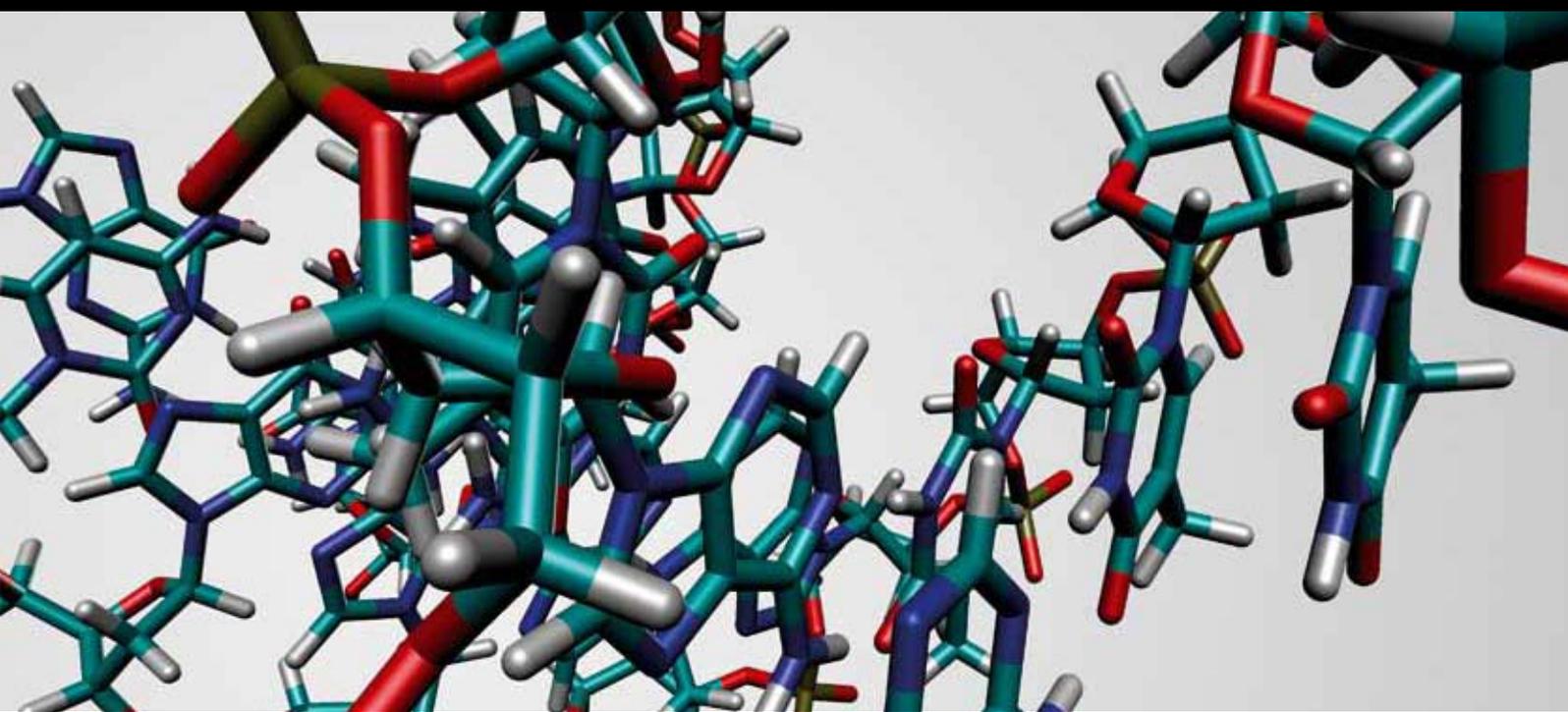


Synthesis and Application of Functional Nucleic Acids

Guest Editors: Daisuke Miyoshi, Friedrich C. Simmel, Souvik Maiti, and Luis A. Marky





Synthesis and Application of Functional Nucleic Acids

Journal of Nucleic Acids

Synthesis and Application of Functional Nucleic Acids

Guest Editors: Daisuke Miyoshi, Friedrich C. Simmel, Souvik Maiti, and Luis A. Marky



Copyright © 2011 SAGE-Hindawi Access to Research. All rights reserved.

