations (2.0  $\mu$ M), refolded by denaturing at 94 °C for 0.5 min to protect modified nucleic acids and subsequently cooled to 25 °C at a rate of 0.5 °C/min. Each CMA-70 solution (50  $\mu$ L each) was mixed with 50  $\mu$ L of a target solution (CPT1, irinotecan and CPT; 0.20  $\mu$ M) in a PBS buffer, and incubated at 25 °C for 1 h. Fluorescence polarization for each mixture described above was recorded every 20–30 s for 20 min with excitation at 372 nm and monitoring at 456 nm of 25 °C. Thus, fluorescence polarization at 8 aptamer concentrations was generated as curve (Figure 2).

### 2.4. Fluorescence Polarization Assay for Time Course Analyses of bApt Complexes

Thrombin was dissolved in distilled water, and then, a 2.0  $\mu$ M thrombin solution in 1  $\times$  PBS was prepared. Similarly, mother liquors of fibrinogen (20  $\mu$ M) and irinotecan (1.0  $\mu$ M) were prepared in 1  $\times$  PBS.

A solution containing bApt (1.0  $\mu$ M) was prepared in 1  $\times$  PBS, and then, bApt was refolded by annealing (preheating at 94 °C for 0.5 min followed by cooling to 25 °C at a rate of 0.5 °C/min). First, the irinotecan solution (63  $\mu$ L, 1.1  $\mu$ M) was placed in a 1 cm cuvette and a started fluorescence polarization measurement. After 20 min, the bApt solution (7  $\mu$ L, 1.0  $\mu$ M) was put in the same cuvette. Subsequently, the thrombin solution (3.5  $\mu$ L, 2.0  $\mu$ M) was added to produce a ternary complex of thrombin, irinotecan, and bApt after 40 min of FP measurement. Finally, fibrinogen solution (3.5  $\mu$ L, 20  $\mu$ M) was added after 60 min of FP measurement and measured FP for another 20 min (Figure 3). The final concentration of thrombin/irinotecan/bApt was 0.10  $\mu$ M and fibrinogen was 1.0  $\mu$ M.

### 2.5. Cell Cultivation and Cell Growth Inhibition Assay

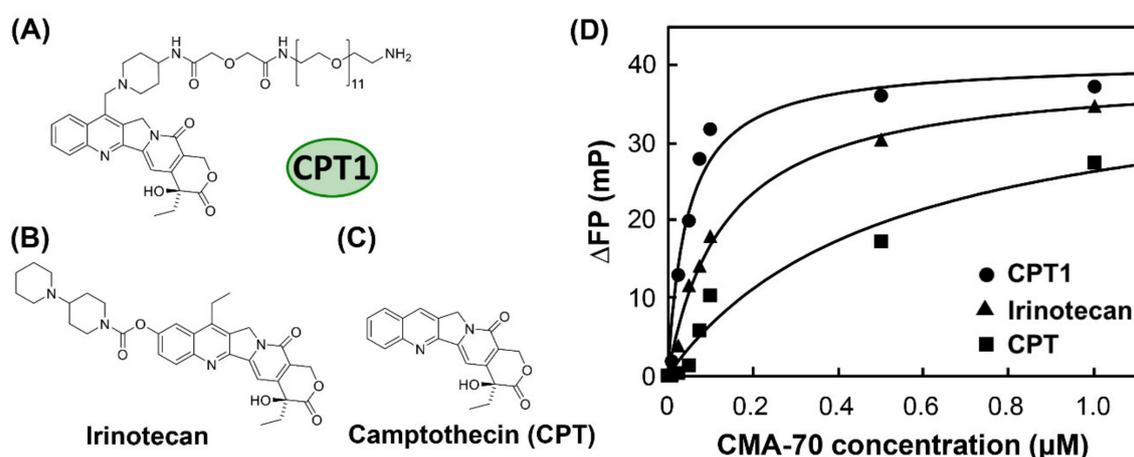
Trypsinized HeLa cells were diluted to a concentration of  $1.0 \times 10^4$  cells/mL with a fresh medium containing DMEM-low glucose, a heat-inactivated serum (Gibco™, FBS, US origin), and an antibiotic solution (penicillin–streptomycin) at a percentage proportion of 90/9/1 (v/v/v%). The medium (300  $\mu$ L) containing the cells was transferred to a 96-well tissue culture plate. Then, the cells were cultured in a humidified incubator containing 95% air and 5% CO<sub>2</sub> at 37 °C for 24 h. After 240  $\mu$ L of the medium was removed, 240  $\mu$ L of the new medium was added to each well. Then, 30  $\mu$ L of the fibrinogen solution (10  $\mu$ M) was added and mixed via gentle pipetting. Subsequently, the solution of the ternary thrombin/CPT1/bApt complex (30  $\mu$ L, 1.0  $\mu$ M) was added and mixed by gentle pipetting. Solutions of the binary thrombin/TBA complex and free CPT1 (30  $\mu$ L, 1.0  $\mu$ M each) were used as controls. Similarly, the abovementioned solutions for the other control experiments were examined. Furthermore, control experiments without fibrinogen (replaced with 1 $\times$  PBS) were conducted. The cells were cultured in a humidified incubator containing 95% air and 5% CO<sub>2</sub> at 37 °C for 48 h.

Next, 300  $\mu\text{L}$  of the medium was replaced with 100  $\mu\text{L}$  of medium containing DMEM-low glucose and CellTracker<sup>TM</sup> Green CMFDA dye in dimethyl sulfoxide (10 mM) at a percentage proportion of 99.8/0.2 (v/v%). The sample was then incubated in a humidified incubator containing 95% air and 5%  $\text{CO}_2$  at 37  $^\circ\text{C}$  for 30 min. After 300  $\mu\text{L}$  of the medium was replaced with 300  $\mu\text{L}$  of 1 $\times$  PBS, the stained cell layer was examined under a BZ-X710 inverted fluorescence microscope to acquire fluorescence images (40 $\times$ ) using excitation light (480 nm) and a cut-off filter (<520 nm) (Figure 4). Three independent experiments were performed.

### 3. Results

#### 3.1. Fluorescence Polarization Assay Using CMA-70

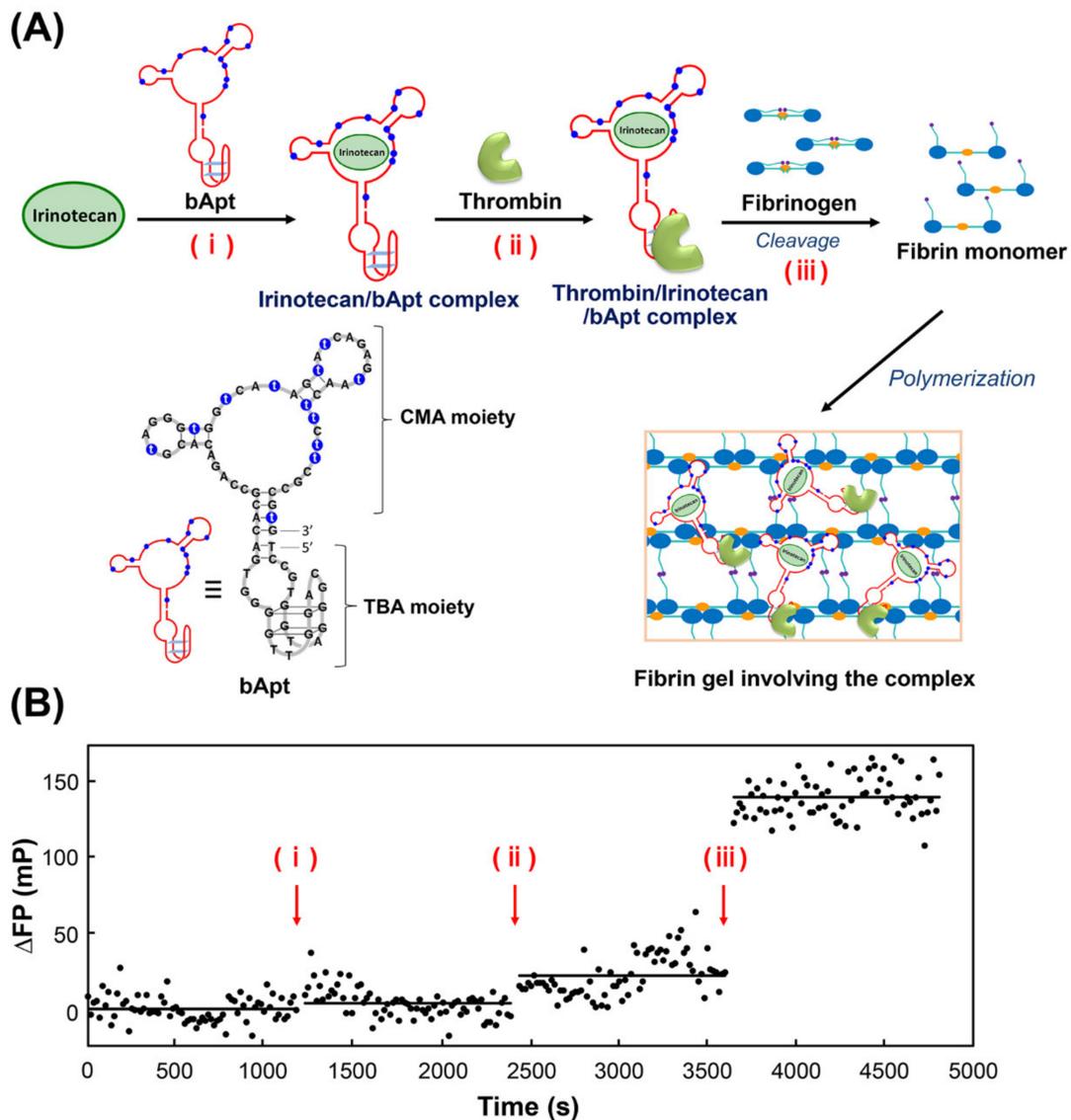
The fluorescence polarization measurement was conducted to examine the binding properties to CPT1, irinotecan, and camptothecin (CPT) (Figure 2A–C) using CMA-70, which is the mother aptamer of the camptothecin-derivative-binding site in bApt. From Figure 2D, we revealed that the fluorescence polarization degree ( $\Delta\text{FP}$ ) of irinotecan increases depending on the CMA-70 concentration, as in CPT1 and CPT. It was indicated that the affinity of CMA-70 for irinotecan exceeded that of CPT.



**Figure 2.** Chemical structures of (A) camptothecin derivative (CPT1), (B) irinotecan, and (C) camptothecin (CPT) and (D) titration curves for CPT1 (●), irinotecan (▲), and CPT (■) polarization versus CMA-70 concentration; polarization was monitored at 456 nm using the excitation wavelength of CPT1, irinotecan, and CPT (372 nm).

#### 3.2. Fluorescence Polarization Assay Using bApt

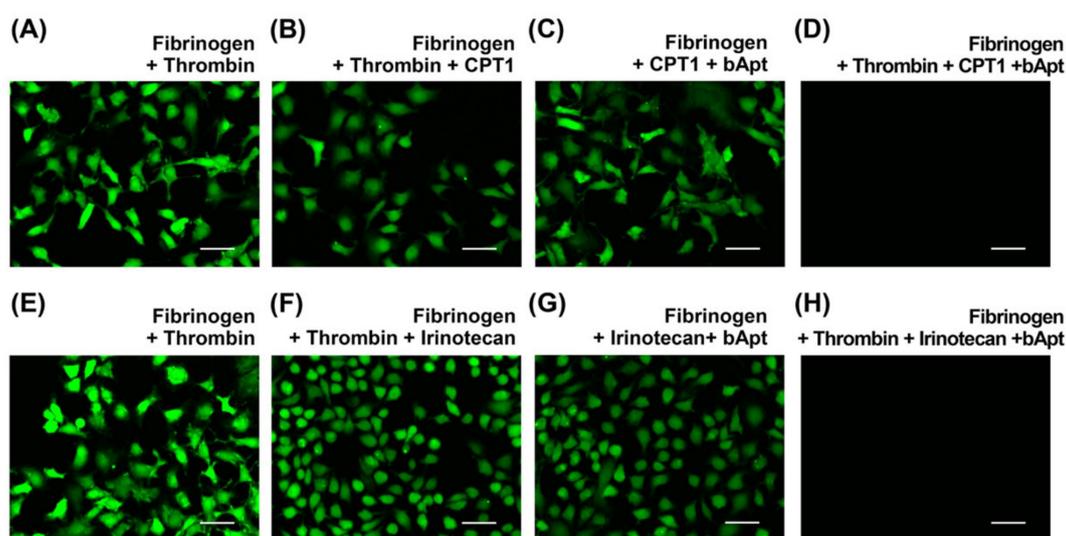
Next, fluorescence polarization measurements of irinotecan were performed to observe the stepwise binding of bApt to its target and the uptake into fibrin gels (Figure 3). Figure 3A shows the time course in differential fluorescence polarization degrees by (i) addition of bApt to irinotecan, (ii) addition of thrombin, and (iii) addition of fibrinogen, followed by polymerization (gelation) of the fibrin monomer generated by thrombin cleavage activity. From Figure 3B, irinotecan bound to the CMA moiety of bApt, as the addition of bApt increased the fluorescence polarization by approximately 4 mP. Furthermore, by adding thrombin, thrombin bound to the TBA moiety of bApt, which increased the fluorescence polarization by approximately 18 mP. Finally, when fibrinogen was added, the fluorescence polarization degree increased by approximately 120 mP, indicating that irinotecan was incorporated into the gel along with bApt during the process of fibrin gel formation.



**Figure 3.** Mechanism of in situ condensation in fibrin gel based on the selective oligonucleotide entrapment in fibrin polymers (SOEF) phenomenon; entrapment of (A) the thrombin/irinotecan/bApt complex during the fibrin gel formation. (B) The titration curves for irinotecan polarization versus bApt concentration; polarization was monitored at 456 nm using an excitation wavelength of 372 nm.

### 3.3. Cell Growth Inhibition Assay

Finally, cell growth inhibition assays were performed using HeLa cells to verify whether bApt could be used as a drug carrier for irinotecan. Fluorescent staining uses CellTracker™, which can selectively stain only live cells. From Figure 4D,H, nonemissive images were obtained only in the presence of thrombin/CPT1/bApt complex or thrombin/irinotecan/bApt complex after 48 h incubation. Furthermore, the fibrin gels containing bApt are not cytotoxic [30]. Thus, it was demonstrated that a significant inhibition of cell proliferation occurs in irinotecan and CPT1 under the SOEF phenomenon.



**Figure 4.** Fluorescence microscopy images of the HeLa cells incubated for 48 h at 37 °C after addition of (A,E) fibrinogen and thrombin, (B,F) fibrinogen, thrombin and CPT1 or irinotecan, (C,G) fibrinogen, CPT1 or irinotecan, and bApt, and (D,H) fibrinogen, thrombin, CPT1 or irinotecan, and bApt. The thrombin, CPT1 or irinotecan, and bApt concentrations are 100 nM each and fibrinogen concentration is 1000 nM. Scale bar = 50  $\mu$ m.

#### 4. Discussion

Fluorescence polarization measurements showed that the CMA part of bApt binds to irinotecan, and further, the irinotecan-binding bApt is efficiently incorporated into fibrin gels by binding the TBA part of bApt to thrombin. Furthermore, cell assays revealed that fibrin gels densely containing irinotecan can cause cancer cell death.

In this study, 29-mer TBA was used at the TBA part of bApt known to have a  $K_d$  value of 0.5 nM for thrombin. This binding to thrombin can delay the blood coagulation reaction by about 16 s compared to the normal condition without any inhibitors [24]. Meanwhile, in the presence of TBAs with different sequences of similar lengths (15–38 mer) with  $K_d$  values of 0.7–126 nM for thrombin, their blood coagulation reactions were delayed by about 2 s [24]. Therefore, the 29-mer TBA sequence used in this study has a relatively high affinity for thrombin and exhibits inhibitory activity although it does not completely inhibit the activity of thrombin.

Currently, phenomena in which thrombin bound to iron oxide nanoparticles is incorporated into fibrin gels were observed with electron microscopes [31]. Those reports agree with the efficient introduction of bApt into fibrin gels due to the high avidity of the TBA to thrombin and the efficient thrombin-mediated uptake of TBA into fibrin gels observed in this study [26,30].

Moreover, in this study, irinotecan was loaded at the CMA site of bApt instead of CPT1. Recently, we have demonstrated that using bApt carrying CPT1, CPT1 was concentrated in a thin layer of about 50  $\mu$ m thickness formed by fibrin gels, and thereby the growth inhibitory effect on cancer cells appeared more significantly about 182 times higher than that of CPT1 only [30]. The inhibitory concentration ( $IC_{50}$ ) values of CPT1 entrapped in fibrin gel ( $7.6 \pm 0.36$  nM) were approximately 170 times lower than that of CPT1 only in the culture medium ( $1300 \pm 600$  nM) [30]. Similarly, in this study, irinotecan was concentrated in a thin layer formed by fibrin gel and assimilated into the cell, resulting in a higher cancer cell growth inhibitory effect when compared with that obtained only using irinotecan. The  $IC_{50}$  values of irinotecan for HeLa cell were  $1320 \pm 130$  nM [2,32]. Therefore, the  $IC_{50}$  values of irinotecan entrapped in fibrin gel were estimated to be approximately several nM to several tens of nM.

For the application of bifunctional molecular recognition biomolecules to pharmaceuticals, an antibody (emicizumab) with two target recognition sites is currently approved for treating hemophilia [33,34]. Conversely, for using a nucleic acid aptamer, it has been reported that

aptamer–aptamer conjugates enable signal control of tyrosine kinase receptor [35] or delivery of doxorubicin to the tumor cell/brain lesions [36], and aptamer–antibody or aptamer–aptamer conjugate can attract T and cancer cells and destroy cancer cells [37–39]; however, they have not been practically used [35–40].

The bApt oligonucleotide has the activity of binding to both the anticancer drug and thrombin, and acts as a carrier for anticancer drugs, which enables the concentration of the anticancer drug in fibrin gels around growing cancer cells. In the future, in the design of bApt, the anticancer drug binding site is compatible with other functionalities. Therefore, the concept and methodology may also be applicable to diseases related to blood coagulation in the future.

## 5. Conclusions

We demonstrated that the approved drug irinotecan can be selectively and efficiently concentrated into fibrin gels. In the future, the bApt we have developed can be applied as a drug carrier to DDS of anticancer agents around cancer tissues where fibrin production is active. Furthermore, creating drug carriers using a combination of approved drugs and their aptamers and their application to DDS are expected to be applied to diseases related to fibrin, such as blood coagulation disorders.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/23/8755/s1>, Figure S1: Metabolism of irinotecan to SN-38, Figure S2: Enzymatic synthesis of bApt through a primer extension reaction, Table S1: Synthetic oligonucleotides used in this study.

**Author Contributions:** Conceptualization, M.K.; methodology, M.K.; validation, H.F. and Y.K.; formal analysis, H.F. and Y.K.; investigation, H.F. and Y.K.; data curation, H.F. and Y.K.; writing—original draft preparation, H.F. and Y.K.; writing—review and editing, H.F., Y.K. and M.K.; supervision, M.K.; project administration, M.K.; funding acquisition, M.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This paper was supported by a Grant for Translational Research Program (H355TS) from the Japan Agency for Medical Research and Development (AMED), and Nihon University Multidisciplinary Research Grant for 2020.

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

## References

1. Shetty, D.; Skorjanc, T.; Olson, M.A.; Trabolsi, A. Self-assembly of stimuli-responsive imine-linked calix[4]arene nanocapsules for targeted camptothecin delivery. *Chem. Commun.* **2019**, *55*, 8876–8879. [CrossRef]
2. Zhou, M.; Liu, M.; He, X.; Yu, H.; Wu, D.; Yao, Y.; Fan, S.; Zhang, P.; Shi, W.; Zhong, B. Synthesis and biological evaluation of novel 10-substituted-7-ethyl-10-hydroxycamptothecin (SN-38) prodrugs. *Molecules* **2014**, *19*, 19718–19731. [CrossRef]
3. Li, X.Q.; Wen, H.Y.; Dong, H.Q.; Xue, W.M.; Pauletti, G.M.; Cai, X.J.; Xia, W.J.; Shi, D.; Li, Y.Y. Self-assembling nanomicelles of a novel camptothecin prodrug engineered with a redox-responsive release mechanism. *Chem. Commun.* **2011**, *47*, 8647–8649. [CrossRef] [PubMed]
4. Pommier, Y. Topoisomerase I inhibitors: Camptothecins and beyond. *Nat. Rev. Cancer* **2006**, *6*, 789–802. [CrossRef] [PubMed]
5. Cheng, K.; Rahier, N.J.; Eisenhauer, B.M.; Gao, R.; Thomas, S.J.; Hecht, S.M. 14-Azacamptothecin: A potent water-soluble topoisomerase I poison. *J. Am. Chem. Soc.* **2005**, *127*, 838–839. [CrossRef] [PubMed]
6. Pizzolato, J.F.; Saltz, L.B. The camptothecins. *Lancet* **2003**, *361*, 2235–2242. [CrossRef]
7. Arimondo, P.B.; Boutorine, A.; Baldeyrou, B.; Bailly, C.; Kuwahara, M.; Hecht, S.M.; Sun, J.S.; Garestier, T.; Helene, C. Design and optimization of camptothecin conjugates of triple helix-forming oligonucleotides for sequence-specific DNA cleavage by topoisomerase I. *J. Biol. Chem.* **2002**, *277*, 3132–3140. [CrossRef]
8. Fujita, K.; Kubota, Y.; Ishida, H.; Sasaki, Y. Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J. Gastroenterol.* **2015**, *21*, 12234–12248. [CrossRef]
9. Iyer, L.; King, C.D.; Whittington, P.F.; Green, M.D.; Roy, S.K.; Tephly, T.R.; Coffman, B.L.; Ratain, M.J. Genetic predisposition to the metabolism of irinotecan (CPT-11). Role of uridine diphosphate glucuronosyltransferase isoform 1A1 in the glucuronidation of its active metabolite (SN-38) in human liver microsomes. *J. Clin. Invest.* **1998**, *101*, 847–854. [CrossRef]