This is a special issue published in volume 2011 of "Journal of Nucleic Acids." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Luigi A. Agrofoglio, France
Hiroyuki Asanuma, Japan
Ashis K. Basu, USA
Ben Berkhout, The Netherlands
Ann L. Beyer, USA
Christiane Branlant, France
Emery H. Bresnick, USA
Tom Brown, UK
Janusz M. Bujnicki, Poland
Jyoti Chattopadhyaya, Sweden
G. L. Dianov, UAE
Aidan Doherty, UK
Joachim Engels, Germany

Ramón Eritja, Spain
Marina Gottikh, Russia
John A. Hartley, UK
Roland K. Hartmann, Germany
Tapas K Hazra, USA
Hyouta Himeno, Japan
Robert Hudson, Canada
Shigenori Iwai, Japan
B. H. Kim, Republic of Korea
Jørgen Kjems, Denmark
Xingguo Liang, Japan
Luis A. Marky, USA
Atsushi Maruyama, Japan

P. E. Nielsen, Germany
Satoshi Obika, Japan
Tatiana S. Oretskaya, Russia
Tobias Restle, Germany
Mitsuo Sekine, Japan
Jacek Stawinski, Sweden
Dmitry A. Stetsenko, UK
Hiroshi Sugiyama, Japan
Zucaï Suo, USA
Akio Takenaka, Japan
Yitzhak Tor, USA
Dmitry Veprintsev, UK
Birte Vester, Denmark

Contents

Synthesis and Application of Functional Nucleic Acids, Daisuke Miyoshi, Friedrich C. Simmel, Souvik Maiti, and Luis A. Marky
Volume 2011, Article ID 547245, 2 pages

Evaluation of Fluorescent Analogs of Deoxycytidine for Monitoring DNA Transitions from Duplex to Functional Structures, Yogini P. Bhavsar, Samantha M. Reilly, and Randy M. Wadkins
Volume 2011, Article ID 986820, 7 pages

On-Chip DNA Methylation Analysis Using Osmium Complexation, Kaori Sugizaki, Tadashi Umemoto, and Akimitsu Okamoto
Volume 2011, Article ID 480570, 5 pages

Influence of a Hairpin Loop on the Thermodynamic Stability of a DNA Oligomer, Jurij Lah, Mojca Seručnik, and Gorazd Vesnaver
Volume 2011, Article ID 513910, 9 pages

A Structurally Variable Hinged Tetrahedron Framework from DNA Origami, David M. Smith, Verena Schüller, Carsten Forthmann, Robert Schreiber, Philip Tinnefeld, and Tim Liedl
Volume 2011, Article ID 360954, 9 pages

Photoswitching of Site-Selective RNA Scission by Sequential Incorporation of Azobenzene and Acridine Residues in a DNA Oligomer, Akinori Kuzuya, Keita Tanaka, and Makoto Komiyama
Volume 2011, Article ID 162452, 8 pages

Use of Nucleic Acid Analogs for the Study of Nucleic Acid Interactions, Shu-ichi Nakano, Masayuki Fujii, and Naoki Sugimoto
Volume 2011, Article ID 967098, 11 pages

Radiolytic Reduction Characteristics of Artificial Oligodeoxynucleotides Possessing 2-Oxoalkyl Group or Disulfide Bonds, Kazuhito Tanabe, Takeo Ito, and Sei-ichi Nishimoto
Volume 2011, Article ID 816207, 7 pages

Coupling Strategies for the Synthesis of Peptide-Oligonucleotide Conjugates for Patterned Synthetic Biomineralization, Joshua D. Carter and Thomas H. LaBean
Volume 2011, Article ID 926595, 8 pages

The Effects of Molecular Crowding on the Structure and Stability of G-Quadruplexes with an Abasic Site, Takeshi Fujimoto, Shu-ichi Nakano, Daisuke Miyoshi, and Naoki Sugimoto
Volume 2011, Article ID 857149, 9 pages

Direct Detection of Thrombin Binding to 8-Bromodeoxyguanosine-Modified Aptamer: Effects of Modification on Affinity and Kinetics, Shou Goji and Jun Matsui
Volume 2011, Article ID 316079, 5 pages