10. Kunimoto, T.; Nitta, K.; Tanaka, T.; Uehara, N.; Baba, H.; Takeuchi, M.; Yokokura, T.; Sawada, S.; Miyasaka, T.; Mutai, M. Antitumor activity of 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxy-camptothecin, a novel water-soluble derivative of camptothecin, against murine tumors. *Cancer Res.* **1987**, *47*, 5944–5947.
11. Gurav, D.; Varghese, O.P.; Hamad, O.A.; Nilsson, B.; Hilborn, J.; Oommen, O.P. Chondroitin sulfate coated gold nanoparticles: A new strategy to resolve multidrug resistance and thromboinflammation. *Chem. Commun.* **2016**, *52*, 966–969. [[CrossRef](#)]
12. Li, J.; Mo, L.; Lu, C.H.; Fu, T.; Yang, H.H.; Tan, W. Functional nucleic acid-based hydrogels for bioanalytical and biomedical applications. *Chem. Soc. Rev.* **2016**, *45*, 1410–1431. [[CrossRef](#)] [[PubMed](#)]
13. Li, H.; Arroyo-Curras, N.; Kang, D.; Ricci, F.; Plaxco, K.W. Dual-reporter drift correction to enhance the performance of electrochemical aptamer-based sensors in whole blood. *J. Am. Chem. Soc.* **2016**, *138*, 15809–15812. [[CrossRef](#)] [[PubMed](#)]
14. McKeague, M.; De Girolamo, A.; Valenzano, S.; Pascale, M.; Ruscito, A.; Velu, R.; Frost, N.R.; Hill, K.; Smith, M.; McConnell, E.M.; et al. Comprehensive analytical comparison of strategies used for small molecule aptamer evaluation. *Anal. Chem.* **2015**, *87*, 8608–8612. [[CrossRef](#)]
15. Coles, D.J.; Rolfe, B.E.; Boase, N.R.; Veedu, R.N.; Thurecht, K.J. Aptamer-targeted hyperbranched polymers: Towards greater specificity for tumours in vivo. *Chem. Commun.* **2013**, *49*, 3836–3838. [[CrossRef](#)]
16. Shoji, A.; Kuwahara, M.; Ozaki, H.; Sawai, H. Modified DNA aptamer that binds the (R)-isomer of a thalidomide derivative with high enantioselectivity. *J. Am. Chem. Soc.* **2007**, *129*, 1456–1464. [[CrossRef](#)]
17. Koeller, J.; Surinach, A.; Arikian, S.R.; Zivkovic, M.; Janeczko, P.; Cockrum, P.; Kim, G. Comparing real-world evidence among patients with metastatic pancreatic ductal adenocarcinoma treated with liposomal irinotecan. *Ther. Adv. Med. Oncol.* **2020**, *12*, 1–8. [[CrossRef](#)]
18. Kciuk, M.; Marciniak, B.; Kontek, R. Irinotecan-Still an Important Player in Cancer Chemotherapy: A Comprehensive Overview. *Int. J. Mol. Sci.* **2020**, *21*, 4919. [[CrossRef](#)] [[PubMed](#)]
19. Riccardi, C.; Napolitano, E.; Platella, C.; Musumeci, D.; Montesarchio, D. G-quadruplex-based aptamers targeting human thrombin: Discovery, chemical modifications and antithrombotic effects. *Pharmacol. Ther.* **2020**, 107649, in press. [[CrossRef](#)] [[PubMed](#)]
20. Chabata, C.V.; Frederiksen, J.W.; Sullenger, B.A.; Gunaratne, R. Emerging applications of aptamers for anticoagulation and hemostasis. *Curr. Opin. Hematol.* **2018**, *25*, 382–388. [[CrossRef](#)] [[PubMed](#)]
21. Takahashi, S.; Sugimoto, N. Volumetric contributions of loop regions of G-quadruplex DNA to the formation of the tertiary structure. *Biophys. Chem.* **2017**, *231*, 146–154. [[CrossRef](#)] [[PubMed](#)]
22. Kasahara, Y.; Irisawa, Y.; Fujita, H.; Yahara, A.; Ozaki, H.; Obika, S.; Kuwahara, M. Capillary electrophoresis-systematic evolution of ligands by exponential enrichment selection of base- and sugar-modified DNA aptamers: Target binding dominated by 2'-O,4'-C-methylene-bridged/locked nucleic acid primer. *Anal. Chem.* **2013**, *85*, 4961–4967. [[CrossRef](#)] [[PubMed](#)]
23. Kasahara, Y.; Irisawa, Y.; Ozaki, H.; Obika, S.; Kuwahara, M. 2',4'-BNA/LNA aptamers: CE-SELEX using a DNA-based library of full-length 2'-O,4'-C-methylene-bridged/linked bicyclic ribonucleotides. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 1288–1292. [[CrossRef](#)] [[PubMed](#)]
24. Tasset, D.M.; Kubik, M.F.; Steiner, W. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J. Mol. Biol.* **1997**, *272*, 688–698. [[CrossRef](#)]
25. Mosesson, M.W. Fibrinogen and fibrin structure and functions. *J. Thromb. Haemost.* **2005**, *3*, 1894–1904. [[CrossRef](#)]
26. Fujita, H.; Inoue, Y.; Kuwahara, M. Selective incorporation of foreign functionality into fibrin gels through a chemically modified DNA aptamer. *Bioorg. Med. Chem. Lett.* **2018**, *28*, 35–39. [[CrossRef](#)]
27. Fujita, H.; Kuwahara, M. Selection of Natural and Base-Modified DNA Aptamers for a Camptothecin Derivative. *Curr. Protoc. Nucleic Acid Chem.* **2016**, *65*, 9.10.1–9.10.19. [[CrossRef](#)]
28. Fujita, H.; Imaizumi, Y.; Kasahara, Y.; Kitadume, S.; Ozaki, H.; Kuwahara, M.; Sugimoto, N. Structural and affinity analyses of g-quadruplex DNA aptamers for camptothecin derivatives. *Pharmaceuticals* **2013**, *6*, 1082–1093. [[CrossRef](#)]
29. Imaizumi, Y.; Kasahara, Y.; Fujita, H.; Kitadume, S.; Ozaki, H.; Endoh, T.; Kuwahara, M.; Sugimoto, N. Efficacy of base-modification on target binding of small molecule DNA aptamers. *J. Am. Chem. Soc.* **2013**, *135*, 9412–9419. [[CrossRef](#)]

30. Kuwahara, M.; Fujita, H.; Kataoka, Y.; Nakajima, Y.; Yamada, M.; Sugimoto, N. In situ condensation of an anti-cancer drug into fibrin gel enabling effective inhibition of tumor cell growth. *Chem. Commun.* **2019**, *55*, 11679–11682. [[CrossRef](#)]
31. Ziv-Polat, O.; Skaat, H.; Shahar, A.; Margel, S. Novel magnetic fibrin hydrogel scaffolds containing thrombin and growth factors conjugated iron oxide nanoparticles for tissue engineering. *Int. J. Nanomed.* **2012**, *7*, 1259–1274. [[CrossRef](#)]
32. Takara, K.; Kitada, N.; Yoshikawa, E.; Yamamoto, K.; Horibe, S.; Sakaeda, T.; Nishiguchi, K.; Ohnishi, N.; Yokoyama, T. Molecular changes to HeLa cells on continuous exposure to SN-38, an active metabolite of irinotecan hydrochloride. *Cancer Lett.* **2009**, *278*, 88–96. [[CrossRef](#)] [[PubMed](#)]
33. Kitazawa, T.; Shima, M. Emicizumab, a humanized bispecific antibody to coagulation factors IXa and X with a factor VIIIa-cofactor activity. *Int. J. Hematol.* **2020**, *111*, 20–30. [[CrossRef](#)] [[PubMed](#)]
34. Sampei, Z.; Igawa, T.; Soeda, T.; Okuyama-Nishida, Y.; Moriyama, C.; Wakabayashi, T.; Tanaka, E.; Muto, A.; Kojima, T.; Kitazawa, T.; et al. Identification and multidimensional optimization of an asymmetric bispecific IgG antibody mimicking the function of factor VIII cofactor activity. *PLoS ONE* **2013**, *8*, e57479. [[CrossRef](#)] [[PubMed](#)]
35. Ueki, R.; Atsuta, S.; Ueki, A.; Sando, S. Nongenetic Reprogramming of the Ligand Specificity of Growth Factor Receptors by Bispecific DNA Aptamers. *J. Am. Chem. Soc.* **2017**, *139*, 6554–6557. [[CrossRef](#)] [[PubMed](#)]
36. Macdonald, J.; Denoyer, D.; Henri, J.; Jamieson, A.; Burvenich, I.J.G.; Pouliot, N.; Shigdar, S. Bifunctional Aptamer-Doxorubicin Conjugate Crosses the Blood-Brain Barrier and Selectively Delivers Its Payload to EpCAM-Positive Tumor Cells. *Nucleic Acid Ther.* **2020**, *30*, 117–128. [[CrossRef](#)]
37. Yang, Y.; Sun, X.; Xu, J.; Cui, C.; Safari Yazd, H.; Pan, X.; Zhu, Y.; Chen, X.; Li, X.; Li, J.; et al. Circular Bispecific Aptamer-Mediated Artificial Intercellular Recognition for Targeted T Cell Immunotherapy. *ACS Nano* **2020**, *14*, 9562–9571. [[CrossRef](#)]
38. Passariello, M.; Camorani, S.; Vetrei, C.; Cerchia, L.; De Lorenzo, C. Novel Human Bispecific Aptamer-Antibody Conjugates for Efficient Cancer Cell Killing. *Cancers* **2019**, *11*, 1268. [[CrossRef](#)]
39. Boltz, A.; Piater, B.; Toleikis, L.; Guenther, R.; Kolmar, H.; Hock, B. Bi-specific aptamers mediating tumor cell lysis. *J. Biol. Chem.* **2011**, *286*, 21896–21905. [[CrossRef](#)]
40. Riccardi, C.; Napolitano, E.; Musumeci, D.; Montesarchio, D. Dimeric and Multimeric DNA Aptamers for Highly Effective Protein Recognition. *Molecules* **2020**, *25*, 5227. [[CrossRef](#)]

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Article

# Effective RNA Regulation by Combination of Multiple Programmable RNA-Binding Proteins

Misaki Sugimoto, Akiyo Suda, Shiroh Futaki  and Miki Imanishi \* 

Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan;

sugimoto.misaki.25c@st.kyoto-u.ac.jp (M.S.); suda.akiyo.3v@kyoto-u.ac.jp (A.S.); futaki@scl.kyoto-u.ac.jp (S.F.)

\* Correspondence: imiki@scl.kyoto-u.ac.jp; Tel.: +81-774-38-3212

Received: 7 September 2020; Accepted: 25 September 2020; Published: 28 September 2020



## Featured Application: RNA targeting.

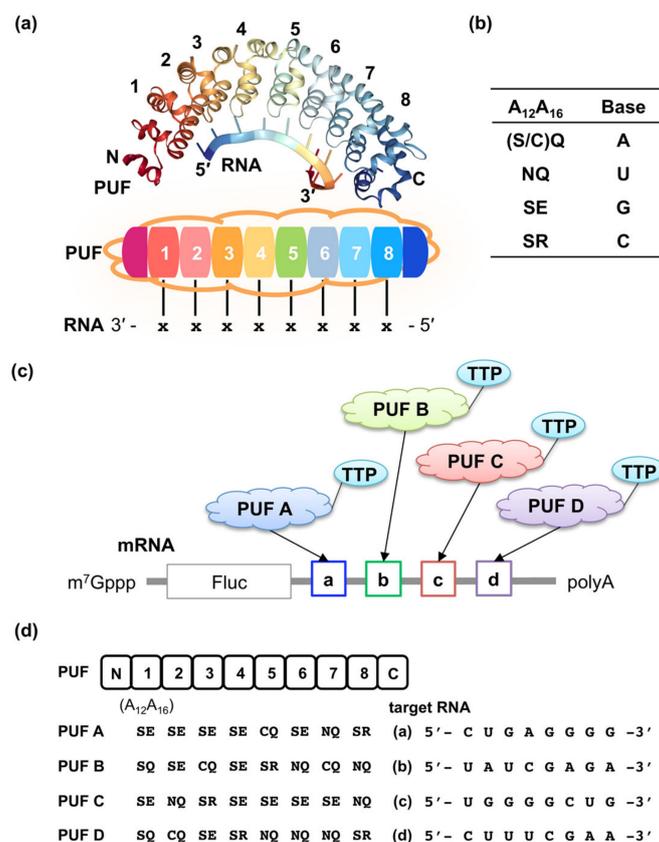
**Abstract:** RNAs play important roles in gene expression through translation and RNA splicing. Regulation of specific RNAs is useful to understand and manipulate specific transcripts. Pumilio and fem-3 mRNA-binding factor (PUF) proteins, programmable RNA-binding proteins, are promising tools for regulating specific RNAs by fusing them with various functional domains. The key question is: How can PUF-based molecular tools efficiently regulate RNA functions? Here, we show that the combination of multiple PUF proteins, compared to using a single PUF protein, targeting independent RNA sequences at the 3' untranslated region (UTR) of a target transcript caused cooperative effects to regulate the function of the target RNA by luciferase reporter assays. It is worth noting that a higher efficacy was achieved with smaller amounts of each PUF expression vector introduced into the cells compared to using a single PUF protein. This strategy not only efficiently regulates target RNA functions but would also be effective in reducing off-target effects due to the low doses of each expression vector.

**Keywords:** RNA binding protein; PUF; RNA regulation

## 1. Introduction

Manipulation of specific RNAs is important for elucidating the roles of specific RNAs and for therapeutic purposes as represented by RNA interference and antisense strategies targeting specific mRNAs. Other than such complementary oligonucleotides, sequence-specific RNA-binding proteins, Pumilio and fem-3 mRNA-binding factor (PUF) proteins and pentatricopeptide repeat (PPR) proteins, and clustered regularly interspaced short palindromic repeats (CRISPR)-based systems, including dCas13 and dCas9-PAMer complexed with guide RNA, have been used for the customized sequence-specific regulation of mRNAs [1–6]. By combining them with various functional domains, such as translational regulation factors [7,8], splicing factors [9–11], RNA editing enzymes [12], and RNA (de)methylation enzymes [13–18], targeted control of RNAs can be achieved. The RNA-binding domains of PUF proteins consist of eight highly homologous structural repeats (repeats 1 to 8), flanked by N- and C-terminal repeats (repeats N and C) [19] (Figure 1a). Each repeat recognizes an RNA base through direct interaction between RNA bases and side chains of amino acids at the 12th and 16th positions [20] (Figure 1b). By changing the amino acid pairs of a PUF repeat, artificial RNA-binding domains corresponding to various 8-nucleotide (nt) RNA sequences can be designed [9,21,22]. As the length of RNA recognized by PUFs (8-nt) is not enough to specify the target RNAs within a huge transcriptome, efforts to increase the length of the RNA-binding sequences have been applied by increasing the number of PUF repeats [21,23–26]. Previous reports also showed the assembly of two PUF proteins in adjacent regions on the RNA to image specific

transcripts [27–30]. By contrast, Cao et al. reported that increasing the number of the PUF-binding sequences in reporter mRNAs results in the effective regulation of reporter genes compared to reporter mRNAs having a single PUF-binding sequence [7]. This was also supported through the regulation of endogenous mRNA by PUFs fused to the RNA decay factor, tristetraprolin (TTP). Abil et al. designed artificial PUF proteins, targeting the 3′ untranslated region (UTR) of vascular endothelial growth factor A (VEGFA) mRNA. TTP-PUF fusion proteins effectively repressed VEGFA production when the particular 8-nt PUF-binding sequence existed at multiple places within the 3′ UTR of the target mRNA [31]. However, it is unusual that many repeated sequences are at the control region of a single mRNA. Instead of increasing the number of target RNA sequences, we intended to use four different PUF proteins, corresponding target RNA sequences that exist at the control region of a reporter RNA (Figure 1c). In this study, we expressed PUF RNA-binding proteins fused with TTP in mammalian cells and enabled the efficient repression of the reporter protein expression by using multiple TTP-PUF proteins.



**Figure 1.** Programmable RNA binding protein, PUF, and regulation of RNA functions. (a) Structure of a PUF protein bound to RNA [PDB ID: 1M8Y] and schematic representation of RNA recognition by a PUF protein. (b) 12th and 16th amino acids (A<sub>12</sub> and A<sub>16</sub>) in each PUF repeat interact with an RNA base. (c) Schematic representation of mRNA regulation by multiple artificial PUF proteins fused with TTP. (d) Design of PUF proteins, PUF (A)–(D), and their recognition RNA sequences, (a)–(d), used in this research.

## 2. Materials and Methods

### 2.1. Plasmid Construction

PUF expression vectors for *E. coli* expression were constructed by a Golden Gate Assembly using the PUF Assembly Kit (Addgene #1000000051) and pET28-GG-PUF as the receiving vector [31]. The amino acid sequences at the 12th and 16th positions of each repeat are shown in Figure 1d.

TTP-PUF expression vectors for mammalian cells were constructed by Golden Gate Assembly using the PUF Assembly Kit and pCMV-TTP(C147R)-GG-PUF as the receiving vector.

Luciferase reporter vectors were constructed, as described in the text that follows. The DNA oligonucleotides containing the corresponding sequences for 4 × RNA (a)–(d), or 1 × (a-b-c-d) and their complementary DNA oligonucleotides (Eurofins genomics) were phosphorylated by T4 polynucleotide kinase (NEB) and annealed; they were inserted into the 3' UTR of the pmirGLO vector (Promega). The sequences of 4 × RNA (a)–(d), or 1 × (a-b-c-d) are shown in Supplementary Figure S1.

## 2.2. Protein Expression and Purification

PUF proteins were expressed in *E. coli* BL21 (DE3) cells (Nippon gene) and purified as described before [23]. Briefly, protein expression was induced by adding 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and *E. coli* cells were incubated at 18 °C overnight. PUF proteins were purified from the *E. coli*-soluble fraction using Ni-NTA HisTrap FF column (GE Healthcare) chromatography followed by ultrafiltration by an Amicon Ultra (10 kDa NMWL) (Merck Millipore).

## 2.3. Electrophoretic Mobility Shift Assay (EMSA)

RNA oligonucleotides, labeled with 6-FAM at the 5'-end, were obtained from FASMAC. Labeled-RNA (10 nM) and three-fold serially diluted concentrations of purified PUF proteins (0–300 nM) were mixed in the reaction buffer containing 10 mM HEPES, 50 mM KCl, 0.5 mM EDTA, 2 mM DTT, 0.01% Tween20, 0.1 mg/mL BSA, 0.02 U/μL RNasin Plus RNase Inhibitor (Promega), and 2.5% Ficoll. The mixture was incubated at 25 °C for 30 min. Free RNA and bound RNA were separated via gel electrophoresis using 8% nondenaturing polyacrylamide gels in 0.5× TBE at 4 °C. After electrophoresis, fluorescently labeled RNAs were detected using the Typhoon scanner RGB system (GE Healthcare). The peak intensity for bound and free labeled-RNA was measured using the ImageQuant TL software (GE Healthcare). The fractions of bound RNA were plotted against the protein concentrations. For calculating equilibrium dissociation constants ( $K_d$ ), the plotted data were fit to the 1:1 binding equation using the Kaleida Graph software (Hulinks) as described before [23].

## 2.4. Cell Culture

NIH3T3 cells (RIKEN cell bank; RCB1862) were maintained in Dulbecco's modified Eagle's medium (DMEM) (Fujifilm Wako Pure Chemical) containing 10% FBS.

## 2.5. Western Blotting Assay

Cells were lysed using the RIPA buffer, and cell lysates were subjected to SDS-PAGE. The proteins were detected using mouse anti-FLAG antibody (Sigma Aldrich; St. Louis, MO, USA, F3165) or mouse anti-β-actin antibody (Sigma Aldrich; AC-74) as primary antibodies, and Goat anti-mouse IgG antibody (GeneTex) as the Horseradish peroxidase (HRP)-labeled secondary antibody. The bands were visualized using an Amersham ECL Prime (GE Healthcare, Chicago, IL, USA).

## 2.6. Luciferase Reporter Assay

Here,  $1.5 \times 10^4$  of NIH3T3 cells were seeded onto 96-well plates the day before transfection. Transfection was performed using Lipofectamine LTX (Thermo Fisher Scientific, Waltham, MA, USA). Transfection mixtures contained 100 ng TTP-PUF expression vector and 1 ng reporter vector. For control, pCMV-TTP(C147R)-GG-PUF (Addgene #1000000051), in which the TTP gene but no PUF-encoding sequences were inserted, was used instead of TTP-PUF expression vectors. Then, 24 h after transfection, cells were lysed using Passive Lysis Buffer (Promega). Firefly luciferase and *Renilla* luciferase activities were measured using the Dual-Luciferase Reporter Assay System (Promega) with measurements taken on a GloMax-Multi Detection System (Promega). Firefly luciferase activity (Fluc) was normalized to *Renilla* luciferase activity (Rluc), and then relative luciferase activities (RLAs), compared to control

samples, were calculated (Equation (1)). The degree of repression by a TTP-PUF relative to the control was indicated by the reciprocal of the relative luciferase activity (Equation (2)).

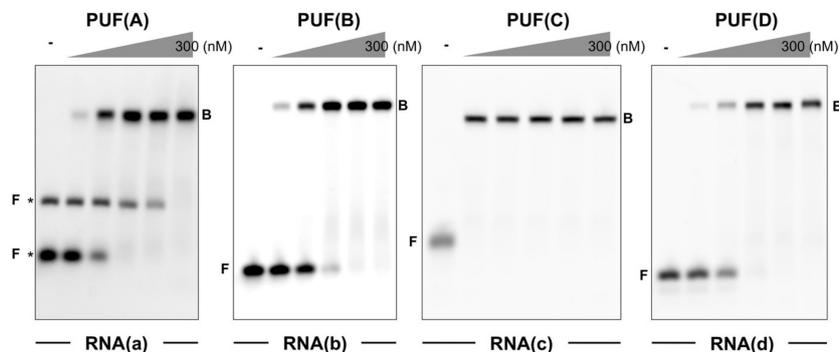
$$RLA_{(TTP-PUF)} = [\text{Fluc}_{(TTP-PUF)} / \text{Rluc}_{(TTP-PUF)}] / [\text{Fluc}_{(\text{control})} / \text{Rluc}_{(\text{control})}] \quad (1)$$

$$\text{Fold repression} = 1 / RLA \quad (2)$$

### 3. Results and Discussion

#### 3.1. RNA-Binding Properties of Programmable RNA-Binding Proteins

Four PUF proteins, PUF (A), (B), (C), and (D), targeting different RNA sequences, RNA (a), (b), (c) and (d), respectively, were designed by assembling PUF repeats corresponding to each RNA base in order (Figure 1d). EMSA revealed the high affinity and specificity of PUF proteins to target RNA sequences. In this assay, the binding of PUF proteins reduced the mobility of the target RNAs, allowing the evaluation of RNA-binding abilities of proteins. PUF (A)–(D) proteins were expressed in *E. coli* cells and purified via affinity chromatography using a Ni-NTA column. PUF (A), (B), (C), and (D) showed band-shifts for RNA (a), (b), (c), and (d), respectively (Figure 2, Figure S2a). The apparent dissociation constants between each PUF protein and corresponding RNA were 10.1 nM (PUF (A)/RNA (a)), 5.3 nM (PUF (B)/RNA (b)), <1 nM (PUF (C)/RNA (c)), and 9.1 nM (PUF (D)/RNA (d)) (Table 1, Figure S2b). Although the  $K_d$  values were somewhat different between PUFs, they were within the reported range of engineered PUF proteins ( $10^{-10}$ – $10^{-8}$  M) [23,31]. Importantly, all PUF (A)–(D) proteins showed high specificity to their corresponding RNA sequences (Figure S2a). RNA (a) showed clear shift bands only when incubated with PUF (A) but not with the other three PUFs. Similarly, RNA (b)–(d) showed clear shift bands only when incubated with their corresponding PUFs.