Use of a Designed Peptide Library to Screen for Binders to a Particular DNA G-Quadruplex Sequence, Keita Kobayashi, Noriko Matsui, and Kenji Usui
Volume 2011, Article ID 572873, 8 pages

The Stability of a Model Substrate for Topoisomerase 1-Mediated DNA Religation Depends on the Presence of Mismatched Base Pairs, William H. Gmeiner, Freddie Salisbury Jr., Chris M. Olsen, and Luis A. Marky

Volume 2011, Article ID 631372, 8 pages

Triplet Analysis That Identifies Unpaired Regions of Functional RNAs, Junji Kawakami, Yoshie Yamaguchi, and Naoki Sugimoto

Volume 2011, Article ID 471843, 6 pages

Synthesis of Specifically Modified Oligonucleotides for Application in Structural and Functional Analysis of RNA, Nico Rublack, Hien Nguyen, Bettina Appel, Danilo Springstube, Denise Strohbach, and Sabine Müller

Volume 2011, Article ID 805253, 19 pages

Commercially Supplied Amine-Modified siRNAs May Require Ultrafiltration prior to Conjugation with Amine-Reactive Compounds, Shannen Lau, Bim Graham, Ben J. Boyd, Colin W. Pouton, and Paul J. White

Volume 2011, Article ID 154609, 5 pages

tRNA Modification and Genetic Code Variations in Animal Mitochondria, Kimitsuna Watanabe and Shin-ichi Yokobori

Volume 2011, Article ID 623095, 12 pages

Control of Aptamer Function Using Radiofrequency Magnetic Field, Kenichi Taira, Koichi Abe, Takayuki Ishibasi, Katsuaki Sato, and Kazunori Ikebukuro

Volume 2011, Article ID 103872, 6 pages

Cell-Specific Aptamers as Emerging Therapeutics, Cindy Meyer, Ulrich Hahn, and Andrea Rentmeister

Volume 2011, Article ID 904750, 18 pages

A ssDNA Aptamer That Blocks the Function of the Anti-FLAG M2 Antibody, Amanda S. Lakamp and Michel M. Ouellette

Volume 2011, Article ID 720798, 11 pages

Non-B DNA Secondary Structures and Their Resolution by RecQ Helicases, Sudha Sharma

Volume 2011, Article ID 724215, 15 pages

Editorial

Synthesis and Application of Functional Nucleic Acids

Daisuke Miyoshi,¹ Friedrich C. Simmel,² Souvik Maiti,³ and Luis A. Marky⁴

¹ Faculty of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20, Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

² Physics Department, Technische Universität München, Am Coulombwall 4a, 85748 Garching, Germany

³ Proteomics and Structural Biology Unit, Institute of Genomics and Integrative Biology, CSIR, Mall Road, New Delhi 110 007, India

⁴ Department of Pharmaceutical Sciences, College of Pharmacy, University of Nebraska Medical Center, Omaha, NE 68198-6025, USA

Correspondence should be addressed to Daisuke Miyoshi, miyoshi@center.konan-u.ac.jp

Received 21 September 2011; Accepted 21 September 2011

Copyright © 2011 Daisuke Miyoshi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Both DNA and RNA can serve purposes beyond the storage and transfer of genetic information. They can display catalytic activity in a variety of chemical reactions, bind to diverse molecules, and respond to chemical stimuli. These functional nucleic acids have been thus employed in a broad spectrum of applications in biotechnology. In fact, the entire field of biotechnology has seen a complete revolution with the introduction and application of functional nucleic acids. Out of the many different types of functional nucleic acids, some of the prime contributions have been from riboswitches which have led to new ways to understand and manipulate biological processes based on molecular recognition coupled to structural rearrangements. Besides that, techniques like RNA interference which rely essentially on using small natural or synthetic functional ribonucleic acids have not only led us to a better understanding and modulation of gene regulation but also improved the rate and quality at which research progresses. Recently, functional nucleic acids have been ascribed to greater promise in nanotechnology as they have opened new avenues ranging from biosensors to computation. This special issue places the spotlight on the crossroads between nucleic acid chemistry and biology especially focusing on the design, synthesis, analysis as well as application of functional nucleic acids.

The first section of this special issue addresses synthesis, modification, characterization, and function of DNA molecules. It was demonstrated that chemical modifications of DNA molecules are promising for the design of biosensors to detect structural changes in nucleic acids and to bind proteins such as thrombin and antibodies, leading to the

controlled inhibition of protein functions. Noteworthy, DNA switches regulated by light, radiation, and magnetic fields are reported. These nucleic acid switches are useful to control functions of nucleic acids by external stimuli. Moreover, biophysical studies of natural and chemically modified nucleic acids should be useful for rational design of functional nucleic acids. The utilizations of DNA in nanotechnology as building blocks for nanostructures and as template for nanoparticles are further reported.

The second section of this special issue presents functional RNA molecules which are either naturally discovered, or developed by combinatorial chemistry named *in vitro* selection or SELEX. The *in vitro* and *in cellulo* applications of such functional RNA molecules are discussed. Moreover, naturally observed and chemical modifications of RNA molecules and their functionalities are presented.

Finally, the last section comprises articles that cover the interactions between nucleic acids and proteins including oligopeptides. The influence of noncanonical DNA secondary structures and base pairs on the interaction with topoisomerases and helicases is presented. These results provide a biophysical and biochemical rationale for the design of functional DNA molecules that can regulate such protein functions. Moreover, the use of a designed peptide library for the development of a particular DNA G-quadruplex sequence is reported.

Collectively, these papers describe the present state of chemistry and biology of functional nucleic acids. Current and ongoing advances in this field will certainly propel our capacity to develop even more sophisticated functional

nucleic acids in the future. In this regard, these articles all provide a perspective into the future of this field.

Acknowledgments

The editors of this issue want to express their deep appreciation to all of the authors and reviewers for their contributions to this issue.

Daisuke Miyoshi
Friedrich C. Simmel
Souvik Maiti
Luis A. Marky

Research Article

Evaluation of Fluorescent Analogs of Deoxycytidine for Monitoring DNA Transitions from Duplex to Functional Structures

Yogini P. Bhavsar, Samantha M. Reilly, and Randy M. Wadkins

Department of Chemistry and Biochemistry, The University of Mississippi, Coulter Hall, Room 409, MS 38677, USA

Correspondence should be addressed to Randy M. Wadkins, rwadkins@olemiss.edu

Received 15 May 2011; Accepted 19 June 2011

Academic Editor: F. C. Simmel

Copyright © 2011 Yogini P. Bhavsar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Topological variants of single-strand DNA (ssDNA) structures, referred to as “functional DNA,” have been detected in regulatory regions of many genes and are thought to affect gene expression. Two fluorescent analogs of deoxycytidine, Pyrrolo-dC (PdC) and 1,3-diaza-2-oxophenoxazine (tC^o), can be incorporated into DNA. Here, we describe spectroscopic studies of both analogs to determine fluorescent properties that report on structural transitions from double-strand DNA (dsDNA) to ssDNA, a common pathway in the transition to functional DNA structures. We obtained fluorescence-detected circular dichroism (FDCD) spectra, steady-state fluorescence spectra, and fluorescence lifetimes of the fluorophores in DNA. Our results show that PdC is advantageous in fluorescence lifetime studies because of a distinct ~2 ns change between paired and unpaired bases. However, tC^o is a better probe for FDCD experiments that report on the helical structure of DNA surrounding the fluorophore. Both fluorophores provide complementary data to measure DNA structural transitions.

1. Introduction

DNA single strands can hybridize to form higher-order functional structures, which include hairpins, triplexes, and quadruplexes [1–4]. The existence and physiological relevance of these secondary structures *in vivo* have been the subject of much controversy. However, several *in vivo* techniques have confirmed the presence of DNA secondary structures in telomeres and regulatory regions of specific genes (e.g., BCL-2, c-myc) [5–8]. Secondary structures may also serve as specific targets recognized by drugs, such as actinomycin D and PIPER, as well as transcription factors, such as Sp1, because of their topological variance from duplex Watson-Crick DNA [9, 10].

On account of their occurrence in regulatory regions and their structural peculiarity, functional DNA structures may serve as biological microswitches for altering transcription by silencing or enhancing gene expression [11, 12]. Exercising control over gene expression by controlling the activation of these switches and/or introducing new switches in biological

circuits could revolutionize medical research and offer new avenues of treating genetic disorders. For exercising such control, it is imperative to understand the factors affecting the mechanism of formation and maintenance of functional structures at a fundamental level.