**Figure 2.** RNA binding of engineered PUF proteins to target RNAs. In the electrophoretic mobility shift assay (EMSA), RNA oligonucleotides, RNA (a)–(d), were incubated with corresponding PUF proteins, PUF (A)–(D), respectively, at the following concentrations (left to right): 0, 3.7, 11, 33, 100, 300 nM, and the mixtures were electrophoresed in native PAGE. “F” and “B” indicate protein-free RNA and protein-bound RNA, respectively. RNA (a) showed two protein-free bands (shown by \*).

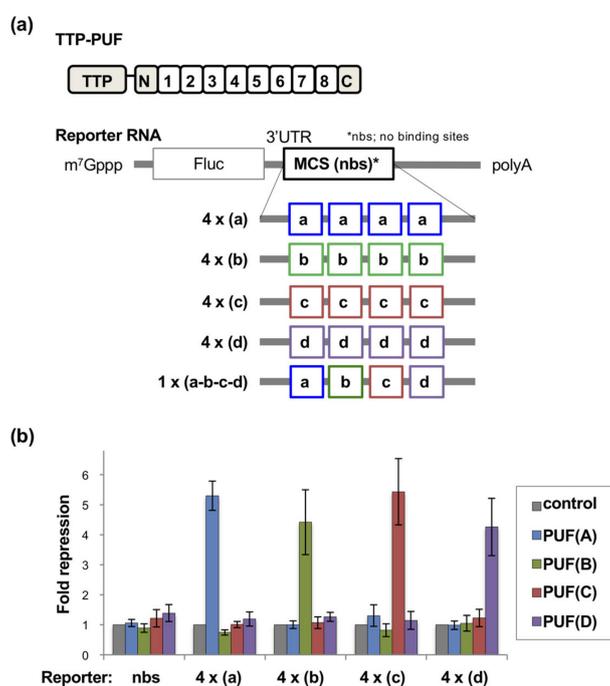
**Table 1.** RNA binding of engineered PUF proteins to specific RNA sequences.

	$K_D$ (nM) <sup>1</sup>			
	RNA(a)	RNA(b)	RNA(c)	RNA(d)
PUF (A)	10.1 ± 1.4	n.d.	n.d.	n.d.
PUF (B)	n.d.	5.3 ± 2.7	n.d.	n.d.
PUF (C)	n.d.	n.d.	<1	n.d.
PUF (D)	n.d.	n.d.	n.d.	9.1 ± 1.4

<sup>1</sup> mean ± SD, n.d.; not detected.

### 3.2. Specific RNA Binding of TTP-PUF within Cells

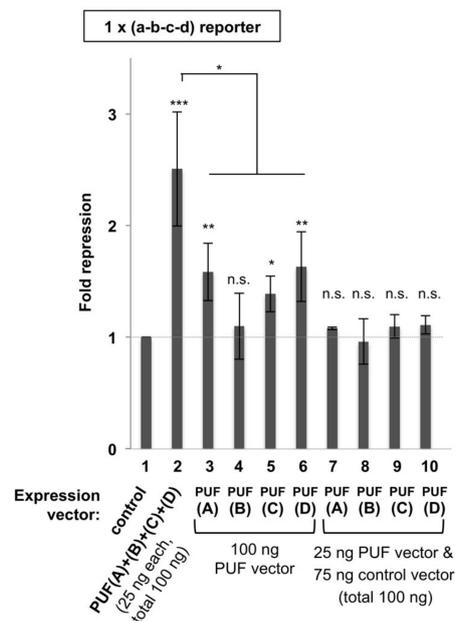
The RNA binding specificities of PUF (A)–(D) for their corresponding target RNAs were also confirmed within living cells using luciferase reporter assays (Figure 3, Figure S1a,b). PUF (A)–(D) fused with tristetraprolin (TTP), an mRNA decay factor, were designed as described before [23,31]. As a control, a TTP expression vector that expresses only TTP without fusing with PUFs was used to estimate the effect of PUF [23,31]. Previous reports showed that a TTP-PUF fusion protein, which can bind to the 3' UTR of the luciferase reporter gene transcript, represses luciferase activity [23,31]. The degree of luciferase repression reflects the RNA binding of TTP-PUF fusion proteins. Here, we prepared reporter vectors that express both firefly luciferase and *Renilla* luciferase genes driven by a PGK and an SV40 promoter, respectively (Figure S1a,b). A four-times repeat sequence of the PUF target sequences, 4 × (a), (b), (c) or (d), was inserted at the 3' UTR of the firefly luciferase gene (Figure 3a). Each TTP-PUF expression or control vectors and a reporter vector were transfected into NIH3T3 cells. In this reporter system, we expected the firefly luciferase activity to be affected by TTP-PUFs, and *Renilla* luciferase activity to reflect the transfection efficiencies. Thus, the firefly luciferase activity, under the expression of each TTP-PUF, was normalized to the *Renilla* luciferase activity of the same sample. The normalized luciferase activity was compared to that of the control sample to evaluate the degree of repression. The expression of the TTP-PUF proteins was confirmed via Western blotting assay (Figure S1c). As shown in Figure 3b, TTP-PUF (A), (B), (C), and (D) showed a 4–6-fold repression in luciferase gene expression with corresponding RNA-binding sequences at the 3' UTR. By contrast, they had little effect on luciferase gene expression without corresponding binding sites at the 3' UTR. None of the TTP-PUFs repressed the luciferase activity in the no binding site (nbs) reporter. These results indicated that the PUF proteins specifically bound to the corresponding RNA sequences at the 3' UTR of the luciferase gene even within cells.



**Figure 3.** Specific binding of engineered PUF proteins to target RNAs in living cells. (a) Schematic representation of TTP-PUF proteins and reporter RNAs. TTP-PUF expression plasmids and luciferase reporter plasmids were co-transfected into NIH3T3 cells. The luciferase reporter genes were transcribed, and reporter RNAs with PUF-binding sequences at the 3' untranslated region (UTR) region should have been produced in the cells. (b) Fold repression of luciferase reporter activities by TTP-PUF proteins. Data are presented as mean ± SD.

### 3.3. Effective Repression of Reporter Genes with Not-Repetitive Sequences by Multiple TTP-PUF Proteins

It is unusual that four or more times sequences exist at the regulatory region of single transcripts. Instead of targeting repeat sequences at the 3' UTR, the synergistic effect of multiple TTP-PUFs was shown through the targeting of multiple RNA sequences at the 3' UTR. As a model of transcripts, a reporter vector with a 1 × (a-b-c-d) sequence at the 3' UTR of the firefly luciferase gene was designed, in which only one binding sequence for each PUF (A)–(D) existed at the 3' UTR (Figure 3a). In contrast to the 4–6-fold repression of the luciferase genes with 4 × binding sequences, each TTP-PUF (A), (C), or (D) showed only about a 1.5-fold repression (Figure 4, lanes 3,5,6), and TTP-PUF (B) alone did not repress luciferase activity (Figure 4, lane 4), maybe because of the low TTP-PUF (B) expression compared to other TTP-PUF proteins (Figure S1c).



**Figure 4.** Combination of TTP-PUF (A)–(D) with 1 × (a-b-c-d) reporter efficiently repressed the luciferase reporter activities at the 3' UTR of the luciferase mRNA. NIH3T3 cells were co-transfected with 1 × (a-b-c-d) reporter (1 ng) and 100 ng (total) of the following expression vectors. Lane 1, control vector; lane 2, mixture of TTP-PUF (A), (B), (C), and (D) (25 ng each); lanes 3–6, 100 ng of TTP-PUF (A) (lane 3), (B) (lane 4), (C) (lane 5), (D) (lane 6); lanes 7–10, 75 ng of control vector and 25 ng of TTP-PUF (A) (lane 7), (B) (lane 8), (C) (lane 9), (D) (lane 10). Data are presented as mean ± SD. (\*\*\*,  $p < 0.005$ ; \*\*,  $p < 0.01$ ; \*,  $p < 0.05$ ; n.s. = not significant).

When NIH3T3 cells were co-transfected with all of the expression vectors (TTP-PUF (A)–(D); 25 ng each) and the 1 × (a-b-c-d) reporter vector, significant repression (2.5-fold) of luciferase activity was observed, even though the amount of each expression vector (25 ng each) was one-fourth of that when used alone (100 ng) (Figure 4, lane 2). The mixture of four TTP-PUF expression vectors had little effect on the reporter lacking a PUF-binding sequence (Figure S3). Furthermore, when the cells were transfected with 25 ng of one of the four TTP-PUF vectors, the same amount used for the mixture, none of them showed significant repression of the 1 × (a-b-c-d) reporter (Figure 4, lanes 7–10). These results indicated that the luciferase activity of the 1 × (a-b-c-d) reporter vector was repressed by the sequence-specific binding of TTP-PUF proteins to their binding sites at the 3' UTR. Additionally, the combination of four TTP-PUF proteins, targeting different RNA sequences, caused effective repression of luciferase activity compared to the use of a single TTP-PUF protein. Differences in repression ratios were observed between reporters with 4 × binding sequences by the corresponding single TTP-PUF (4–6-fold) (Figure 3b) and the 1 × (a-b-c-d) reporter by the mixture of TTP-PUF (A)–(D) (2.5-fold) (Figure 4). One plausible reason for the difference is avidity effects for the

4 × proximity binding sites. The overall RNA structure of the PUF binding sequences can affect the binding of TTP-PUFs. Further optimization will increase the synergistic effects of multiple functional PUF proteins.

So far, synergistic gene activation or repression was shown by a combination of multiple artificial transcription regulators based on transcription activator-like effectors (TALEs) and CRISPR-dCas9 [32,33]. The results using a combination of multiple RNA binding proteins are in accordance with these reports targeting DNA.

#### 4. Conclusions

In this study, four independent PUF proteins were designed and shown to recognize the target RNA sequence with high specificity both in vitro and in cellulo. By combining these four PUF proteins, we showed the use of multiple TTP-PUF proteins in the effective repression of target gene expression, even though the amount of each transfected expression vector was smaller than that of a single PUF protein. Although further verification is needed to understand the off-target effects, we expect that this strategy will reduce off-target effects because of a lower dose of expression vectors. Sequence-specific RNA-binding molecules, like PUFs, can exert desired functions at specific RNA regions when fused with various functional domains. As the importance of RNAs in life science becomes clear, there is an increasing demand for methods to control specific RNAs. Combination of multiple PUF proteins would be a promising strategy to control endogenous RNAs effectively. Furthermore, this concept might be applied to other RNA-targeting molecular tools, such as PPR proteins and CRISPR-Cas13 based systems.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/19/6803/s1>, Figure S1: Plasmids used for luciferase reporter assays and expression of TTP-PUFs, Figure S2: Representative results of EMSA, Figure S3: Effect of the mixture of TTP-PUFs on the nbs reporter.

**Author Contributions:** Conceptualization, M.I. and M.S.; investigation, M.S. and A.S.; writing—M.I. and S.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by JSPS/KAKENHI (19H02850 to M.I.).

**Acknowledgments:** We kindly thank Huimin Zhao for plasmids to construct PUFs (Addgene #1000000051).

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

#### References

1. Pei, Y.; Lu, M. Programmable RNA manipulation in living cells. *Cell. Mol. Life Sci.* **2019**, *76*, 4861–5867. [[CrossRef](#)] [[PubMed](#)]
2. Hall, T.M. De-coding and re-coding RNA recognition by PUF and PPR repeat proteins. *Curr. Opin. Struct. Biol.* **2016**, *36*, 116–121. [[CrossRef](#)] [[PubMed](#)]
3. Chen, Y.; Varani, G. Engineering RNA-binding proteins for biology. *FEBS J.* **2013**, *280*, 3734–3754. [[CrossRef](#)] [[PubMed](#)]
4. Terns, M.P. CRISPR-Based Technologies: Impact of RNA-Targeting Systems. *Mol. Cell* **2018**, *72*, 404–412. [[CrossRef](#)] [[PubMed](#)]
5. Filipovska, A.; Rackham, O. Designer RNA-binding proteins: New tools for manipulating the transcriptome. *RNA Biol.* **2011**, *8*, 978–983. [[CrossRef](#)]
6. Mackay, J.P.; Font, J.; Segal, D.J. The prospects for designer single-stranded RNA-binding proteins. *Nat. Struct. Mol. Biol.* **2011**, *18*, 256–261. [[CrossRef](#)]
7. Cao, J.; Arha, M.; Sudrik, C.; Schaffer, D.V.; Kane, R.S. Bidirectional regulation of mRNA translation in mammalian cells by using PUF domains. *Angew. Chem. Int. Ed.* **2014**, *53*, 4900–4904. [[CrossRef](#)]
8. Campbell, Z.T.; Valley, C.T.; Wickens, M. A protein-RNA specificity code enables targeted activation of an endogenous human transcript. *Nat. Struct. Mol. Biol.* **2014**, *21*, 732–738. [[CrossRef](#)]

9. Dong, S.; Wang, Y.; Cassidy-Amstutz, C.; Lu, G.; Bigler, R.; Jezyk, M.R.; Li, C.; Hall, T.M.; Wang, Z. Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains. *J. Biol. Chem.* **2011**, *286*, 26732–26742. [[CrossRef](#)]
10. Wang, Y.; Cheong, C.G.; Hall, T.M.; Wang, Z. Engineering splicing factors with designed specificities. *Nat. Methods* **2009**, *6*, 825–830. [[CrossRef](#)]
11. Du, M.; Jillette, N.; Zhu, J.J.; Li, S.; Cheng, A.W. CRISPR artificial splicing factors. *Nat. Commun.* **2020**, *11*, 2973. [[CrossRef](#)] [[PubMed](#)]
12. Cox, D.B.T.; Gootenberg, J.S.; Abudayyeh, O.O.; Franklin, B.; Kellner, M.J.; Joung, J.; Zhang, F. RNA editing with CRISPR-Cas13. *Science* **2017**, *358*, 1019–1027. [[CrossRef](#)] [[PubMed](#)]
13. Liu, X.M.; Zhou, J.; Mao, Y.; Ji, Q.; Qian, S.B. Programmable RNA N<sup>6</sup>-methyladenosine editing by CRISPR-Cas9 conjugates. *Nat. Chem. Biol.* **2019**, *15*, 865–871. [[CrossRef](#)] [[PubMed](#)]
14. Shinoda, K.; Suda, A.; Otonari, K.; Futaki, S.; Imanishi, M. Programmable RNA methylation and demethylation using PUF RNA binding proteins. *Chem. Commun.* **2020**, *56*, 1365–1368. [[CrossRef](#)]
15. Rau, K.; Roesner, L.; Rentmeister, A. Sequence-specific m<sup>6</sup>A demethylation in RNA by FTO fused to RCas9. *RNA* **2019**, *25*, 1311–1323. [[CrossRef](#)]
16. Wilson, C.; Chen, P.J.; Miao, Z.; Liu, D.R. Programmable m<sup>6</sup>A modification of cellular RNAs with a Cas13-directed methyltransferase. *Nat. Biotechnol.* **2020**. [[CrossRef](#)]
17. Li, J.; Chen, Z.; Chen, F.; Xie, G.; Ling, Y.; Peng, Y.; Lin, Y.; Luo, N.; Chiang, C.M.; Wang, H. Targeted mRNA demethylation using an engineered dCas13b-ALKBH5 fusion protein. *Nucleic Acids Res.* **2020**, *48*, 5684–5694. [[CrossRef](#)]
18. Zhao, J.; Li, B.; Ma, J.; Jin, W.; Ma, X. Photoactivatable RNA N<sup>6</sup>-methyladenosine editing with CRISPR-Cas13. *Small* **2020**, *16*, e1907301. [[CrossRef](#)]
19. Wang, X.; Zamore, P.D.; Hall, T.M. Crystal structure of a Pumilio homology domain. *Mol. Cell* **2001**, *7*, 855–865. [[CrossRef](#)]
20. Wang, X.; McLachlan, J.; Zamore, P.D.; Hall, T.M. Modular recognition of RNA by a human pumilio-homology domain. *Cell* **2002**, *110*, 501–512. [[CrossRef](#)]
21. Filipovska, A.; Razif, M.F.; Nygard, K.K.; Rackham, O. A universal code for RNA recognition by PUF proteins. *Nat. Chem. Biol.* **2011**, *7*, 425–427. [[CrossRef](#)] [[PubMed](#)]
22. Cheong, C.G.; Hall, T.M. Engineering RNA sequence specificity of Pumilio repeats. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13635–13639. [[CrossRef](#)] [[PubMed](#)]
23. Shinoda, K.; Tsuji, S.; Futaki, S.; Imanishi, M. Nested PUF proteins: Extending target RNA elements for gene regulation. *Chembiochem* **2018**, *19*, 171–176. [[CrossRef](#)] [[PubMed](#)]
24. Adamala, K.P.; Martin-Alarcon, D.A.; Boyden, E.S. Programmable RNA-binding protein composed of repeats of a single modular unit. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E2579–E2588. [[CrossRef](#)]
25. Zhao, Y.Y.; Mao, M.W.; Zhang, W.J.; Wang, J.; Li, H.T.; Yang, Y.; Wang, Z.; Wu, J.W. Expanding RNA binding specificity and affinity of engineered PUF domains. *Nucleic Acids Res.* **2018**, *46*, 4771–4782. [[CrossRef](#)]
26. Criscuolo, S.; Gatti Iou, M.; Merolla, A.; Maragliano, L.; Cesca, F.; Benfenati, F. Engineering REST-specific synthetic PUF proteins to control neuronal gene expression: A combined experimental and computational study. *ACS Synth. Biol.* **2020**, *9*, 2039–2054. [[CrossRef](#)]
27. Ozawa, T.; Natori, Y.; Sato, M.; Umezawa, Y. Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat. Methods* **2007**, *4*, 413–419. [[CrossRef](#)]
28. Yoshimura, H.; Inaguma, A.; Yamada, T.; Ozawa, T. Fluorescent probes for imaging endogenous beta-actin mRNA in living cells using fluorescent protein-tagged pumilio. *ACS Chem. Biol.* **2012**, *7*, 999–1005. [[CrossRef](#)]
29. Yamada, T.; Yoshimura, H.; Inaguma, A.; Ozawa, T. Visualization of nonengineered single mRNAs in living cells using genetically encoded fluorescent probes. *Anal. Chem.* **2011**, *83*, 5708–5714. [[CrossRef](#)]
30. Ozawa, T.; Umezawa, Y. Genetically-encoded fluorescent probes for imaging endogenous mRNA in living cells. *Methods Mol. Biol.* **2011**, *714*, 175–188.
31. Abil, Z.; Denard, C.A.; Zhao, H. Modular assembly of designer PUF proteins for specific post-transcriptional regulation of endogenous RNA. *J. Biol. Eng.* **2014**, *8*, 7. [[CrossRef](#)] [[PubMed](#)]

32. Nomura, W.; Matsumoto, D.; Sugii, T.; Kobayakawa, T.; Tamamura, H. Efficient and orthogonal transcription regulation by chemically inducible artificial transcription factors. *Biochemistry* **2018**, *57*, 6452–6459. [[CrossRef](#)] [[PubMed](#)]
33. Perez-Pinera, P.; Ousterout, D.G.; Brunger, J.M.; Farin, A.M.; Glass, K.A.; Guilak, F.; Crawford, G.E.; Hartemink, A.J.; Gersbach, C.A. Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat. Methods* **2013**, *10*, 239–242. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



## Article

# An RNA Triangle with Six Ribozyme Units Can Promote a *Trans*-Splicing Reaction through Trimerization of Unit Ribozyme Dimers

Junya Akagi <sup>1</sup>, Takahiro Yamada <sup>1</sup>, Kumi Hidaka <sup>2</sup>, Yoshihiko Fujita <sup>3</sup>, Hirohide Saito <sup>3</sup>, Hiroshi Sugiyama <sup>2,4</sup>, Masayuki Endo <sup>2,4</sup> , Shigeyoshi Matsumura <sup>1</sup> and Yoshiya Ikawa <sup>1,\*</sup> 

<sup>1</sup> Department of Chemistry, Graduate School of Science and Engineering, University of Toyama, Gofuku 3190, Toyama 930-8555, Japan; clspinachaj0759@gmail.com (J.A.); m2041213@ems.u-toyama.ac.jp (T.Y.); smatsumu@sci.u-toyama.ac.jp (S.M.)

<sup>2</sup> Department of Chemistry, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan; hidaka.kumi.8z@kyoto-u.ac.jp (K.H.); hs@kuchem.kyoto-u.ac.jp (H.S.); endo@kuchem.kyoto-u.ac.jp (M.E.)

<sup>3</sup> Department of Life Science Frontiers, Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto 606-8502, Japan; yoshihiko.fujita@cira.kyoto-u.ac.jp (Y.F.); hirohide.saito@cira.kyoto-u.ac.jp (H.S.)

<sup>4</sup> Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto 606-8502, Japan

\* Correspondence: yikawa@sci.u-toyama.ac.jp; Tel.: +81-76-445-6599



**Citation:** Akagi, J.; Yamada, T.; Hidaka, K.; Fujita, Y.; Saito, H.; Sugiyama, H.; Endo, M.; Matsumura, S.; Ikawa, Y. An RNA Triangle with Six Ribozyme Units Can Promote a *Trans*-Splicing Reaction through Trimerization of Unit Ribozyme Dimers. *Appl. Sci.* **2021**, *11*, 2583. <https://doi.org/10.3390/app11062583>

Academic Editors: Tamaki Endoh and Chih-Ching Huang

Received: 29 November 2020

Accepted: 11 March 2021

Published: 14 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Ribozymes are catalytic RNAs that are attractive platforms for the construction of nanoscale objects with biological functions. We designed a dimeric form of the *Tetrahymena* group I ribozyme as a unit structure in which two ribozymes were connected in a tail-to-tail manner with a linker element. We introduced a kink-turn motif as a bent linker element of the ribozyme dimer to design a closed trimer with a triangular shape. The oligomeric states of the resulting ribozyme dimers (kUrds) were analyzed biochemically and observed directly by atomic force microscopy (AFM). Formation of kUrd oligomers also triggered *trans*-splicing reactions, which could be monitored with a reporter system to yield a fluorescent RNA aptamer as the *trans*-splicing product.