Numerous fluorescent analogs of DNA bases have been evaluated for examining the subtleties of DNA transitions [13–15]. Locating an appropriate probe that could map the mechanistic aspects of transition of double stranded (ds) DNA to single stranded (ss) secondary structure is a first step toward monitoring the formation of complex, higher-order structures since ssDNA is an intermediate in the pathway to functional structures. In the following report, we present a comparative study of two deoxycytidine analogs, PdC and tC^o (Figure 1), to evaluate their suitability as fluorescent reporter probes for DNA transitions.

The properties of pyrrolocytosine and the effects of base stacking and hydrogen bonding on its quantum yield in nucleic acids have been previously evaluated [16]. Based on this prior work, we reported the use of PdC to determine

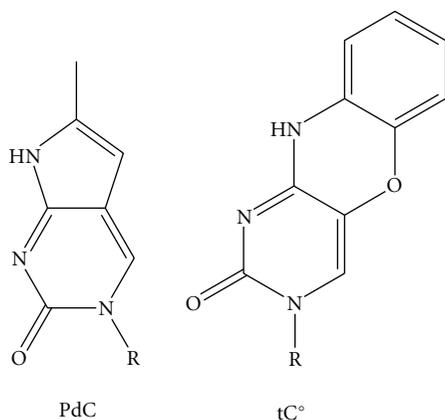


FIGURE 1: Structures of PdC and tC°. R indicates the linkage to deoxyribose.

hairpin formation in short oligos (16 nucleotides) [17]. The fluorophore was stable in these oligos and did not perturb DNA structure. This result implied that PdC might be a useful probe for investigating more complex DNA functional structures in detail. However, to overcome certain limitations of PdC, we also explored another recently characterized fluorescent base analog, tC°.

The tC° has a quantum yield five-times greater than PdC ($Q_f = 0.30$ for tC°) and a molar absorptivity maximum of $9000 \text{ M}^{-1} \text{ cm}^{-1}$ at 360 nm [18]. The absorbance wavelength of tC° is similar to PdC, which allows both to be easily distinguishable from DNA. Both of these fluorophores have been reported to be quenched, relative to the single strand, when base-paired with guanine, making it easy to determine when the DNA is in the duplex form [17, 19]. In addition, unlike PdC, tC° has only minor variations in fluorescent properties caused by surrounding bases [20]. However, like PdC, tC° induces little or no changes in stability upon incorporation into dsDNA. High quantum yield, retention of the original configuration of DNA, and quenching when base-paired suggested that tC° would be a useful fluorescent probe for mapping the transition of duplex DNA to a functional DNA structure spectroscopically. In this paper, we compare spectral properties of tC° and PdC for use as reporters of DNA conformation. The DNA sequence used (Table 1) is known to form a cruciform structure in the cloning vector pBR322 under superhelical duress [21, 22]. This sequence was chosen since the results from the work described here will provide a context for interpreting data in future studies of the PdC and tC° incorporated into this supercoiled plasmid. Hence the effect of supercoiling on functional structure formation and conformation can be potentially probed with these fluorescent bases.

2. Materials and Methods

2.1. Materials and Equipment. All DNA oligonucleotides (sequences shown in Table 1), including those incorporating PdC and tC°, were obtained from Midland Certified Reagent Co. (Midland, TX, USA). All oligos were dissolved in

TABLE 1: The 113-mer DNA sequence that is part of the cruciform found in the terminator of the Amp^R cassette in the cloning vector pBR322 under superhelical strain. The location of the substitution of cytidine by the fluorescent analogs marked by an asterisk.