**Keywords:** catalytic RNA; group I ribozyme; RNA nanostructure; RNA nanotechnology; RNA-protein complex; *trans*-splicing

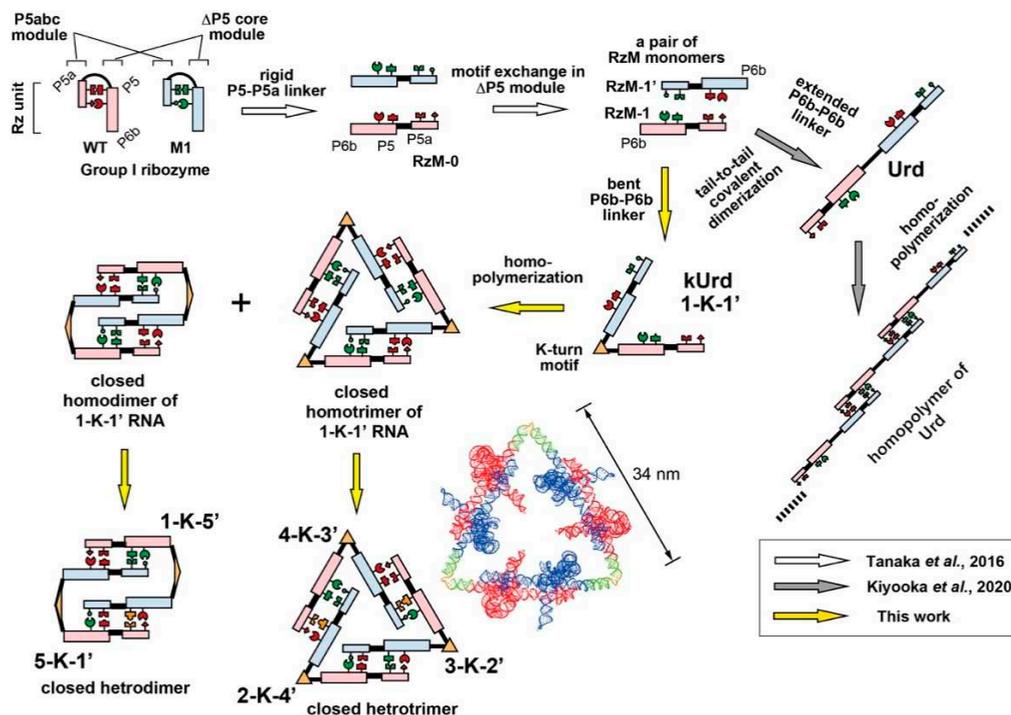
## 1. Introduction

Folded polypeptides work not only in monomeric states but also in assembled states, including homo- and heterooligomers [1–3]. In the course of enzyme evolution, symmetrical protein homooligomers with polygonal and polyhedral shapes emerged from their monomeric ancestors, presumably due to the advantages in their biological properties over the respective monomeric states, including enzymatic activities and structural stability [1–3]. Interfaces between monomer units play a key role in the assembly of protein oligomers. The artificial design of protein oligomers with symmetrical shapes has been explored with the use of protein–protein interaction interfaces extracted from naturally occurring protein oligomers, generating the field of protein nanostructure research [4–6]. Recent challenges in this field include *de novo* design of artificial protein–protein interaction interfaces that can serve as modular structural units for versatile protein nanostructure design [7,8]. Among the various forms of symmetrical protein homooligomers, symmetrical protein homodimers can be regarded as the simplest molecular parts for the rational design and assembly of protein structures [9–11].

RNA is a biopolymer that can fold into defined 3D structures with biological functions. Structured RNAs are known to perform many types of protein-like functions, including roles as catalysts and receptors [12–14]. Large and complex 3D RNA structures frequently formed through noncovalent assembly of structural modules, which can fold locally within a single polynucleotide chain [15]. Intramolecular assembly of RNA structural modules is

supported and maintained by RNA–RNA interaction interfaces [15,16]. Their association is usually strong enough to reconstitute the catalytic activity in a multimolecular format, in which separately prepared modular RNAs are assembled noncovalently by RNA–RNA interaction interfaces [17–25].

We recently engineered an RNA–RNA interaction interface of the *Tetrahymena* group I ribozyme to design a prototype of ribozyme modules (abbreviated as RzM-0, Figure 1) with self-dimerizing capability [26,27]. The RNA–RNA interface in the RzM-0 homodimer is supported by two sets of tertiary interactions between the  $\Delta$ P5 core ribozyme module and P5abc module, which are connected by flexible linker elements (P5-P5a linker) in the parent ribozyme and form strong intramolecular associations (WT group I ribozyme, Figure 1). We rationally replaced the flexible P5-P5a with a rigid duplex to prevent intramolecular  $\Delta$ P5-P5abc assembly (Figure S1C), resulting in homodimerization of the resulting variant (RzM-0 in Figure 1) in a head-to-head manner by forming RzM:RzM interface 0:0 [26,27]. A pair of symmetrical  $\Delta$ P5-P5abc interfaces yielding RzM:RzM homodimer was then engineered to asymmetrical recognition interfaces, such as RzM:RzM interface 1:1', yielding RzM:RzM heterodimers (Figure 1). We then connected a pair of RzM monomers in a tail-to-tail manner at their P6b region, enabling the resulting covalent dimer (termed unit ribozyme dimer and abbreviated as Urd) to homooligomerize in a directional manner (Figure 1) [28]. Unit ribozyme dimers (Urds) homooligomerized to form an open chain assembly (Figure 1). The extents of oligomerization were programmable by designing the specificity of the assembly interfaces between two ribozyme modules (RzMs) in Urds [28]. We also expected that the parent Urds could be modified to produce closed forms of symmetrical oligomers if we used a bent P6b-P6b linker. In this study, we constructed and examined a new series of Urds, which were designed to form closed trimers with a triangular shape.



**Figure 1.** Design of closed oligomers based on engineered group I ribozyme dimers. Scheme of modular redesign to generate unit ribozyme dimers (kUrds) and their oligomerization. White arrows indicate structural redesign to prepare RzM RNAs and their noncovalent dimerization as reported by Tanaka et al. in 2016 [26]. Gray arrows indicate rational redesign for covalent dimerization of RzM RNAs reported by Kiyooka et al. in 2020 [28]. Yellow arrows indicate structural redesign to construct kUrd 1-K-1' in this study. Sequences and secondary structures of wild-type ribozyme, M1 variant ribozyme and their structural elements are shown in Figure S1.

## 2. Materials and Methods

### 2.1. Molecular Design

Three-dimensional structural models of kink-turn unit ribozyme dimer (abbreviated as kUrd) **1-K-1'** and its closed homotrimer were constructed using the crystal structure of a shortened form of the *Tetrahymena* group I ribozyme (PDB ID: 1X8W) [29], a model 3D structure of its full-length form [27] and the KT-15 kink-turn motif (PDB ID: 1JJ2) [30]. Molecular modeling was performed using Discovery Studio (BIOVIA, San Diego, CA, USA), protocol of which has been described with an example to construct a 3D model structure of the RzM:RzM dimer [27]. In this study, the modeling was continued to connect three RzM:RzM dimers and three KT-15 motifs by pairs of A-form RNA duplexes (15 bp/10 bp).

### 2.2. Plasmid Construction and RNA Preparation

Plasmids encoding the sequences of  $\alpha$ - and  $\beta$ -chains of kUrd RNAs were derived in two steps from plasmids encoding chimeric constructs of the wild-type *Tetrahymena* group I intron and its M1 variant-type intron [31]. PCR-based mutagenesis was used in each step of the stepwise plasmid construction. We first changed the sequences of P5-P5a linker elements of the starting plasmids to form the rigid P5-P5a duplex to yield RzM:RzM interfaces. The resulting plasmids were further modified by replacing the sequences of P6b elements to introduce the KT-15 kink-turn as the bent linker element. The resulting plasmids were used as templates for PCR. For each PCR, the sense primer contained the T7 promoter sequence. PCR-amplified DNA templates were used for synthesis of  $\alpha$ - and  $\beta$ -chains of kUrd RNAs by in vitro transcription with T7 RNA polymerase. Transcription reactions were performed for 4.5 h at 37 °C in the presence of nucleotide triphosphates (1 mM each), 15 mM Mg<sup>2+</sup>, 40 mM Tris-Cl (pH 7.5), 10 mM dithiothreitol and 2 mM spermidine. The DNA template in the reaction mixture was removed by DNase treatment for 30 min. The transcribed RNA was purified on 4% denaturing polyacrylamide gels. 3'-End labeling of RNAs with BODIPY fluorophore was performed according to the published protocol in which the diol moiety in the 3' end ribose was oxidized by NaIO<sub>4</sub> to produce two aldehyde moieties, which were then connected with a primary amino group in the BODIPY derivative in the presence of NaBH<sub>3</sub>CN [32].

### 2.3. Preparation of L7Ae Protein and L7Ae-EGFP Fusion Protein

L7Ae protein and L7Ae-EGFP fusion protein were prepared according to the published protocol [33,34]. Briefly, the pET 28-b vector was used for cloning and construction of the recombinant protein L7Ae from *Archaeoglobus fulgidus* and its EGFP fusion protein. *Escherichia coli* BL-21 (DE3) (pLysS) cells were then used for protein production.

### 2.4. Electrophoretic Mobility Shift Assay (EMSA) of kUrd Oligomers

To analyze homooligomerization of kUrd **1-K-1'**, aqueous solution (18  $\mu$ L) containing its  $\alpha$ - and  $\beta$ -chain RNAs (0.14  $\mu$ M) was heated at 80 °C for 2.5 min and then cooled to 37 °C. The 10 $\times$  concentrated buffer (2  $\mu$ L) containing 300 mM Tris-Cl (pH 7.5) and 100–200 mM Mg(OAc)<sub>2</sub> was added to the RNA solution. The resulting solution contained 30 mM Tris-Cl (pH 7.5), 10–20 mM Mg(OAc)<sub>2</sub> and 0.125  $\mu$ M each  $\alpha$ - and  $\beta$ -chain RNA to form 0.125  $\mu$ M **1-K-1'**. To analyze heterotrimers, three kUrd solutions (6.67  $\mu$ L each), each of which contained 0.375  $\mu$ M each  $\alpha$ - and  $\beta$ -chain RNA, were prepared separately and their mixed solution (20  $\mu$ L) was then prepared. To analyze heterodimers, two kUrd solutions (10  $\mu$ L each), each of which contained 0.25  $\mu$ M each  $\alpha$ - and  $\beta$ -chain RNA, were prepared separately and their mixed solution (20  $\mu$ L) was then prepared. The resulting solution (20  $\mu$ L) containing kUrd homo- or heterooligomers was incubated at 37 °C for 30 min and then 4 °C for 30 min and 6 $\times$  loading buffer containing 50% glycerol and 0.1% xylene cyanol (4  $\mu$ L) were added. The samples (24  $\mu$ L) were loaded onto a 5% non-denaturing polyacrylamide gel (29:1 acrylamide:bisacrylamide) containing 50 mM Tris-acetate (pH 7.5) and 5–25 mM Mg(OAc)<sub>2</sub>. Electrophoresis was performed at 4 °C, 200 V for the initial 5 min, followed by 75 V for 5 or 12 h. Gels were analyzed using a Pharos FX fluorimager (BioRad,

Hercules, CA, USA). Each RzM:RzM interface consists of a pair of  $\Delta P5/P5abc$  interfaces. The binding affinity between  $\Delta P5$  core module and P5abc module varied considerably depending on the identity of the  $\Delta P5/P5abc$  interface [35,36]. Stability of the interface is improved by increasing  $Mg^{2+}$  concentration [35,36]. In each EMSA, we therefore tuned the concentration of  $Mg^{2+}$ .

### 2.5. Atomic Force Microscopy

Atomic force microscopy (AFM) was performed on a high-speed AFM (Nano Live Vision; RIBM, Tsukuba, Japan) according to the protocol reported previously [37]. The sample solutions containing kUrd RNAs were diluted to a final concentration of ~80 nM in folding buffer containing 17.5 mM  $Mg^{2+}$  and 2  $\mu L$  of the sample was deposited onto the mica surface of the AFM.

### 2.6. Electrophoretic Mobility Shift Assay (EMSA) of the kUrd Trimer with L7Ae Protein

Three kUrd solutions (4.1  $\mu L$  each), each of which contained 30 mM Tris-Cl (pH 7.5), 17.5 mM  $Mg(OAc)_2$  and 0.61  $\mu M$  each  $\alpha$ - and  $\beta$ -chain RNA, were prepared separately according to the protocol for kUrd oligomer formation. To each kUrd solution (4.1  $\mu L$ ) was added 5 $\times$  concentrated binding buffer (1.33  $\mu L$ ) containing 50 mM HEPES-KOH (pH 7.5), 750 mM KCl and 7.5 mM  $MgCl_2$ . To this solution, was added a solution (1.25  $\mu L$ ) of L7Ae (or L7Ae-EGFP) protein with appropriate concentration and the resulting solution (6.68  $\mu L$ ) was incubated at 37  $^{\circ}C$  for 30 min. Three kUrd+L7Ae solutions were then mixed. The resulting solution (20  $\mu L$ ) containing three kUrds (0.125  $\mu M$  each) and L7Ae was incubated at 37  $^{\circ}C$  for 30 min and then 4  $^{\circ}C$  for 30 min and 6 $\times$  loading buffer containing 50% glycerol and 0.1% xylene cyanol (4  $\mu L$ ) was added. The samples (24  $\mu L$ ) were analyzed in a same manner to EMSA of kUrd oligomers.

### 2.7. Substrate Cleavage Reaction

To prepare a set of kUrd monomer solutions, aqueous solutions containing appropriate pairs of  $\alpha$ - and  $\beta$ -chain RNAs (final concentration of each kUrd: 0.25  $\mu M$ ) were heated at 80  $^{\circ}C$  for 5 min and then cooled to 37  $^{\circ}C$ . Then, 10 $\times$  concentrated reaction buffer (final concentration: 30 mM Tris-Cl, pH 7.5 and 3 mM or 5 mM  $MgCl_2$ ) and 2 mM guanosine triphosphate (final concentration: 0.2 mM) were added to the RNA solution and incubated at 37  $^{\circ}C$  for 30 min to form a given kUrd monomer. The resulting kUrd monomer solutions were then mixed to afford a kUrd oligomer solution, which was additionally incubated at 37  $^{\circ}C$  for 30 min. Ribozyme reactions were started by adding substrate-a (5'-FAM-GGCCUCUAAAAA-3', final concentration: 0.25  $\mu M$ ) conjugated with carboxyfluorescein (FAM) at its 5' end [37]. The resulting solution containing 0.25  $\mu M$  each kUrd and 0.25  $\mu M$  substrate-a was kept at 37  $^{\circ}C$ . Aliquots were taken at given time points and the mixtures were electrophoresed on 15% denaturing polyacrylamide gels. The substrate and reaction products were analyzed using a Pharos FX fluorimager. All of the ribozyme activity assays were repeated at least twice and representative gel images are shown.

### 2.8. Trans-Splicing Monitored by Fluorescence of Spinach RNA-DFHBI

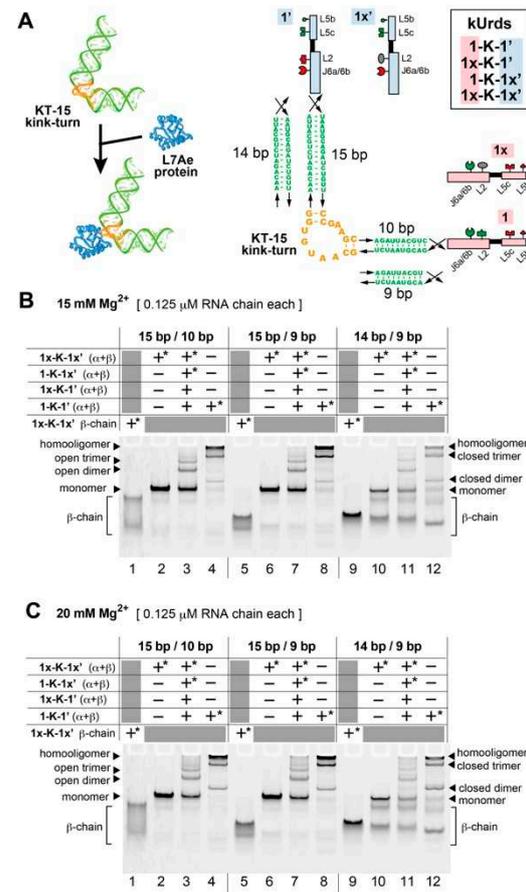
To prepare three kUrd monomer solutions, aqueous solutions (17  $\mu L$ ) containing appropriate pairs of  $\alpha$ - and  $\beta$ -chain RNAs (final concentration of each kUrd: 0.25  $\mu M$ ) were heated at 80  $^{\circ}C$  for 5 min and then cooled to 37  $^{\circ}C$ . Then, 10 $\times$  concentrated reaction buffer (2  $\mu L$ , final concentration: 40 mM Tris-Cl, pH 7.5, 125 mM KCl, 5 mM  $MgCl_2$  and 10  $\mu M$  DFHBI) and 2 mM guanosine triphosphate (1  $\mu L$ , final concentration: 0.2 mM) were added to the RNA solution and incubated at 37  $^{\circ}C$  for 30 min to form a given kUrd monomer. The resulting kUrd monomer solutions were then mixed to afford a kUrd oligomer solution (60  $\mu L$ ) containing 0.25  $\mu M$  each kUrd and initiate the *trans*-splicing reaction. The solution was transferred quickly to the wells of a microplate reader with a mineral oil overlay. The plate was then incubated at 37  $^{\circ}C$  in a plate reader (Infinite F200 Pro; Tecan, Männedorf,

Switzerland), which was pre-warmed at 37 °C and recorded fluorescence intensity of the sample solution.

### 3. Results

#### 3.1. Design of Unit Ribozyme Dimers (Urds) with a Bent Linker

To design a new series of Urds to form closed trimers, we needed a suitable RNA motif consisting of two helices with fixed (~60°) angles as a bent linker unit. We selected the kink-turn RNA motif (Figure 1 and Figure S2) [30], which is one of the most extensively studied RNA motif families and provides a sharp (~60°) bend to the helical axis (Figure 2A). Kink-turn motifs have been identified in a diverse range of naturally occurring RNA structures with various functions [30]. Here, we chose the KT-15 kink-turn motif (Figure 2A), which not only serves as an RNA structural element but also acts as a recognition element for the ribosomal protein, L7Ae (Figure 2A) [30]. This RNA–protein complex has also been used as a module in synthetic RNA–protein nanostructures [33,34,38,39] and also synthetic RNA switches [40–43].



**Figure 2.** Structures of kUrds and their assembly analyzed by EMSA. (A) 3D structures of KT-15 kink-turn motif (orange) with a pair of duplexes (green) and its binding with L7Ae protein (blue). Sequence and secondary structures of the linker elements of kUrds with three distinct pairs of duplexes. Small red and green icons indicate RNA motifs. Mutations in L2 elements to disrupt RzM:RzM interaction are also shown as elliptic gray icons. Sequences and secondary structures of these RNA motifs are shown in Figure S1. (B) EMSA of 1-K-1' homopolymers formed in the presence of 15 mM Mg<sup>2+</sup>. Asterisks indicate RNA chains labeled with BODIPY fluorophore. (C) EMSA of 1-K-1' homopolymers formed in the presence of 20 mM Mg<sup>2+</sup>. Asterisks indicate RNA chains labeled with BODIPY fluorophore.

### 3.2. Homooligomerization of kUrds

A new class of Urd variants bearing the KT-15 motif was designated as kink-turn unit ribozyme dimer (abbreviated as kUrd, Figure 1). To optimize kUrd linkers, we compared three linker elements, each of which had the KT-15 motif flanked by a pair of duplexes with 15 bp/10 bp, 15 bp/9 bp, or 14 bp/9 bp (Figure 2A). The three distinct kUrd linkers were used to connect two RzM monomers (RzM-1 and RzM-1') in a tail-to-tail manner, yielding three **1-K-1'** RNAs (Figure 2A and Figure S2). The three **1-K-1'** were designed to homotrimerize to form closed trimers (Figure 1). For each **1-K-1'**, we prepared **1x-K-1'** and **1-K-1x'** to eliminate assembly ability of the RzM on one side by introducing a UUCG tetraloop [44,45] into the L2 element (where x designates L2 UUCG, see Figures S1D and S3). We also prepared **1x-K-1x'** mutant for each **1-K-1'** to eliminate assembly ability of RzMs on both sides (Figure S3). We analyzed three **1-K-1'** kUrds with distinct pairs of helices by electrophoretic mobility shift assay (EMSA) to characterize their assembly properties. Each **1-K-1'** kUrd was composed of two RNA strands ( $\alpha$ - and  $\beta$ -RNA chains) (Figures S2B and S3, Table S1). Preliminary experiment showed that the two RNA strands assembled efficiently to form the kUrd monomer (Figure S4). We first labeled the 3'-end of the  $\beta$ -chain RNA with the BODIPY fluorophore [32] to visualize kUrd and its assembly in native gels in the presence of 10 mM Mg<sup>2+</sup> (Figure S5A), 15 mM Mg<sup>2+</sup> (Figure 2B) and 20 mM Mg<sup>2+</sup> (Figure 2C).

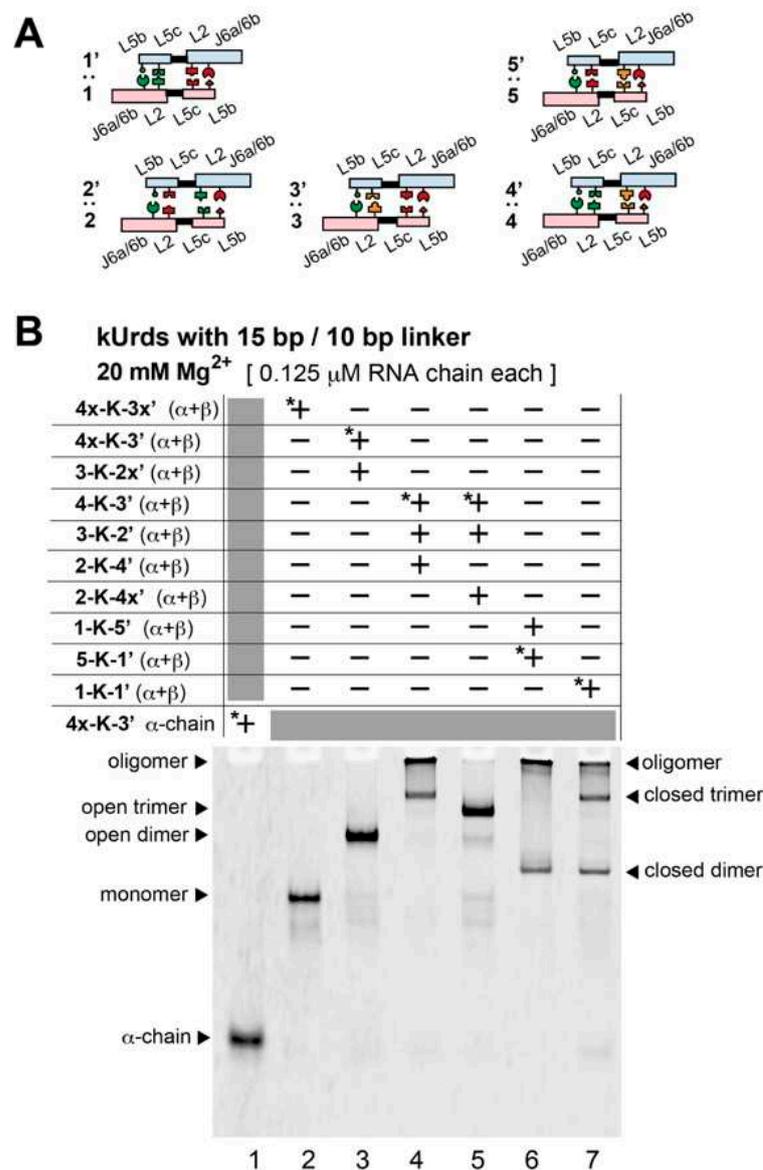
BODIPY-labeled  $\beta$ -chain RNA of **1x-K-1x'**, which can also be used as  $\beta$ -chain RNA of **1-K-1x'**, showed broad and multiple bands (lanes 1, 5 and 9 in Figure 2B,C). The  $\beta$ -chain RNA, however, showed a sharp and retarded band in the presence of an equimolar amount of the partner  $\alpha$ -chain RNA (lanes 2, 6 and 10 in Figure 2B,C), indicating assembly of the  $\alpha$ - and  $\beta$ -chains RNAs to form a structured kUrd **1x-K-1x'** monomer. Introduction of two L2-UUCG loops into both of the RzMs in kUrd afforded **1x-K-1x'** to selectively form a monomeric state (lanes 2, 6 and 10 in Figure 2B,C). We then confirmed the assembly state of **1-K-1'** and its mutants by preparing a solution containing equal amounts of four kUrds (**1x-K-1x'**, **1x-K-1'**, **1-K-1x'** and **1-K-1'**), prepared by mixing two  $\alpha$ -chain RNAs and two  $\beta$ -chain RNAs in a single tube (lanes 3, 7 and 11 in Figure 2B,C). BODIPY-labeled  $\beta$ -chain RNAs for **1x-K-1x'** and **1-K-1x'** selectively visualized kUrd monomer (**1x-K-1x'**), kUrd dimer (**1x-K-1':1-K-1x'**) and kUrd open trimer (**1x-K-1':1-K-1':1-K-1x'**). In addition, to kUrd monomers, two new bands with lower mobilities were observed (lanes 3, 7 and 11 in Figure 2B,C), which were expected to be a kUrd dimer consisting mainly of **1x-K-1':1-K-1x'** and a kUrd open trimer consisting mainly of **1x-K-1':1-K-1':1-K-1x'**.

We then analyzed the solution containing only **1-K-1'** (lanes 4, 8 and 12 in Figure 2B,C), which was expected to form homooligomers, including a closed form homotrimer. The dominant band was retained in the gel slot, suggesting the formation of high-molecular weight oligomers. Two new bands, which were not seen in lanes 3, 7 and 11, were also observed. The lower mobility band would correspond to the desired closed trimer containing six ribozyme units (lanes 4, 8 and 12 in Figure 2B,C). The higher mobility band could be a closed dimer of Urds, which was not predicted in molecular modeling of **1-K-1'**. In the presence of 10 mM Mg<sup>2+</sup>, the unit kUrd appeared to form efficiently (lanes 2, 6 and 10 in Figure S5A, see also Figure S5B,C) but oligomeric states of kUrds were formed less efficiently than those with 15 mM Mg<sup>2+</sup> (Figure S5B,C). Formation of oligomeric states of kUrds seemed enhanced in the presence of 20 mM Mg<sup>2+</sup> compared to 15 mM Mg<sup>2+</sup> (Figure S5B). Among the three kUrds with distinct pairs of duplexes (15 bp/10 bp, 15 bp/9 bp and 14bp/9 bp) in the kink-turn linker elements, the 15 bp/10 bp and 15 bp/9 bp linkers worked better than the 14 bp/9 bp linker, which produced multiple bands in the presence of 20 mM Mg<sup>2+</sup> (Figure 2C). In the following experiments, we used the 15 bp/10 bp pair as the kink-turn linker element.

### 3.3. Block Copolymerization of kUrds

Based on the homooligomerization of **1-K-1'** forming the desired closed trimer and unexpected closed dimer, we then examined selective formation of a closed trimer and

a closed dimer through copolymerization of two or three different kUrds (Figure 1). We substituted two RzMs in 1-K-1' to prepare a trio of Urds (3-K-2', 2-K-4' and 4-K-3') for triblock-type kUrd copolymer. For this purpose, we introduced three RzM:RzM interactions 2:2', 3:3' and 4:4' (Figure 3A and Figure S1D), which were designed to be orthogonal to one another in their recognition specificity [28]. In a similar manner, a pair of Urds (1-K-5' and 5-K-1') were also designed for diblock-type kUrd copolymer (Figure 3A and Figure S1D), which would selectively yield a closed dimer. In block copolymerization by two or three distinct kUrds, we first prepared each kUrd separately by assembling its  $\alpha$ - and  $\beta$ -RNA chain RNAs (Figures S2B and S4). Solutions of preformed kUrds were then mixed to form block copolymers of kUrds.



**Figure 3.** Assembly of a trio of kUrds to form a closed heterotrimer and a pair of kUrds to form a closed heterodimer. **(A)** Five distinct pairs of RzM:RzM interfaces, three of which were used in a closed heterotrimer configuration and the remaining two of which were used in a closed heterodimer configuration. Sequences and secondary structures of RNA motifs specifying RzM:RzM interfaces are shown in Figure S3. **(B)** EMSA of triblock and diblock kUrd copolymers formed in the presence of 20 mM Mg<sup>2+</sup>. Asterisks indicate RNA chains labeled with BODIPY fluorophore.

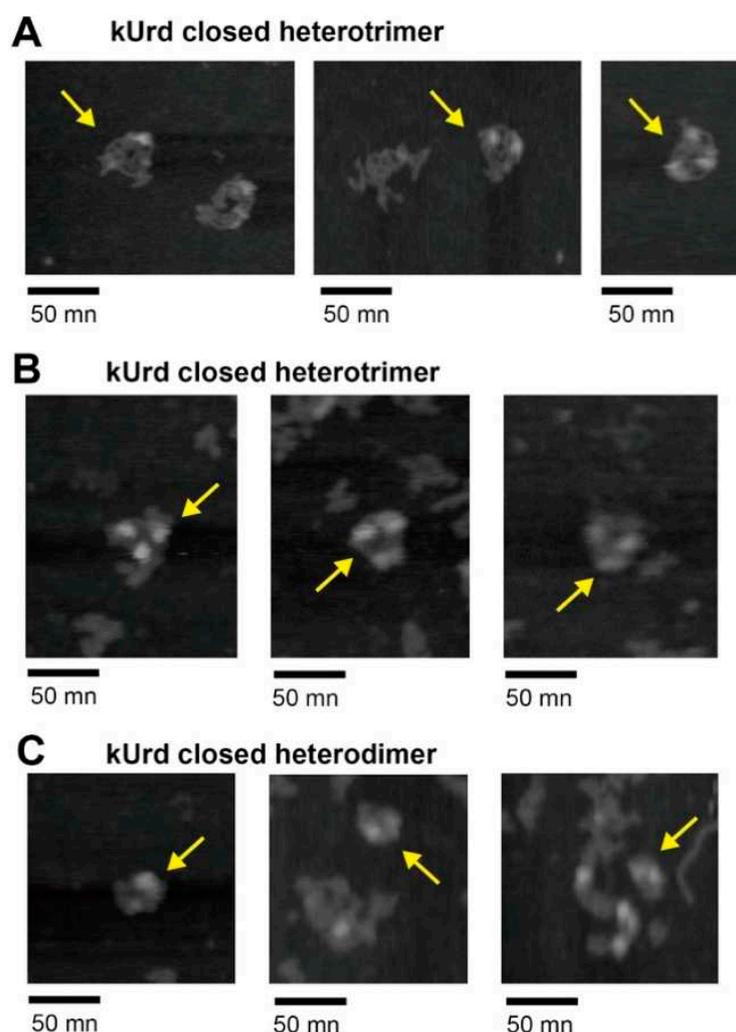
We first examined copolymerization of **3-K-2'**, **2-K-4'** and **4-K-3'** to obtain the selective formation of a closed trimer (Figure 3B). To distinguish the closed trimer from the open trimer, we also examined **4-K-3x'**, with which ring closure would be blocked by disruption of the **3:3'** interface. As seen in lane 4 in Figure 3B, assembly of three kUrds provided a migrating band the mobility of which was close to that of the band corresponding to the possible closed trimer in homopolymerization of **1-K-1'** (lane 7 in Figure 3B). Higher oligomerization of kUrds seen in **1-K-1'** homopolymers (lane 7 in Figure 3B) was still seen in triblock copolymers (lane 4 in Figure 3B). In the presence of **4-K-3x'** in place of **4-K-3'**, no band was retained in the sample well (lane 5 in Figure 3B), consistent with the molecular design to form the open block trimer (**3-K-2':2-K-4':4-K-3x'**) lacking further oligomerization ability. Selective formation of the open form of kUrd trimer was suggested by the single migrating band (lane 5 in Figure 3B), which migrated slower than the open kUrd dimer (**4x-K-3:3-K-2x'**, see lane 3 in Figure 3B) but faster than the possible closed trimers (lanes 5 and 7 in Figure 3B). Importantly, in the trio of kUrd designed for triblock copolymer (lane 4 in Figure 3B), no band corresponding to the possible closed dimer of **1-K-1'** (see lane 7 in Figure 3B) was observed. Bands corresponding to the open dimer and monomer were not seen in lane 4 (Figure 3B), suggesting that triblock oligomerization proceeded efficiently. Comparison between two trios of kUrds with/without ring closure ability (lanes 4 and 5 in Figure 3B) suggested that the closed trimer is thermodynamically stable and the dissociation of RzM interfaces in the closed trimer occurred poorly.

The pair of kUrds for diblock copolymer also formed oligomers with no migration (lane 6 in Figure 3B). This mixture also yielded a band the mobility of which corresponded to the possible closed dimer of **1-K-1'** (lanes 6 and 7 in Figure 3B). No band was seen in lane 6 at the position corresponding to that of the possible closed trimer. In the closed heterotrimer consisting of three distinct kUrds, each kUrd was formed by  $\alpha$ - and  $\beta$ -chain RNAs. As the kUrd closed heterotrimer contained six distinct RNA chains, we labeled one of six RNA chains with BODIPY fluorophore and monitored the mobility of the closed trimer on native gels (Figure S6). Regardless of the identity of the BODIPY-labeled RNA chain, all EMSA showed similar behavior, supporting the formation of the closed heterotrimer by assembly of three kUrds.

### 3.4. Atomic Force Microscopy Analysis of kUrd Heterotrimer and Heterodimer

Based on the observation that three kUrds are likely to form closed trimers in both homopolymerization and triblock copolymerization, we visualized the molecular structures of the heterotrimer (**3-K-2':2-K-4':4-K-3'**) and heterodimer (**1-K-5':5-K-1'**) by atomic force microscopy (AFM) [33,37–39]. In the presence of 17.5 mM  $Mg^{2+}$ , a sample containing three kUrds gave AFM images containing objects with triangular shapes (Figure 4A,B). The triangular objects were observed with limited frequency. We measured lengths of 21 sides of seven triangular objects observed in AFM images. While their lengths varied between 28.1 nm–39.6 nm, their average value was 34.5 nm. This value was closely similar to that predicted by molecular modeling of the closed trimer (approximately 34 nm, Figure 1). In some of the AFM images of kUrd triangles, we observed triangular shapes as three bright regions (Figure 4B), heights of which were 6–8 nm (Figure S7A).

We then analyzed AFM images of a solution containing **1-K-5'** and **5-K-1'** that formed a closed heterodimer, **1-K-5':5-K-1'** (Figure 4C). Due to the limited efficacy of the closed dimer formation (Figure 3B), AFM provided the corresponding images inefficiently. A few RNA objects, however, were found to reflect closed dimers (Figure 4C), in which two bright spots were observed in each object. Heights of bright spots were 5–7 nm (Figure S7B). We also analyzed AFM images of two mutant kUrds with disrupted L2 interacting motifs (Figure S8). They retained some of RNA interacting motifs but were not able to form kUrd dimers (Figure S8A). The two mutant kUrds formed some aggregates on mica surfaces but gave no bright spots in the presence of 17.5 mM  $Mg^{2+}$  (Figure S8C,D).

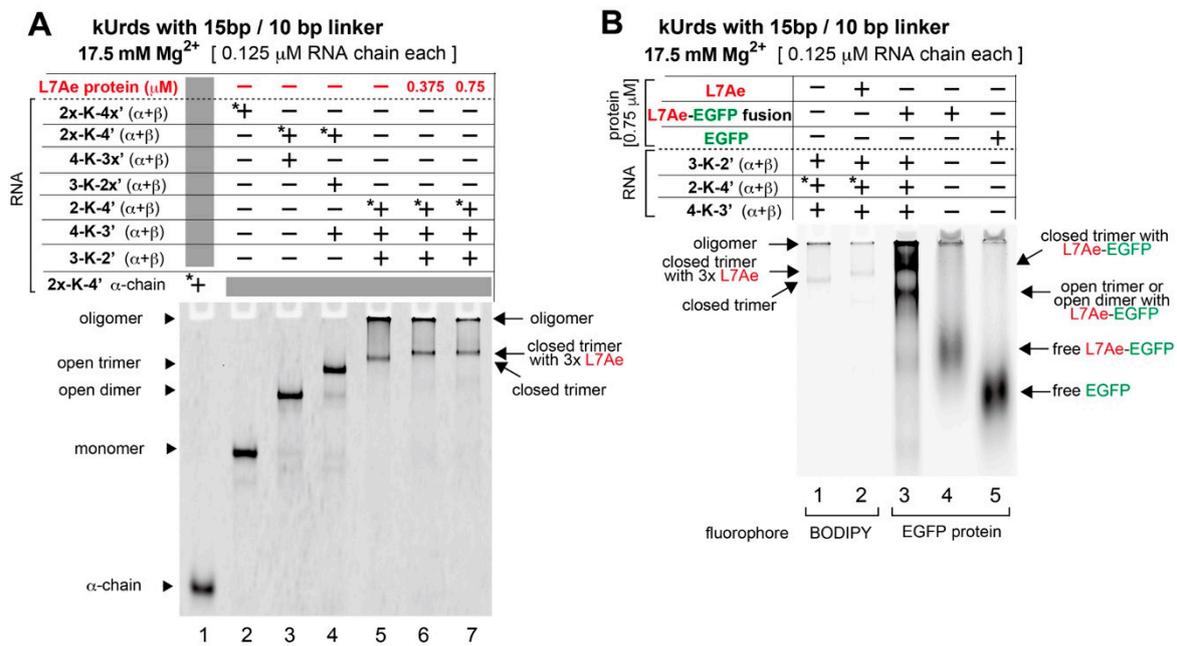


**Figure 4.** AFM imaging of closed kUrd heterotrimers and closed heterodimers. (A,B) AFM of closed kUrd trimers in the presence of 17.5 mM  $Mg^{2+}$ . Yellow arrows indicate closed kUrd trimers. (C) AFM of closed kUrd dimers in the presence of 17.5 mM  $Mg^{2+}$ . Yellow arrows indicate closed kUrd dimers.

### 3.5. Decoration of the kUrd Closed Trimer with Kink-Turn Binding Protein L7Ae

We then decorated the kUrd closed trimer with the RNA binding protein, L7Ae, for which the KT-15 kink-turn motif serves as a binding element (Figure 2A) [30]. We performed EMSA of the heterotrimer (3-K-2':2-K-4':4-K-3') in the absence and presence of L7Ae protein. In the presence of each RNA chain at 0.125  $\mu$ M (2.5 pmol each in 20  $\mu$ L of solution), theoretical amounts of the KT-15 motif in the resulting three distinct kUrd units (7.5 pmol in 20  $\mu$ L of solution) was also 0.375  $\mu$ M (7.5 pmol in 20  $\mu$ L of solution).

In the presence of 0.375  $\mu$ M L7Ae protein, slight but distinct retardation of the closed trimer was observed (lane 6 in Figure 5A). A twofold molar excess amount of L7Ae protein (0.75  $\mu$ M) over the theoretical amount of kink-turn motif (0.375  $\mu$ M) did not cause further retardation of the band (lane 7 in Figure 5A). The mobility of the trimer band did not change in the presence of a greater excess of L7Ae protein (Figure S9). These observations suggested that the retarded band corresponded to the RNA–protein complex consisting of the kUrd closed trimer possessing three KT-15 kink-turns and three L7Ae protein molecules. In the kUrd trimer–L7Ae complex, fluorescent visualization of L7Ae proteins was also examined using an engineered L7Ae protein fused to enhanced green fluorescent protein (EGFP) (Figure 5B) [46].



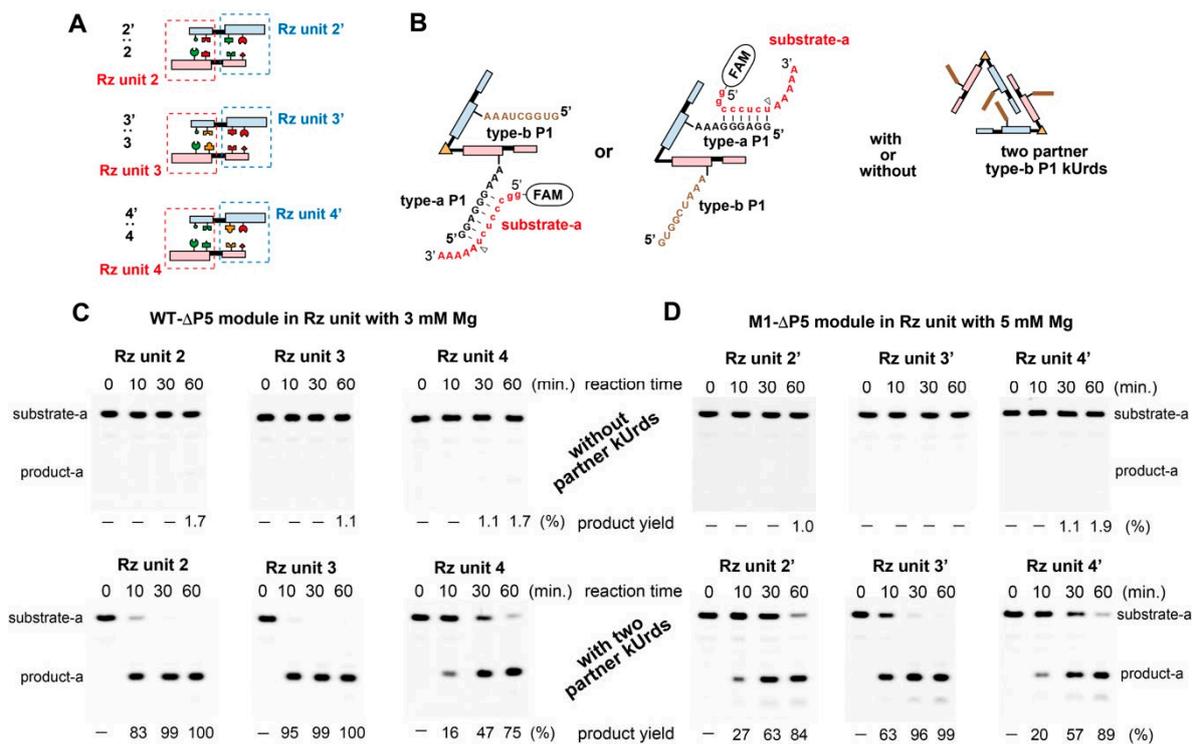
**Figure 5.** Complex formation of L7Ae protein with KT-15 kink-turn motifs in a closed kUrd heterotrimer. (A) EMSA of a closed kUrd heterotrimer in the absence and presence of L7Ae protein. (B) EMSA of a closed kUrd heterotrimer in the absence and presence of L7Ae protein and L7Ae-EGFP fusion protein.

L7Ae-EGFP fusion protein, which showed a broad band in free form (lane 4 in Figure 5B), shifted significantly in the presence of the kUrd heterotrimer (lane 3 in Figure 5B). In lane 3, the major band migrated more slowly than the free kUrd heterotrimer (lane 1) and also more slowly than the complex of closed trimer and L7Ae protein (lane 2). The major band in lane 3 corresponded to a complex of the closed trimer and L7Ae-EGFP fusion protein. The faster migrating band in lane 3 may correspond to an RNA-protein complex containing an open trimer or an open dimer because it was observed very weakly as a byproduct in assembly of the closed trimer (lane 4 in Figure S9B).

### 3.6. Catalytic Activities of kUrds and Their Oligomers

Although closed trimers and closed dimers were not formed exclusively without higher oligomers, it was interesting to see the catalytic abilities of ribozyme units in kUrds because their assembly would produce an active ribozyme unit consisting of the ΔP5 core ribozyme module and P5abc activator module (Figures 1 and 6A) [18].

As the ΔP5 core ribozyme can be activated by transient (weak) association of the P5abc module, activity assay required a lower Mg<sup>2+</sup> concentration than EMSA [36]. In the substrate cleavage reactions by the wild-type ΔP5 module (Figure S1A) and M1 type ΔP5 module (Figure S1B), the efficiencies of P5abc-dependent activation were optimum in the presence of 3 mM Mg<sup>2+</sup> and 5 mM Mg<sup>2+</sup>, respectively. In the closed heterotrimer, six ΔP5-P5abc ribozyme units were formed, in which three had the wild-type ΔP5 modules (units 2, 3 and 4) with the remaining three having the M1 type ΔP5 module (units 2', 3' and 4') (Figure 6A). We examined activation of the ΔP5 core ribozyme by the P5abc module upon formation of kUrd oligomers including the closed trimers.