Sequence
5'-TGA GGT TAA GGG ATT TTG GTC ATG AGA TTA TCA AAA
AGG ATC TTC* ACC TAG ATC CTT TTA ATT AAA AAT GAA
GTT TTA AAT CAA TCT AAA GTA TAT ATG AGT AAA CTT
GGG C -3'

TE buffer (10 mM Tris, 1 mM EDTA, pH 8) made from spectroscopic-grade reagents obtained from Fisher Scientific. Purity of each oligo was assessed using agarose gels, which showed no other higher or lower molecular weight contaminations. CD and FDCD experiments were performed using a circular dichroism instrument (Model 202SE, AVIV Biomedical Inc. Lakewood, NJ, USA) and fluorescence lifetime measurements were conducted on a multifrequency cross-correlation phase and modulation fluorometer (model K2, ISS Inc., Champaign, IL, USA).

2.2. Duplex Oligos. After ensuring the purity of the single-strand PdC, tC°, and wildtype (ssPdC, sstC°, and ssWT; the WT contains no fluorophore) oligos, each was mixed with a slight molar excess of their complementary strand. This was heated to 80°C for 20 minutes, and then cooled to room temperature to anneal strands. The double-strand oligo formation was confirmed by observing the difference in migration of ds- and ss-oligos on agarose gels and compared to a DNA ladder of the known length. In addition, duplex was the only strand observed after staining with the SYBR Gold, which stains both ss- and dsDNA.

2.3. CD and UV Absorbance. CD spectra were recorded using a quartz cuvette having an optical path length of 1 cm. All oligos analyzed were in 700 μL volume at concentration of 0.35 μM . Wild-type oligos were analyzed to ensure that the signals were not modified by the fluorophores. Each data point was averaged over an integration time of 1 s per nm. CD signals were collected from 25°C to 90°C over the range of 220 nm to 420 nm to check for any change in the global conformation of the single-stranded and double-stranded WT, PdC, and tC° oligos. While performing CD scans, the total UV absorbance of each oligo was also collected.

2.4. FDCD. FDCD spectra were collected using a quartz cuvette of 1 cm path length. Oligos analyzed were in a total volume of 70 μL and at the concentration of 8.5 μM . Each FDCD scan was averaged over 15 seconds per nm in order to increase signal-to-noise ratio. The excitation wavelengths were from 300 nm to 420 nm at temperatures ranging from 25°C to 90°C. Emission was obtained by using a 450 nm cut-on filter.

2.5. Steady-State Fluorescence. Steady-state fluorescence spectra were analyzed for ss- and ds- (WT, PdC, and tC°) oligos. All the oligos had concentration of 0.35 μM in TE