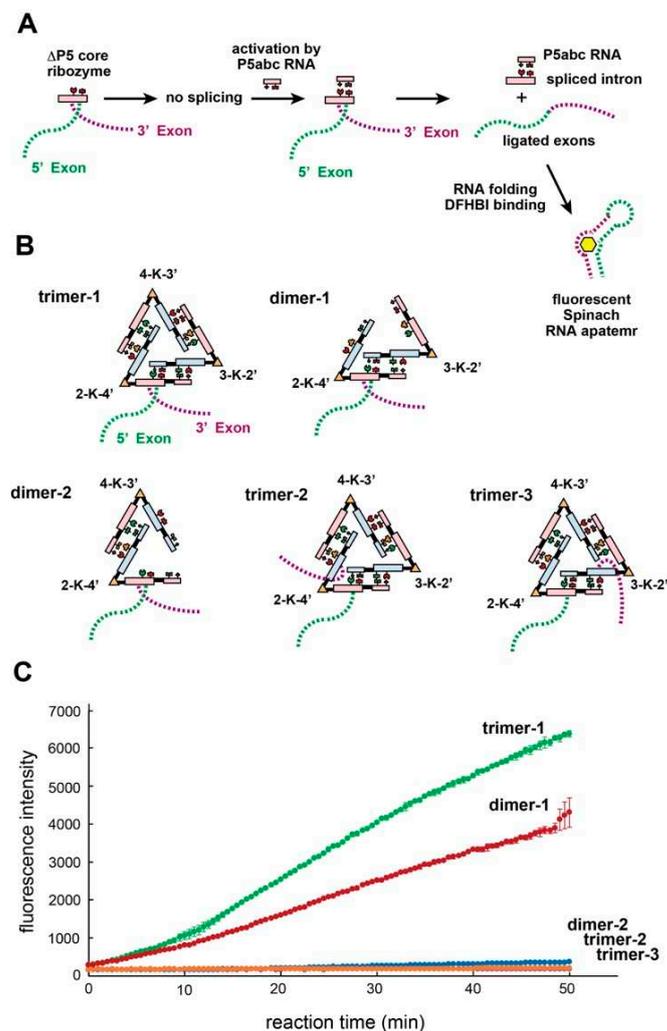
**Figure 6.** Activation of ribozyme unit upon assembly of kUrds. (A) Formation of two ribozyme units in each RzM:RzM interface. One ribozyme unit (2, 3, or 4) formed the wild-type ribozyme and the other ribozyme unit (2', 3', or 4') formed the M1 type ribozyme. (B) Sequences of substrate-a and two types of P1 binding elements. Location of P1 element in the secondary structure of the *Tetrahymena* group I ribozyme are shown in Figure S1. (C) Cleavage of substrate-a by the wild-type ribozyme unit in a given kUrd without (top) or with (bottom) its partner kUrds in the presence of 3 mM Mg<sup>2+</sup>. (D) Cleavage of substrate-a by the M1 type ribozyme unit in a given kUrd without (top) or with (bottom) its partner kUrds in the presence of 5 mM Mg<sup>2+</sup>.

We developed an assay system with which each  $\Delta P5$ -P5abc ribozyme unit in kUrd oligomers could be analyzed separately. In the  $\Delta P5$ -P5abc ribozyme unit, the P1 substrate recognition element for the cleavage reaction could be introduced in a modular manner. In EMSA and AFM analysis,  $\alpha$ - and  $\beta$ -RNA chain RNAs commonly possessed a P1 recognition element for type-a substrate. To analyze the single ribozyme units in kUrds and in their oligomers selectively, we employed the second P1 element recognizing type-b substrate (Figure 6B) [47]. The substrate recognition of type-a and type-b P1 elements are highly orthogonal to each other [37]. The type-a substrate-P1 pair was introduced selectively to one of six ribozyme units, while the remaining five units were designed to possess type-b P1 elements (Figure 6B). The cleavage reaction of type-a substrate thus, enabled us to selectively monitor the catalytic ability of a ribozyme unit with a type-a P1 element (Figure 6B).

In the isolated state of each kUrd, its wild-type  $\Delta P5$  ribozyme module possessing the type-a P1 element showed no activity in the presence of 3 mM Mg<sup>2+</sup> (Figure 6C, top). Addition of the partner kUrds in which the  $\Delta P5$  ribozyme module possessed only type-b P1 elements, however, induced the catalytic ability to cleave the type-a substrate (Figure 6C, bottom). In a similar manner, the M1 type ribozyme unit in each kUrd was also inactive with 5 mM Mg<sup>2+</sup> (Figure 6D, top) but activated by the addition of partner kUrds (Figure 6D, bottom). Such activation behavior was primarily conducted by the association of  $\Delta P5$  catalytic modules in each kUrd with P5abc activator modules with the partner kUrds. It should be noted that moderate differences were observed in the catalytic activities among the three ribozyme units sharing the wild-type (unit 2, 3 and 4) or M1 type (unit 2', 3' and 4')  $\Delta P5$  module (Figure 6C,D). The catalytic activity of the  $\Delta P5$ -P5abc ribozyme unit was governed by the identity of the  $\Delta P5$  module and also by the identity of the RNA–RNA interface between the  $\Delta P5$  module and the P5abc module (Figure 6A) [35,36].

### 3.7. *Trans*-Splicing Reaction Promoted by *kUrd* Oligomer Formation

The intron form of  $\Delta P5$  core ribozyme module can perform the self-splicing reaction to yield ligated exons in the presence of the P5abc activator module (Figure 7A) [18]. The pair of  $\Delta P5$  core ribozyme modules in each *kUrd* were composed of  $\alpha$ - and  $\beta$ -chain RNAs. In the intron form of wild-type  $\Delta P5$  core ribozyme module on one side of *kUrd*, 5' and 3' exons were provided by  $\alpha$ - and  $\beta$ -chain RNAs, respectively (Figure S10A). In the intron form of the M1 type  $\Delta P5$  core ribozyme module on the opposite side of *kUrd*, 5' and 3' exons were provided by  $\beta$ - and  $\alpha$ -chain RNAs, respectively (Figure S6B). The organization of  $\alpha$ - and  $\beta$ -chain RNAs in the intron form of the  $\Delta P5$  core ribozyme modules would conduct *trans*-splicing reaction joining 5' and 3' exons, which were provided by two separate RNA chains, to produce a mature exon sequence (Figure S10A,B) [31,48]. *Trans*-splicing promoted by group I ribozyme has been recognized as a promising strategy to repair mRNAs of disease-related genes bearing undesirable mutations [48–50]. Therefore, we examined *trans*-splicing reaction catalyzed by a *kUrd* in triblock copolymers (3-K-2':2-K-4':4-K-3') including the closed trimer.



**Figure 7.** *Trans*-splicing reaction by the ribozyme unit 2 formed in the closed *kUrd* heterotrimer. (A) Parent *cis*-splicing reaction of the bimolecular group I ribozyme. 5' and 3' exons were both provided by the unimolecular  $\Delta P5$  ribozyme RNA, the activity of which is repressed in the absence of the P5abc RNA. Splicing reaction produced Spinach RNA as the ligated exons. Spinach RNA captured DFHBI and induced its fluorescence. (B) Closed *kUrd* trimers and their partial dimers used for analysis of the *trans*-splicing reaction. (C) Time-dependent increases in fluorescence of Spinach RNA–DFHBI complex. Splicing reactions were carried out at 37 °C.

We installed 5' and 3' exon sequences and their recognition elements into the  $\alpha$ - and  $\beta$ - chain RNAs of **2-K-4'** kUrd, respectively (Figure S10C and Table S2). Exon sequences were appended to the ribozyme module consisting of the wild-type  $\Delta$ P5 ribozyme module activated by the P5abc module provided by **3-K-2'** kUrd through the **2:2'** interface (Figure 7B and Figure S10C). We used a shortened derivative of the Spinach RNA aptamer as a ligated exon sequence dissected in precursor form (Figure 7A) [51]. Spinach RNA is an RNA receptor for DFHBI and its fluorescence is induced by restricting the molecular motion of DFHBI in the complex form [52]. The progression of *trans*-splicing reaction can be monitored in the presence of DFHBI [51,53]. To evaluate the effects of kUrd oligomer formation on the *trans*-splicing reaction, we prepared a set of kUrd variants some of which possessed 5' and 3' exon sequences (Figure 7B and Figure S10, see also Table S2). A trio of kUrds (**3-K-2':2-K-4':4-K-3'**) to form a triblock oligomer (**trimer-1**) exhibited an increase in emission (green plot in Figure 7C). We also examined two sets of triblock oligomers (**trimer-2** and **trimer-3**) in which the 5' exon and 3' exon were installed to incorrect RNA chains and the two exons were unable to meet in one  $\Delta$ P5 ribozyme unit. The two sets of kUrds showed no increase in fluorescence from Spinach RNA complexed with DFHBI (purple plot and blue plot in Figure 7C). These data indicated that the correct assignments of 5' and 3' exons in six RNA chains are crucial in the oligomeric state of kUrds.

To evaluate the effects of triblock oligomerization (**3-K-2':2-K-4':4-K-3'**) in the splicing proficient trimer (**trimer-1**), we assayed a solution containing **2-K-4'** and **3-K-2'**, which form **dimer-1**, in which the intron form of the ribozyme unit in **2-K-4'** was formed correctly and **2-K-4'** was correctly activated by the P5abc module in **3-K-2'** (Figure 7B). The difference in activity between the presence and absence of the third kUrd unit (**4-K-3'**) mainly reflected the oligomerization ability. The solution of **dimer-1** exhibited emission (red plot in Figure 7C), which was, however, about two thirds that of the solution containing three kUrds. This observation suggested that **4-K-3'** improved the activity of the splicing ribozyme unit through the formation of oligomeric states. To examine the contribution of the P5abc module provided by **3-K-2'**, we depleted the **3-K-2'** unit from the three kUrd components. The sample solution containing **2-K-4'** and **4-K-3'** to form **dimer-2** showed no increase in emission (orange plot in Figure 7C), indicating the importance of activation by the P5abc module to conduct full splicing ability of the  $\Delta$ P5 ribozyme module possessing 5' and 3' exons.

#### 4. Discussion

In this study, we performed molecular modeling of kUrd to design a closed trimer, the formation of which was confirmed by EMSA and AFM analyses. On the other hand, significant formation of higher oligomers indicated that the current structural design was insufficient partly due to the 3D structure of the full-length *Tetrahymena* ribozyme used for the molecular design in this study. In molecular modeling of kUrds (a dimeric form of an engineered *Tetrahymena* ribozyme connected by the kink-turn motif), we used a model 3D structure of the full-length ribozyme and the X-ray crystallographic structure of a shortened form of the ribozyme lacking some peripheral elements [29]. There may have been some deviations from the full-length *Tetrahymena* ribozyme 3D structure, which has yet to be solved by X-ray crystallography. Improvement of 3D structural design, however, may be possible because a new 3D model structure of the full-length ribozyme was constructed recently by combining cryo-EM data, chemical probing data and molecular modeling [54]. Although the resolution of this new model is modest (6.8 Å), it will be useful to refine the design of ribozyme-based nanostructures.

The formation of the closed dimer is also an issue to be resolved for precise design of kUrd-based RNA nanostructures. Formation of the closed kUrd dimer, which was unexpected from molecular modeling, suggested the structural flexibility of kUrd RNAs and entropic effects in RNA assembly. The flexible nature of the kink-turn RNA motifs has been analyzed and reported previously [55–57]. Therefore, we hypothesized that the formation of the closed dimer may also be supported by the structural flexibility of the

KT-15 kink-turn motif. Distinct from our previous Urd with an extended linker, each kUrd can accept L7Ae protein at its linker element bearing the KT-15 kink-turn motif. EMSA of the heterotrimer in the presence of L7Ae suggested the formation of KT-15–L7Ae complex at each corner of the triangle-shaped closed kUrd trimer. We also performed AFM analysis of the heterotrimer in the presence of L7Ae but no significant differences were observed between images of kUrds with and without L7Ae. This was conceivable because L7Ae is much smaller than kUrd RNA. We also found that L7Ae protein strongly inhibited the catalytic ability of kUrds. Collaboration between L7Ae protein and the *Tetrahymena* ribozyme unit in catalysis is still an open issue.

In this study, we mainly analyzed the closed trimer and closed dimer of kUrds. The closed trimer contained six ribozyme units and three kink-turn motifs that interacted with L7Ae proteins. In reactions promoted by the ribozyme unit in kUrd oligomers, *trans*-splicing reaction proceeded more efficiently in the presence of three kUrds forming a heterotrimer than two kUrds providing the core  $\Delta$ P5 ribozyme and P5abc activator. Although the extent of improvement was moderate, this was a promising result as an initial step toward RNA processing in vivo controlled by ribozyme-based nanostructures. The kUrd trimer containing six group I ribozyme units consists of nearly 2400 nucleotides, which is comparable to bacterial 16S and 23S ribosomal RNAs. The size of the triangle shape (with an edge of approximately 34 nm) was also comparable to the bacterial ribosome (approximately 20 nm in diameter). Although the structure and function of kUrd-based nanostructures are still primitive, this study represents an initial step toward artificial construction of complex RNA nanostructures capable of playing crucial roles in naturally occurring biological systems. *Trans*-splicing ribozymes have been investigated as a promising tool to repair mutated mRNAs causing genetic disorders [49,50]. Ribozyme dimers employed in this study may be used to repair two distinct mutated mRNA simultaneously and cooperatively. Two ribozyme units in a dimer also target the same mRNA sequence, to which two *trans*-splicing ribozyme units perform distinct splicing reactions to yield different products. For example, one ribozyme unit repairs the mutation of the target mRNA whereas the other ribozyme unit attaches a fluorescent RNA aptamer to the target mRNA for its fluorescent imaging. Fabrication of closed trimers using ribozyme dimers and KT-15 kink-turn would enable us to attach various protein components through L7Ae protein that binds to KT-15 [38,39]. Such ribozyme-based RNA-protein nanoparticle can be further decorated by RNA-based sensors and therapeutic RNA modules such as RNA aptamers and siRNAs, with which next generation of RNA-based nanomedicines may be developed.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/2076-3417/11/6/2583/s1>, Figure S1: Sequences and secondary structures of the *Tetrahymena* group I ribozymes and their structural elements used for modular engineering in this study, Figure S2: Scheme of rational redesign of unimolecular *Tetrahymena* ribozymes to generate kUrd 1-K-1', Figure S3: Kink-turn unit ribozyme dimers (kUrds) employed in this study, Figure S4: Formation of kUrds through the assembly of their  $\alpha$ -chain and  $\beta$ -chain RNAs, Figure S5: Oligomerization of 1-K-1' and its mutants, Figure S6: EMSA of triblock copolymers formed in the presence of 20 mM Mg<sup>2+</sup>, Figure S7: AFM images of the closed heterotrimer and heterodimer and their cross sections, Figure S8: AFM images of the closed heterodimer and its mutant monomers, Figure S9: Complex formation of L7Ae proteins with KT-15 kink-turn motifs in a closed kUrd heterotrimer, Figure S10: Closed kUrd trimers and their partial dimers used for analysis of the *trans*-splicing reaction.

**Author Contributions:** Conceptualization, Y.I.; methodology, Y.I., H.S. (Hirohide Saito), H.S. (Hiroshi Sugiyama), M.E. and S.M.; investigation, J.A., K.H. and Y.F.; writing—original draft preparation, Y.I. and T.Y.; writing—review and editing, H.S. (Hirohide Saito), H.S. (Hiroshi Sugiyama), M.E. and S.M.; supervision, Y.I., H.S. (Hirohide Saito), H.S. (Hiroshi Sugiyama) and M.E. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by University of Toyama Discretionary Funds of the President “Toyama RNA Collaborative Research” (to Y.I. and S.M.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

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

## References

- Goodsell, D.S.; Olson, A.J. Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105–153. [[CrossRef](#)]
- Lesieur, C. The assembly of protein oligomers: Old stories and new perspectives with graph theory. In *Oligomerization of Chemical and Biological Compounds*; Lesieur, C., Ed.; IntecOpen: London, UK, 2014. [[CrossRef](#)]
- Jones, S.; Thornton, J.M. Protein-protein interactions: A review of protein dimer structures. *Prog. Biophys. Mol. Biol.* **1995**, *63*, 31–65. [[CrossRef](#)]
- Glover, D.J.; Clark, D.S. Protein calligraphy: A new concept begins to take shape. *ACS Cent. Sci.* **2016**, *2*, 438–444. [[CrossRef](#)]
- Sun, H.; Luo, Q.; Hou, C.; Liu, J. Nanostructures based on protein self-assembly: From hierarchical construction to bioinspired materials. *Nano Today* **2017**, *14*, 16–41. [[CrossRef](#)]
- Kuan, S.L.; Bergamini, F.R.G.; Weil, T. Functional protein nanostructures: A chemical toolbox. *Chem. Soc. Rev.* **2018**, *47*, 9069–9105. [[CrossRef](#)] [[PubMed](#)]
- Boyken, S.E.; Chen, Z.; Groves, B.; Langan, R.A.; Oberdorfer, G.; Ford, A.; Gilmore, J.M.; Xu, C.; DiMaio, F.; Pereira, J.H.; et al. De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* **2016**, *352*, 680–687. [[CrossRef](#)]
- Fallas, J.A.; Ueda, G.; Sheffler, W.; Nguyen, V.; McNamara, D.E.; Sankaran, B.; Pereira, J.H.; Parmeggiani, F.; Brunette, T.J.; Cascio, D.; et al. Computational design of self-assembling cyclic protein homo-oligomers. *Nat. Chem.* **2017**, *9*, 353–360. [[CrossRef](#)]
- Kuhlman, B.; O'Neill, J.W.; Kim, D.E.; Zhang, K.Y.; Baker, D. Conversion of monomeric protein L to an obligate dimer by computational protein design. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10687–10691. [[CrossRef](#)]
- Stranges, P.B.; Machius, M.; Miley, M.J.; Tripathy, A.; Kuhlman, B. Computational design of a symmetric homodimer using beta-strand assembly. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20562–20567. [[CrossRef](#)] [[PubMed](#)]
- Mou, Y.; Huang, P.S.; Hsu, F.C.; Huang, S.J.; Mayo, S.L. Computational design and experimental verification of a symmetric protein homodimer. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10714–10719. [[CrossRef](#)]
- Fedor, M.J.; Williamson, J.R. The catalytic diversity of RNAs. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 399–412. [[CrossRef](#)]
- Weinberg, C.E.; Weinberg, Z.; Hammann, C. Novel ribozymes: Discovery, catalytic mechanisms, and the quest to understand biological function. *Nucleic Acids Res.* **2019**, *47*, 9480–9494. [[CrossRef](#)] [[PubMed](#)]
- Breaker, R.R. Prospects for riboswitch discovery and analysis. *Mol. Cell* **2011**, *4*, 867–879. [[CrossRef](#)]
- Butcher, S.E.; Pyle, A.M. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res.* **2011**, *44*, 1302–1311. [[CrossRef](#)] [[PubMed](#)]
- Ishikawa, J.; Fujita, Y.; Maeda, Y.; Furuta, H.; Ikawa, Y. GNRA/receptor interacting modules: Versatile modular units for natural and artificial RNA architectures. *Methods* **2011**, *54*, 226–238. [[CrossRef](#)]
- Jarrell, K.A.; Dietrich, R.C.; Perlman, P.S. Group II intron domain 5 facilitates a *trans*-splicing reaction. *Mol. Cell. Biol.* **1988**, *8*, 2361–2366. [[CrossRef](#)]
- van der Horst, G.; Christian, A.; Inoue, T. Reconstitution of a group I intron self-splicing reaction with an activator RNA. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 184–188. [[CrossRef](#)]
- Engelhardt, M.A.; Doherty, E.A.; Knitt, D.S.; Doudna, J.A.; Herschlag, D. The P5abc peripheral element facilitates preorganization of the *Tetrahymena* group I ribozyme for catalysis. *Biochemistry* **2000**, *39*, 2639–2651. [[CrossRef](#)]
- Doudna, J.A.; Cech, T.R. Self-assembly of a group I intron active site from its component tertiary structural domains. *RNA* **1995**, *1*, 36–45.
- Ikawa, Y.; Shiraiishi, H.; Inoue, T. *Trans*-activation of the *Tetrahymena* ribozyme by its P2-2.1 domains. *J. Biochem.* **1998**, *123*, 528–533. [[CrossRef](#)]
- Rahman, M.M.; Matsumura, S.; Ikawa, Y. Effects of molecular crowding on a bimolecular group I ribozyme and its derivative that self-assembles to form ribozyme oligomers. *Biochem. Biophys. Res. Commun.* **2018**, *507*, 136–141. [[CrossRef](#)]
- Pan, T. Higher order folding and domain analysis of the ribozyme from *Bacillus subtilis* ribonuclease P. *Biochemistry* **1995**, *34*, 902–909. [[CrossRef](#)]
- Loria, A.; Pan, T. Domain structure of the ribozyme from eubacterial ribonuclease P. *RNA* **1996**, *2*, 551–563.
- Kim, H.; Poelling, R.R.; Leeper, T.C.; Meyer, M.A.; Schmidt, F.J. In vitro transactivation of *Bacillus subtilis* RNase P RNA. *FEBS Lett.* **2001**, *506*, 235–238. [[CrossRef](#)]
- Tanaka, T.; Matsumura, S.; Furuta, H.; Ikawa, Y. Tecto-GIRz: Engineered group I ribozyme the catalytic ability of which can be controlled by self-dimerization. *ChemBioChem* **2016**, *17*, 1448–1455. [[CrossRef](#)] [[PubMed](#)]
- Tanaka, T.; Ikawa, Y.; Matsumura, S. Rational engineering of a modular group I ribozyme to control its activity by self-dimerization. *Methods Mol. Biol.* **2017**, *1632*, 325–340. [[PubMed](#)]

28. Kiyooka, R.; Akagi, J.; Hidaka, K.; Sugiyama, H.; Endo, M.; Matsumura, S.; Ikawa, Y. Catalytic RNA nano-objects formed by self-assembly of group I ribozyme dimers serving as unit structures. *J. Biosci. Bioeng.* **2020**, *130*, 253–259. [[CrossRef](#)] [[PubMed](#)]
29. Guo, F.; Gooding, A.R.; Cech, T.R. Structure of the Tetrahymena ribozyme: Base triple sandwich and metal ion at the active site. *Mol. Cell* **2004**, *16*, 351–362.
30. Klein, D.J.; Schmeing, T.M.; Moore, P.B.; Steitz, T.A. The kink-turn: A new RNA secondary structure motif. *EMBO J.* **2001**, *20*, 4214–4221. [[CrossRef](#)] [[PubMed](#)]
31. Tanaka, T.; Hirata, Y.; Tominaga, Y.; Furuta, H.; Matsumura, S.; Ikawa, Y. Heterodimerization of group I ribozymes enabling exon recombination through pairs of cooperative *trans*-splicing reactions. *ChemBioChem* **2017**, *18*, 1659–1667. [[CrossRef](#)]
32. Ikawa, Y.; Moriyama, S.; Furuta, H. Facile syntheses of BODIPY derivatives for fluorescent labeling of the 3' and 5' ends of RNAs. *Anal. Biochem.* **2008**, *378*, 166–170. [[CrossRef](#)]
33. Ohno, H.; Kobayashi, T.; Kabata, R.; Endo, K.; Iwasa, T.; Yoshimura, S.H.; Takeyasu, K.; Inoue, T.; Saito, H. Synthetic RNA-protein complex shaped like an equilateral triangle. *Nat. Nanotechnol.* **2011**, *6*, 116–120. [[CrossRef](#)] [[PubMed](#)]
34. Ohuchi, S.J.; Sagawa, F.; Sakamoto, T.; Inoue, T. A trifunctional, triangular RNA-protein complex. *FEBS Lett.* **2015**, *589*, 2424–2428. [[CrossRef](#)] [[PubMed](#)]
35. Tanaka, T.; Furuta, H.; Ikawa, Y. Installation of orthogonality to the interface that assembles two modular domains in the Tetrahymena group I ribozyme. *J. Biosci. Bioeng.* **2014**, *117*, 407–412. [[CrossRef](#)]
36. Rahman, M.M.; Matsumura, S.; Ikawa, Y. Artificial RNA motifs expand the programmable assembly between RNA modules of a bimolecular ribozyme leading to application to RNA nanostructure design. *Biology* **2017**, *6*, 37. [[CrossRef](#)] [[PubMed](#)]
37. Oi, H.; Fujita, D.; Suzuki, Y.; Sugiyama, H.; Endo, M.; Matsumura, S.; Ikawa, Y. Programmable formation of catalytic RNA triangles and squares by assembling modular RNA enzymes. *J. Biochem.* **2017**, *161*, 451–462. [[CrossRef](#)]
38. Fujita, Y.; Furushima, R.; Ohno, H.; Sagawa, F.; Inoue, T. Cell-surface receptor control that depends on the size of a synthetic equilateral-triangular RNA-protein complex. *Sci. Rep.* **2014**, *4*, 6422. [[CrossRef](#)]
39. Osada, E.; Suzuki, Y.; Hidaka, K.; Ohno, H.; Sugiyama, H.; Endo, M.; Saito, H. Engineering RNA-protein complexes with different shapes for imaging and therapeutic applications. *ACS Nano.* **2014**, *8*, 8130–8140. [[CrossRef](#)] [[PubMed](#)]
40. Saito, H.; Kobayashi, T.; Hara, T.; Fujita, Y.; Hayashi, K.; Furushima, R.; Inoue, T. Synthetic translational regulation by an L7Ae-kink-turn RNP switch. *Nat. Chem. Biol.* **2010**, *6*, 71–78. [[CrossRef](#)]
41. Wroblewska, L.; Kitada, T.; Endo, K.; Siciliano, V.; Stillo, B.; Saito, H.; Weiss, R. Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery. *Nat. Biotechnol.* **2015**, *33*, 839–841. [[CrossRef](#)]
42. Matsuura, S.; Ono, H.; Kawasaki, S.; Kuang, Y.; Fujita, Y.; Saito, H. Synthetic RNA-based logic computation in mammalian cells. *Nat. Commun.* **2018**, *9*, 4847. [[CrossRef](#)]
43. Ohno, H.; Akamine, S.; Saito, H. Synthetic mRNA-based systems in mammalian cells. *Adv. Biosyst.* **2020**, *4*, e1900247. [[CrossRef](#)] [[PubMed](#)]
44. Ennifar, E.; Nikulin, A.; Tishchenko, S.; Serganov, A.; Nevskaya, N.; Garber, M.; Ehresmann, B.; Ehresmann, C.; Nikonov, S.; Dumas, P. The crystal structure of UUCG tetraloop. *J. Mol. Biol.* **2000**, *304*, 35–42. [[CrossRef](#)]
45. Hall, K.B. Mighty tiny. *RNA* **2015**, *21*, 630–631. [[CrossRef](#)] [[PubMed](#)]
46. Zhang, G.; Gurtu, V.; Kain, S.R. An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem. Biophys. Res. Commun.* **1996**, *227*, 707–711. [[CrossRef](#)]
47. Campbell, T.B.; Cech, T.R. Mutations in the Tetrahymena ribozyme internal guide sequence: Effects on docking of the P1 helix into the catalytic core and correlation with catalytic activity. *Biochemistry* **1996**, *35*, 11493–11502. [[CrossRef](#)]
48. Ikawa, Y.; Matsumura, S. Engineered group I ribozymes as RNA-based modular tools to control gene expression. In *Applied RNA Bioscience*; Masuda, S., Izawa, S., Eds.; Springer: Berlin, Germany, 2018; pp. 203–220.
49. Müller, U.F. Design and experimental evolution of *trans*-splicing group I intron ribozymes. *Molecules* **2017**, *22*, 75. [[CrossRef](#)]
50. Lee, C.H.; Han, S.R.; Lee, S.W. Therapeutic applications of group I intron-based *trans*-splicing ribozymes. *Wiley Interdiscip. Rev. RNA* **2018**, *9*, e1466. [[CrossRef](#)]
51. Furukawa, A.; Maejima, T.; Matsumura, S.; Ikawa, Y. Characterization of an RNA receptor motif that recognizes a GCGA tetraloop. *Biosci. Biotechnol. Biochem.* **2016**, *80*, 1386–1389. [[CrossRef](#)] [[PubMed](#)]
52. Paige, J.S.; Wu, K.Y.; Jaffrey, S.R. RNA mimics of green fluorescent protein. *Science* **2011**, *333*, 642–646. [[CrossRef](#)]
53. Furukawa, A.; Tanaka, T.; Furuta, H.; Matsumura, S.; Ikawa, Y. Use of a fluorescent aptamer RNA as an exonic sequence to analyze self-splicing ability of a group I intron from structured RNAs. *Biology* **2016**, *5*, 43. [[CrossRef](#)] [[PubMed](#)]
54. Kappel, K.; Zhang, K.; Su, Z.; Watkins, A.M.; Kladwang, W.; Li, S.; Pintilie, G.; Topkar, V.V.; Rangan, R.; Zheludev, I.N.; et al. Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **2020**, *17*, 699–707. [[CrossRef](#)] [[PubMed](#)]
55. Matsumura, S.; Ikawa, Y.; Inoue, T. Biochemical characterization of the kink-turn RNA motif. *Nucleic Acids Res.* **2003**, *31*, 5544–5551. [[CrossRef](#)]
56. Goody, T.A.; Melcher, S.E.; Norman, D.G.; Lilley, D.M. The kink-turn motif in RNA is dimorphic, and metal ion-dependent. *RNA* **2004**, *10*, 254–264. [[CrossRef](#)] [[PubMed](#)]
57. Shi, X.; Huang, L.; Lilley, D.M.; Harbury, P.B.; Herschlag, D. The solution structural ensembles of RNA kink-turn motifs and their protein complexes. *Nat. Chem. Biol.* **2016**, *12*, 146–152. [[CrossRef](#)]

Article

# Chloro-Substituted Naphthyridine Derivative and Its Conjugate with Thiazole Orange for Highly Selective Fluorescence Sensing of an Orphan Cytosine in the AP Site-Containing Duplexes

Chun-xia Wang <sup>†</sup>, Yusuke Sato , Takashi Sugimoto, Norio Teramae and Seiichi Nishizawa \*

Department of Chemistry, Graduate School of Science, Tohoku University, Sendai 980-8578, Japan; cxwang@iccas.ac.cn (C.-x.W.); yusuke.sato.a7@tohoku.ac.jp (Y.S.); sgmt6600@gmail.com (T.S.); norio.teramae.d3@tohoku.ac.jp (N.T.)

\* Correspondence: seiichi.nishizawa.c8@tohoku.ac.jp

<sup>†</sup> College of New Energy and Materials, China University of Petroleum Beijing, Beijing 102249, China.

Received: 21 May 2020; Accepted: 13 June 2020; Published: 16 June 2020



**Abstract:** Fluorescent probes with the binding selectivity to specific structures in DNAs or RNAs have gained much attention as useful tools for the study of nucleic acid functions. Here, chloro-substituted 2-amino-5,7-dimethyl-1,8-naphthyridine (CINaph) was developed as a strong and highly selective binder for target orphan cytosine opposite an abasic (AP) site in the DNA duplexes. CINaph was then conjugated with thiazole orange (TO) via an alkyl spacer (CINaph–TO) to design a light-up probe for the detection of cytosine-related mutations in target DNA. In addition, we found the useful binding and fluorescence signaling of the CINaph–TO conjugate to target C in AP site-containing DNA/RNA hybrid duplexes with a view toward sequence analysis of microRNAs.

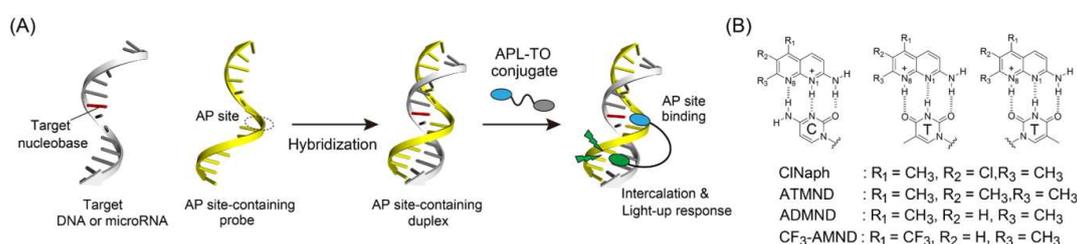
**Keywords:** fluorescent probe; conjugate; abasic site; DNA; microRNA

## 1. Introduction

Much attention has been paid to the design of fluorescent probes capable of selective binding to specific structures in DNAs and RNAs [1,2]. This class of fluorescent probes has been designed to bind abasic (apyrimidinic or apurinic; AP) sites [3–8], bulges [9,10], mismatched sites [11,12] and overhanging structures [13–15], by which the biological functions of these noncanonical base-pairs have been examined. Moreover, these fluorescent probes have great potential to be applied for the analysis of target DNAs and RNAs based on their binding-induced fluorescence signaling. These probes also serve as the affinity-labeling agents in the label-free assays for detecting various analytes [16–18].

We developed AP site-binding ligands (APLs) that are conjugated with thiazole orange (TO) for the fluorescence sensing of orphan nucleobases in DNA/DNA duplexes (Figure 1A) [19,20]. APLs can form the pseudo-base pairing with target orphan nucleobases, which allows the selective binding to the target nucleobase opposite the AP site [3–7]. On the other hand, TO unit connected with the APL unit through an appropriate linker can function as a fluorescent intercalator [21,22]. The TO unit alone shows negligible fluorescence, but its fluorescence is greatly enhanced upon intercalation into the base pairs near the AP site. APL–TO conjugates thus enable the light-up sensing of target orphan nucleobase in the AP site-containing DNA (AP–DNA) duplexes. Such light-up probes are useful for a more sensitive analysis compared to fluorescence quenching probes. We succeeded in the detection of single base mutation in target DNA sequences based on the binding and light-up functions of APL–TO conjugates in combination with an AP–DNA probe (cf. Figure 1A). In addition, these conjugates were applicable to the sequence-selective analysis of microRNAs, a class of small

non-coding RNAs associated with various diseases and cancers [23], by the selective binding of conjugates to target orphan nucleobases in the AP site-containing DNA/RNA (AP–DNA/RNA) hybrids. The properties of nucleobase-selectivity and fluorescence emission wavelength are tunable by adopting suitable APLs and cyanine dyes, respectively, which enables the development of multiplex analysis of target sequences by the simultaneous use of these conjugates in a single solution [19]. In this context, 2-amino-5,6,7-trimethyl-1,8-naphthyridine (ATMND, Figure 1B) was previously used as an APL unit in the conjugate with TO unit for the light-up sensing of target orphan C in the AP–DNA and AP–DNA/RNA hybrid duplexes [19]. The binding selectivity of ATMND toward target C can be rationalized by the complementary base pairing of the N1-protonated form of ATMND with target C through hydrogen bonding (Figure 1B) [3]. However, ATMND inherently has only moderate selectivity for C over T because of the possible protonation at N8 that allows the recognition of T (Figure 1B). This should limit the application of the conjugate for the sequence analysis of target DNAs and microRNAs such as the discrimination of C against T or U.



**Figure 1.** (A) Schematic illustration of the analysis of target DNA or microRNA sequences based on the binding of abasic (AP) site-binding ligands conjugated with thiazole orange conjugates (APL–TO) in combination with an AP site-containing probe for hybridization; (B) Chemical structures of chloro-substituted 2-amino-5,7-dimethyl-1,8-naphthyridine (ClNaph) and its related derivatives. Proposed binding of these ligands with cytosine and thymine (two possible patterns) are also shown.

This work describes the enhanced binding selectivity of a 2-amino-1,8-naphthyridine derivative for orphan C by the incorporation of a chloro group. We previously found that the introduction of the electron-withdrawing trifluoromethyl group into 2-amino-7-methyl-1,8-naphthyridine (AMND) led to remarkably enhanced C/T selectivity due to more favorable protonation at N1 compared to N8 [24]. However, the obtained CF<sub>3</sub>-AMND (Figure 1B) showed much weaker affinity compared to ATMND. Herein, 2-amino-5,7-dimethyl-1,8-naphthyridine (ADMND), an AMND derivative with an additional methyl group, was explored as the scaffold for the introduction of an electron-withdrawing group considering the fact that ADMND can bind more strongly to target C than AMND [3]. The resulting compound carrying the chloro group at position 6 in the naphthyridine ring (ClNaph) was found to exhibit the strong binding affinity as well as the high selectivity for target C. The use of ClNaph as an APL unit in the conjugate with TO was shown to be useful for the design of a light-up probe toward target C over other nucleobases in the AP–DNA duplexes and AP–DNA/RNA hybrids with a view toward the detection of C-related mutations in DNA and microRNA sequences.

## 2. Materials and Methods

### 2.1. Materials

All of the DNAs and RNAs were purchased from Nihon Gene Research Laboratories, Inc. (Sendai, Japan) and Sigma-Genosys (Hokkaido, Japan), respectively. The other reagents were commercially available and used without further purification. The concentration of DNAs and RNAs were determined from the molar extinction coefficient at 260 nm ( $\epsilon_{260}$ ) according to the literature [25]. Water was deionized ( $\geq 18.0$  M $\Omega$  cm specific resistance) by an Elix 5 UV water purification system and a Milli-Q synthesis A10 system (Millipore Corp., Bedford, MA, USA). The other reagents were purchased from standard suppliers and used without further purification. <sup>1</sup>H NMR spectra were measured with a JEOL

ECA-600 spectrometer at 500 MHz. High-resolution ESI-MS spectra were measured with a Bruker APEX III mass spectrometer.

Unless otherwise mentioned, all measurements were performed in 10-mM sodium cacodylate buffer solutions (pH 7.0) containing 100-mM NaCl, 1.0-mM EDTA and ethanol (<2%). Before the measurements, target duplex-containing samples were annealed as follows: heated at 75 °C for 10 min and gradually cooled to 5 °C (3 °C/min), after which the solution temperature was raised again to 20 °C. The probe was then added to the samples.

## 2.2. Probe Synthesis

CINaph: 2,6-diaminopyridine (3.17 g, 29.1 mmol) in phosphoric acid (50 mL) was added to 3-chloropentane-2,4-dione (3.99 g, 29.6 mmol) and the reaction mixture was stirred at 90 °C for 24 h. After neutralization with NaOH, the mixture was filtered. The brown residue was extracted with 300 mL of CHCl<sub>3</sub> three times and the organic phase was dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated *in vacuo*, affording CINaph as a dark brown solid (3.57 g, 17.2 mmol, 59.2%). <sup>1</sup>H-NMR (CDCl<sub>3</sub>): 8.03 (d, 1H, *J* = 8.4 Hz), 6.76 (d, 1H, *J* = 9.2 Hz), 2.75 (s, 3H), 2.64 (s, 3H). High resolution ESI-MS calcd ([M + H]<sup>+</sup>) 207.0563; found, 207.0569.

CINaph-TO and CINaph-TO2 conjugates: The conjugates were synthesized according to our previous report [19]. Briefly, aminoethyl group was incorporated into CINaph and the resulting derivative was attached with the carboxylate-terminated decanyl (C10) spacer-containing TO derivative [26,27]. CINaph-TO conjugate, <sup>1</sup>H-NMR (DMSO-*d*<sub>6</sub>) 8.80 (d, 1H, *J* = 8.8 Hz), 8.60 (d, 1H, *J* = 7.2 Hz), 8.13 (d, 1H, *J* = 8.8 Hz), 8.05 (d, 1H, *J* = 7.6 Hz), 8.00 (t, 1H, *J* = 7.6 Hz), 7.80 (d, 1H, *J* = 8.4 Hz), 7.76 (t, 1H, *J* = 8.0 Hz), 7.63 (t, 1H, *J* = 8.4 Hz), 7.41 (t, 1H, *J* = 8.0 Hz), 7.36 (d, 1H, *J* = 7.6 Hz), 6.93 (s, 1H), 6.78 (d, 1H, *J* = 9.2 Hz), 4.58 (t, 2H, *J* = 7.2 Hz), 3.44–3.47 (q, 4H), 2.58 (s, 3H), 2.54 (s, 3H), 2.05 (t, 2H, *J* = 7.2 Hz), 1.84 (t, 2H), 1.42 (t, 2H), 1.34–1.16 (q, 14H). High resolution ESI-MS calcd ([M + H]<sup>+</sup>) 707.3299; found, 707.3292. CINaph-TO2 conjugate, <sup>1</sup>H-NMR (DMSO-*d*<sub>6</sub>) 8.72 (d, 1H, *J* = 8.8 Hz), 8.61 (d, 1H, *J* = 6.8 Hz), 8.07–8.01 (m, 4H), 7.78–7.74 (m, 2H), 7.62–7.58 (t, 1H, *J* = 7.6 Hz), 7.43–7.37 (m, 2H), 6.92 (s, 1H), 6.77 (d, 2H, *J* = 4.0 Hz), 4.61 (t, 2H), 4.17 (s, 3H), 3.50–3.42 (m, 4H), 2.58 (s, 3H), 2.55 (s, 3H), 2.00 (t, 2H, *J* = 7.2 Hz), 1.59–1.12 (m, 18H). High resolution ESI-MS calcd ([M + H]<sup>+</sup>) 707.3299; found, 707.3286.

## 2.3. Fluorescent Measurements

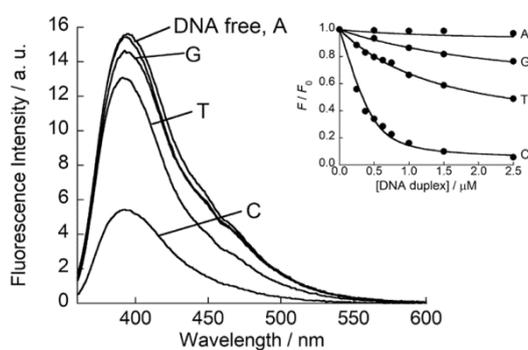
Fluorescence spectra were measured with a JASCO model FP-6500 spectrofluorophotometer equipped with a thermoelectrically temperature-controlled cell holder (Japan Spectroscopic Co., Ltd., Tokyo, Japan) using a 3 × 3 mm quartz cell.

The dissociation constant (*K*<sub>d</sub>) of the probe was determined at 20 °C by fluorescent titration experiments. The changes in fluorescence intensity were analyzed by nonlinear least-squares regression based on a 1:1 binding isotherm [19]. Errors in the *K*<sub>d</sub> values are the standard deviations obtained from three independent experiments (*N* = 3).

## 3. Results and Discussion

First, we examined the binding ability of CINaph (500 nM) to target orphan nucleobases in model 21-meric AP-DNA duplexes (500 nM; 5'-d(GCA GCT CCC AXA GTC TCC TCG)-3'/3'-d(CGT CGA GGG TNT CAG AGG AGC)-5', X = AP site (Spacer C3, a propyl residue), N = target nucleobase; G, C, A or T) by the fluorescence measurements. As shown in Figure 2, CINaph showed the emission with the maximum at 393 nm in the absence of DNAs. The addition of AP-DNA duplexes caused a decrease in the fluorescence intensity of CINaph. The largest fluorescence quenching response was observed for target C, which indicates that CINaph shows the preferential binding to target C compared to target G, A and T. The binding affinity of CINaph was assessed by the fluorescence titration experiments (Inset of Figure 2). CINaph showed the quenching response for target DNA duplexes in a concentration-dependent manner and the resulting titration curves were analyzed by a

1:1 binding model for the determination of the dissociation constants ( $K_d$ ). The  $K_d$  value for C was obtained as  $70 \pm 5.3$  nM. Significantly, this affinity is much stronger compared to the parent ADMND and the previously developed CF<sub>3</sub>-AMND and is almost comparable to ATMND (Table 1). In addition, CINaph showed useful binding selectivity toward target C, in which the  $K_d$  value for target C was two orders of magnitude smaller than those for other target nucleobases ( $K_d$ /nM: T,  $2020 \pm 240$ , G and A > 5000). The observed selectivity for C over T is much superior to ADMND and ATMND [3] whereas it is slightly moderate in comparison with CF<sub>3</sub>-AMND [24]. We reason that this is due to the more favorable protonation of CINaph at the N1 position for the complementary base-pairing with orphan C compared to the N8 protonation (Figure 1B), where the electron-withdrawing effect of the chloro group would be essential as was observed for CF<sub>3</sub>-AMND [24]. We note that the type of the AP site affects the binding of CINaph. The use of Spacer C3 (propyl residue) allowed the stronger binding to target C compared to the tetrahydrofuranyl residue (dSpacer), presumably due to less steric hindrance (Figure S3). The observed strong and highly selective binding properties for target C should render CINaph a useful APL unit in the conjugate with TO unit.



**Figure 2.** Fluorescence response of CINaph (500 nM) to target AP–DNA duplexes (500 nM). Inset: Titration curves for the binding of CINaph (500 nM) to target nucleobases. Measurements were done in solutions buffered to pH 7.0 (10-mM sodium cacodylate) containing 100-mM NaCl and 1.0-mM EDTA.  $F$  and  $F_0$  denote the fluorescence intensities of CINaph in the presence and absence of DNA duplexes. Excitation, 350 nm. Analysis, 393 nm. Temperature, 20 °C.

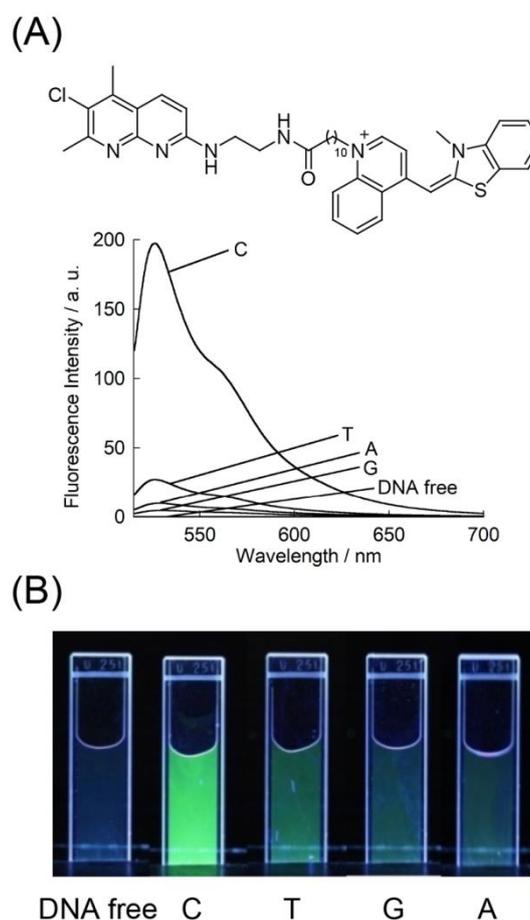
**Table 1.** Dissociation constants ( $K_d$ ) of CINaph and its related derivatives for target C in 21-meric AP–DNA duplexes <sup>a</sup>.

	$K_d$ /nM
CINaph	$70 \pm 5.3$
ADMND <sup>b</sup>	164
ATMND <sup>b</sup>	53
CF <sub>3</sub> -AMND <sup>c</sup>	1400

<sup>a</sup>  $K_d$  values measured in solutions buffered to pH 7.0 (10-mM sodium cacodylate) containing 100-mM NaCl and 1.0-mM EDTA. Temperature, 20 °C. <sup>b</sup> Values taken from [3]. <sup>c</sup> Value taken from [24].

CINaph was coupled with the quinolone ring of the TO unit through a long alkyl (C10)-linker according to our previous report [19], which affords the CINaph–TO conjugate (Figure 3A). We measured the fluorescence response of the TO unit of the conjugate (500 nM) for the same AP–DNA duplexes (500 nM) as those used for the examination of CINaph (cf. Figure 2). As shown in Figure 3A, TO unit shows negligible emission in the absence of DNAs due to the free rotation of the benzothiazole and quinolone rings [21]. In contrast, we observed a remarkable light-up response of the TO unit for target C-containing AP–DNA duplex, in which the fluorescence intensity at 526 nm increased by 207-fold. This can be explained by the restriction of the rotation of the TO unit by intercalation into the duplex region (cf. Figure 1A) [19]. In addition, the fluorescence titration experiments revealed that the  $K_d$  value of CINaph–TO conjugate for target C reached  $6.6 \pm 0.3$ -nM (Figure S4). This affinity is

one order of magnitude stronger than CINaph itself (cf. Table 1). Apparently, the conjugation with the TO unit led to the enhanced binding affinity for target C in the AP–DNA duplexes. It should be noted that CINaph–TO conjugate has high binding selectivity for target C over other three nucleobases (Figure S4:  $K_d$  values for T, G and A > 1500 nM). Considering that TO lacks the selectivity for the orphan nucleobases [19], the CINaph unit is responsible for the observed C-selectivity of the conjugate. This was confirmed by the fluorescence quenching response of the CINaph unit with high selectivity to target C (Figure S5). These results showed the useful binding and light-up functions of the CINaph–TO conjugate for the detection of C-related mutations in target DNA sequences. It is also noteworthy that the strong emission of CINaph–TO conjugate for target C can be seen even with the naked eyes under UV light irradiation (Figure 3B). Fluorescence response for target C is clearly distinguishable from those for other target nucleobases, which thus facilitates a simple and rapid analysis of target DNA sequences.

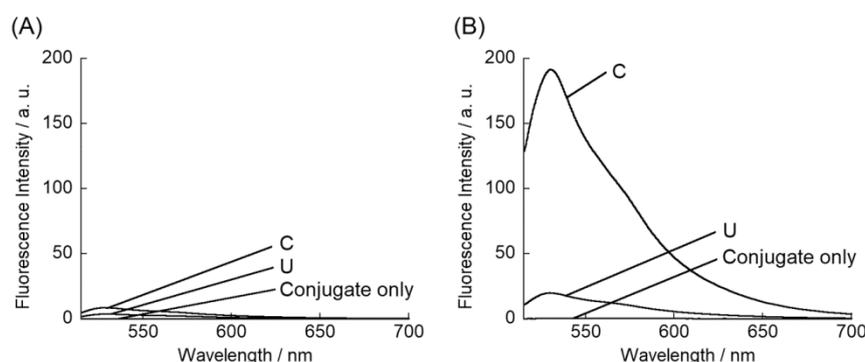


**Figure 3.** (A) Fluorescence response of CINaph–TO conjugate (500 nM) to target AP–DNA duplexes (500 nM). Excitation, 504 nm. Temperature, 20 °C. (B) Images of fluorescence emission of CINaph–TO conjugate (500 nM) in the absence and presence of target AP–DNA duplexes (500 nM) under excitation light with the wavelength of 365 nm, obtained by a digital camera at room temperature. Other solution conditions are the same as those given in Figure 2.

We found that the kind of heterocycles of the TO unit, benzothiazole or quinolone rings, that connect to CINaph unit in the conjugate affected the binding and fluorescence signaling abilities for the binding to AP–DNA duplexes. When the C10-linker was appended to the benzothiazole ring, the resulting conjugate (CINaph–TO2) showed the selective light-up response for target C in the AP–DNA duplexes (Figure S6); however, the degree of the response was smaller compared to CINaph–TO (cf. Figure 3A). This is attributable to the reduced binding affinity of CINaph–TO2 conjugate to target C

( $K_d = 71$  nM). This indicates that the conjugation of CINaph unit into the benzothiazole ring hinders the effective intercalation of the TO unit for DNA duplexes as reported in the literature [27]. In addition, we observed moderate selectivity of CINaph–TO2 conjugate for target C over other three nucleobases in AP–DNA duplexes (Figure S6) while the reason for this is unclear yet. Hence, the connection of CINaph unit to the quinolone ring of the TO unit is effective for the strong binding and large light-up response of the conjugate for target C in the AP–DNA duplexes.

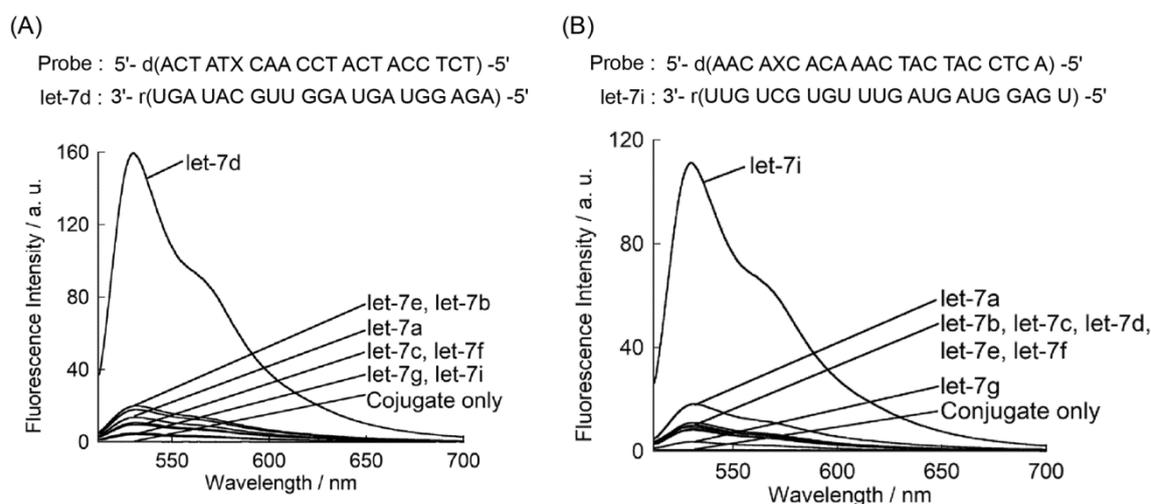
As described above, the CINaph–TO conjugate can serve as a useful light-up probe for target C in the AP–DNA duplexes. Meanwhile, we noticed weak binding of the conjugate for AP–RNA duplexes (Figure 4A and Figure S7). The binding affinity was estimated as  $> 1300$ -nM for target C in the AP–RNA duplex (5′-r(GCA GCU CCC AXA GUC UCC UCG)-3′/3′-r(CGU CGA GGG UCU CAG AGG AGC)-5′, X = AP site (Spacer C3), C = target nucleobase), which was three orders of magnitude larger than that for the AP–DNA duplex (cf. Figure S4). We consider that this arises from the preferential binding of the CINaph–TO conjugate to B-formed AP–DNA duplexes relative to A-formed AP–RNA duplexes [6]. However, when targeting C in single stranded RNA, the use of AP–DNA probe is highly effective. The resulting duplex is the hybrid between RNA and AP–DNA ((5′-d(GCA GCT CCC AXA GTC TCC TCG)-3′/3′-r(CGU CGA GGG UCU CAG AGG AGC)-5′, X = AP site (Spacer C3), C = target nucleobase) that can adopt A-form/B-form intermediate structure [28], under the same measurement conditions used to examine AP–RNA duplexes. As shown in Figure 4B, we observed the significant light-up response of the conjugate for the AP–DNA/RNA hybrid, in which the response was 22-fold larger compared to that for the AP–RNA duplex. The  $K_d$  for target C in the hybrid was obtained as 22 nM (Figure S7), where the affinity is much superior to that for AP–RNA duplexes. Importantly, the CINaph–TO conjugate retains its selectivity for target C over U opposite the AP site in the DNA/RNA hybrid, as was observed in AP–DNA duplexes (cf. Figure 3B). These results suggest the potential use of CINaph–TO conjugate for the detection of microRNAs in combination with AP–DNA hybridization probe (cf. Figure 1A).



**Figure 4.** Fluorescence response of CINaph–TO conjugate (500 nM) for target C or U in AP site-containing duplexes (500 nM): (A), AP–RNA duplex and (B) AP–DNA/RNA hybrid. Other solution conditions are the same as those given in Figure 2. Excitation, 504 nm. Temperature, 20 °C.

We performed the preliminary experiments for detection of microRNAs based on the binding-induced light-up response of CINaph–TO conjugate for AP–DNA/RNA hybrids. A 21-meric AP–DNA probe was designed for the selective detection of the let-7d sequence among let-7 family [29], as shown in Figure 5A. Hybridization between this probe and let-7d allows the construction of the AP–DNA/RNA hybrid containing an orphan C opposite an AP site. Meanwhile, the hybridization with other let-7 sequences leads to the formation of orphan U-containing hybrids with several mismatch base pairs (Table S1). We observed the significant light-up response of CINaph–TO for the let-7d-containing hybrid (Figure 5A). This response is much larger than those for other let-7 sequences, which clearly shows that CINaph–TO conjugate enables the selective detection probe for let-7d over other let-7 members. The binding affinity of the conjugate for the let-7d-containing hybrid ( $K_d = 23$  nM) was

found to be comparable to that for the model AP–DNA/RNA hybrid (cf. Figure S8). Our assay can be applied to the analysis of various microRNAs by using the AP site-containing DNA probe whose sequence is designed so as to be complementary to the target sequence. Figure 5B shows the fluorescence response of the conjugate in combination with the 22-meric AP–DNA probe for let-7i. Selective light-up detection for let-7i was achieved due to the selective binding of the conjugate to target C in the AP–DNA/RNA hybrid formed between the AP–DNA probe and let-7i. These results show the applicability of the CINaph–TO conjugate for the selective detection of target microRNA sequences based on the hybridization for the construction of the AP–DNA/RNA hybrid as well as the binding-induced light-up response of the conjugate for the orphan C in the resulting hybrid.



**Figure 5.** Selective detection of (A) let-7d and (B) let-7i sequences by CINaph–TO conjugate in combination with AP–DNA probe (X = AP site (Spacer C3)). Sequences of other let-7 members are shown in Table S1. [CINaph–TO], [target microRNA], [AP–DNA probe] = 500 nM. Other solution conditions for the fluorescence response of the conjugate are the same as those given in Figure 2. Excitation, 504 nm. Temperature: 20 °C.

#### 4. Conclusions

In summary, we report that CINaph served as a strong and highly selective binder for the orphan C opposite an AP site in DNA duplexes. In addition, we demonstrated the usefulness of CINaph as the APL unit in the conjugate with a TO unit for the design of light-up probes for the detection of C-related mutations in DNA and microRNA sequences. These results obtained could provide insights needed to design this class of light-up conjugates suitable for the analysis of DNA and microRNA sequences. As shown in our previous works [19,20], the spacer length and structure have a large impact on the binding and fluorescence sensing abilities of APL–TO conjugates. We will examine these concerns for further improvement of the APL–TO conjugates in order to develop the practical assays.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/12/4133/s1>, Figure S1: <sup>1</sup>H NMR spectra of the conjugates; Figure S2: ESI-MS spectra of the conjugates; Figure S3: Fluorescence response of CINaph to target dSpacer-containing AP–DNA duplexes. Inset: Titration curves for the binding of CINaph to target nucleobases; Figure S4: Titration curves for the binding of CINaph–TO conjugate to target nucleobases; Figure S5: Fluorescence response of CINaph unit in the conjugate to target AP–DNA duplexes; Figure S6: Chemical structure of CINaph–TO2 conjugate and its fluorescence response to target AP–DNA duplexes; Figure S7: Titration curves of the binding of CINaph–TO conjugate to target C in the AP–DNA/RNA hybrid and AP–RNA duplex; Table S1: Sequences of let-7 members used in this study. Figure S8. Titration curves for the binding of CINaph–TO conjugate to the let-7d/DNA probe hybrid.

**Author Contributions:** C.-x.W. and Y.S. conceived and designed the experiments. C.-x.W. and T.S. performed the experiments. C.-x.W. and Y.S. wrote the study. N.T. and S.N. supervised the project. All authors have read and agreed to the published version of the manuscript.

**Funding:** Y.S., N.T. and S.N. acknowledge the support from Scientific Research (S) (No. 22225003) from Japan Society for the Promotion of Science (JSPS).

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

## References

1. Granzhan, A.; Kotera, N.; Teulade-Fichou, M.-P. Finding needles in a haystack: Recognition of mismatched base pairs in DNA by small molecules. *Chem. Soc. Rev.* **2014**, *43*, 3630–3665. [[CrossRef](#)] [[PubMed](#)]
2. Song, G.; Ren, J. Recognition and regulation of unique nucleic acid structures by small molecules. *Chem. Commun.* **2010**, *46*, 7283–7294. [[CrossRef](#)]
3. Sato, Y.; Nishizawa, S.; Yoshimoto, K.; Seino, T.; Ichihashi, T.; Morita, K.; Teramae, N. Influence of substituent modifications on the binding of 2-amino-1,8-naphthyridines to cytosine opposite an AP site in DNA duplexes: Thermodynamic characterization. *Nucleic Acids Res.* **2009**, *37*, 1411–1422. [[CrossRef](#)] [[PubMed](#)]
4. Qing, D.; Xu, C.-Y.; Sato, Y.; Yoshimoto, K.; Nishizawa, S.; Teramae, N. Enhancement of the binding ability of a ligand for nucleobase recognition by introducing a methyl group. *Anal. Sci.* **2006**, *22*, 201–203.
5. Wang, C.-X.; Sato, Y.; Kudo, M.; Nishizawa, S.; Teramae, N. Ratiometric fluorescent signaling of small molecules, environmentally sensitive dye conjugates for detecting single base mutations in DNA. *Chem. Eur. J.* **2012**, *18*, 9481–9484. [[CrossRef](#)] [[PubMed](#)]
6. Sato, Y.; Ichihashi, T.; Nishizawa, S.; Teramae, N. Strong and selective binding of amiloride to an abasic site in RNA duplexes: Thermodynamic characterization and microRNA detection. *Angew. Chem. Int. Ed.* **2012**, *51*, 6339–6372. [[CrossRef](#)]
7. Sato, Y.; Toriyabe, Y.; Nishizawa, S.; Teramae, N. 2,4-Diamino-6,7-dimethylpteridine as a fluorescent ligand for binding and sensing an orphan cytosine in RNA duplexes. *Chem. Commun.* **2013**, *49*, 9983–9985. [[CrossRef](#)]
8. Fakhari, M.A.; Rokita, S.E. A new solvatochromic fluorophore for exploring nonpolar environments created by biopolymers. *Chem. Commun.* **2011**, *47*, 4222–4224. [[CrossRef](#)] [[PubMed](#)]
9. Suda, H.; Kobori, A.; Zhang, J.; Hayashi, G.; Nakatani, K. *N'*-Bis(3-aminopropyl)-2,7-diamino-1,8-naphthyridine stabilized a single pyrimidine bulge in duplex DNA. *Bioorg. Med. Chem.* **2005**, *13*, 4507–4512. [[CrossRef](#)] [[PubMed](#)]
10. Verma, R.K.; Takei, F.; Nakatani, K. Synthesis and photophysical properties of fluorescence molecular probe for turn-on-type detection of cytosine bulge DNA. *Org. Lett.* **2016**, *18*, 3170–3173. [[CrossRef](#)]
11. Arambula, J.F.; Ramisetty, S.R.; Baranger, A.M.; Zimmerman, S.C. A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16068–16073. [[CrossRef](#)] [[PubMed](#)]
12. Zeglis, B.M.; Barton, J.K. A mismatch-selective bifunctional rhodium-oregon green conjugate: A fluorescent probe for mismatched DNA. *J. Am. Chem. Soc.* **2006**, *128*, 5654–5655. [[CrossRef](#)] [[PubMed](#)]
13. Sato, T.; Sato, Y.; Iwai, K.; Kuge, S.; Nishizawa, S.; Teramae, N. Synthetic fluorescent probes capable of selective recognition of 3'-overhanging nucleotides for siRNA delivery imaging. *Chem. Commun.* **2015**, *51*, 1421–1424. [[CrossRef](#)]
14. Tanabe, T.; Sato, T.; Sato, Y.; Nishizawa, S. Design of a fluorogenic PNA probe capable of simultaneous recognition of 3'-overhang and double-stranded sequences of small interfering RNAs. *RSC Adv.* **2018**, *8*, 42095–42099. [[CrossRef](#)]
15. Sato, Y.; Kaneko, M.; Sato, T.; Nakata, S.; Takahashi, Y.; Nishizawa, S. Enhanced binding affinity of siRNA overhang-binding fluorescent probes by conjugation with cationic oligopeptides for improved analysis of the siRNA delivery process. *ChemBioChem* **2019**, *20*, 408–414. [[CrossRef](#)] [[PubMed](#)]
16. Xu, Z.; Morita, K.; Sato, Y.; Dai, Q.; Nishizawa, S.; Teramae, N. Label-free aptamer-based sensor using abasic site-containing DNA and a nucleobase-specific fluorescent ligand. *Chem. Commun.* **2009**, *42*, 6445–6447. [[CrossRef](#)] [[PubMed](#)]
17. Sato, Y.; Nishizawa, S.; Teramae, N. Label-free molecular beacon system based on DNAs containing abasic sites and fluorescent ligands that bind abasic sites. *Chem. Eur. J.* **2011**, *17*, 11650–11656. [[CrossRef](#)]
18. Sato, Y.; Zhang, Y.; Nishizawa, S.; Seino, T.; Nakamura, K.; Li, M.; Teramae, N. Competitive assay for theophylline based on an abasic site-containing DNA duplex aptamer and a fluorescent ligand. *Chem. Eur. J.* **2012**, *18*, 12719–12724. [[CrossRef](#)]

19. Sato, Y.; Kudo, M.; Toriyabe, Y.; Kuchitsu, S.; Wang, C.-X.; Nishizawa, S.; Teramae, N. Abasic site-binding ligands conjugated with cyanine dyes for “off-on” fluorescence sensing of orphan nucleobases in DNA duplexes and DNA-RNA hybrids. *Chem. Commun.* **2014**, *50*, 515–517. [[CrossRef](#)] [[PubMed](#)]
20. Sato, Y.; Saito, H.; Aoki, D.; Teramae, N.; Nishizawa, S. Lysine linkage in abasic site-binding ligand-thiazole orange conjugates for improved binding affinity to orphan nucleobases in DNA/RNA hybrids. *Chem. Commun.* **2016**, *52*, 14446–14449. [[CrossRef](#)]
21. Nygren, J.; Svanvik, N.; Kubista, M. The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* **1998**, *46*, 39–51. [[CrossRef](#)]
22. Armitage, B.A. Cyanine dye-DNA interactions: Intercalation, groove binding, and aggregation. *Top. Curr. Chem.* **2005**, *253*, 55–76.
23. Dong, H.; Lei, J.; Ding, L.; Wen, Y.; Ju, H.; Zhang, X. MicroRNA: Function, detection, and bioanalysis. *Chem. Rev.* **2013**, *113*, 6207–6233. [[CrossRef](#)]
24. Sato, Y.; Zhang, Y.; Seino, T.; Sugimoto, T.; Nishizawa, S.; Teramae, N. Highly selective binding of naphthyridine with a trifluoromethyl group to cytosine opposite an abasic site in DNA duplexes. *Org. Biomol. Chem.* **2012**, *10*, 4003–4006. [[CrossRef](#)] [[PubMed](#)]
25. Puglisi, J.D.; Tinoco, I. Absorbance melting curves of RNA. *Methods Enzymol.* **1989**, *180*, 304–325.
26. Carreon, J.R.; Stewart, K.M.; Mahon, K.P., Jr.; Shin, S.; Kelly, S.O. Cyanine dye conjugates as probes for live cell imaging. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5182–5185. [[CrossRef](#)]
27. Carreon, J.R.; Mahon, K.P., Jr.; Kelly, S.O. Thiazole orange-peptide conjugates: Sensitivity of DNA binding to chemical structure. *Org. Lett.* **2004**, *6*, 517–519. [[CrossRef](#)]
28. Shaw, N.N.; Arya, D.P. Recognition of the unique structure of DNA:RNA hybrids. *Biochimie* **2008**, *90*, 1026–1039. [[CrossRef](#)]
29. Johnson, S.M.; Grosshans, H.; Shingara, J.; Byrom, M.; Jarvis, R.; Cheng, A.; Labourier, E.; Reinert, K.L.; Brown, D.; Slack, F.J. RAS is regulated by the let-7 microRNA family. *Cell* **2005**, *120*, 635–647. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



MDPI  
St. Alban-Anlage 66  
4052 Basel  
Switzerland  
Tel. +41 61 683 77 34  
Fax +41 61 302 89 18  
[www.mdpi.com](http://www.mdpi.com)

*Applied Sciences* Editorial Office  
E-mail: [applsci@mdpi.com](mailto:applsci@mdpi.com)  
[www.mdpi.com/journal/applsci](http://www.mdpi.com/journal/applsci)





MDPI  
St. Alban-Anlage 66  
4052 Basel  
Switzerland

Tel: +41 61 683 77 34  
Fax: +41 61 302 89 18

[www.mdpi.com](http://www.mdpi.com)



ISBN 978-3-0365-1513-7