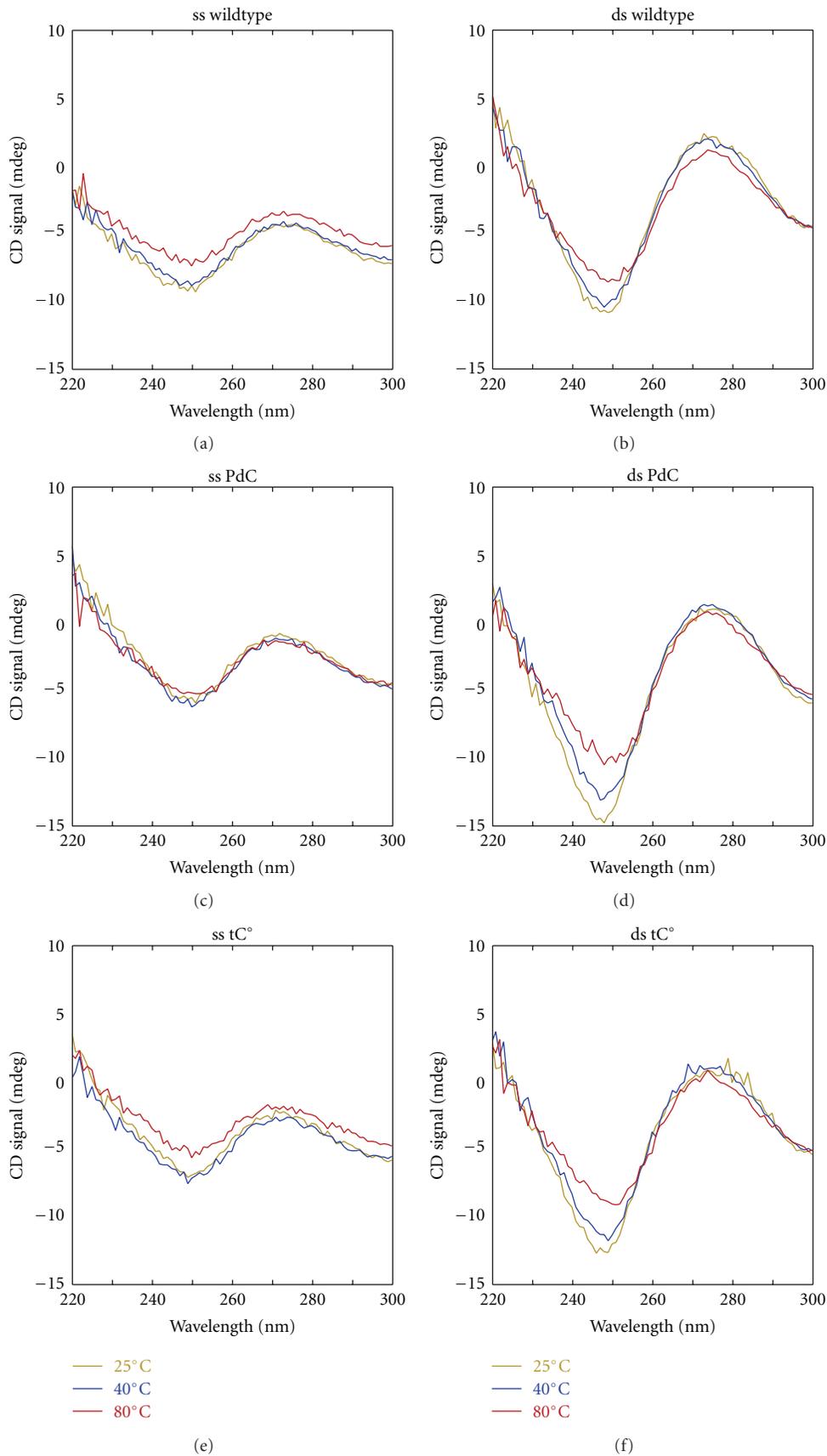


FIGURE 2: The characteristic CD spectrum of wild type ss- and ds-oligos as a function of temperature. All samples were at 0.35 μ M concentration.

TABLE 2: Fluorescence lifetimes of PdC and tC° in single- and double-strand DNA. The τ_1 and τ_2 are obtained by fitting a double exponential decay curve to the data, and f_1 and f_2 are fraction of fluorescence corresponding to lifetimes τ_1 and τ_2 , respectively. The χ^2 indicate the measure of error in the fit. The value $\langle\tau\rangle$ is average lifetime calculated by $\sum f_i\tau_i^2/\sum f_i\tau_i$.

Oligo	τ_1 (ns)	f_1	τ_2 (ns)	f_2	χ^2	$\langle\tau\rangle$ (ns)
ss tC°	5.3	0.9	1.7	0.1	2.0	5.1
ds tC°	5.1	0.8	1.8	0.2	1.5	4.8
ss PdC	5.2	0.7	0.7	0.3	1.9	5.1
ds PdC	3.3	0.6	0.1	0.3	1.4	3.2

buffer (10 mM Tris, 1 mM EDTA, pH 8). Excitation scans were collected over the wavelength range of 220–400 nm, with the emission wavelength fixed at 450 nm. Emission spectra were taken over the wavelength ranging from 400 nm to 580 nm with the excitation wavelength fixed at 350 nm. The ss- and dsWT oligos were used as a baseline signal and subtracted from the steady-state fluorescence data.

2.6. Fluorescence Lifetime Measurements. Fluorescence lifetime was measured for both ss- and ds-oligos using a frequency domain lifetime instrument. The cuvettes used were of 1 cm path length. Excitation wavelengths were 350 nm and 368 nm for PdC and tC°, respectively. Data were collected from 10–240 MHz frequencies with 100 iterations at each frequency and 40 total frequencies from which the lifetimes were calculated. Both the individual decays (fitted to 2 components) and average lifetimes are reported.

3. Results and Discussion

3.1. Incorporation of Fluorescent DC Analogs Does Not Perturb DNA Structure. In order to assure that no structural perturbations were introduced by insertion of the fluorophores in the oligos, changes in absorbance at 260 nm were monitored at varying temperatures to determine duplex melting temperatures as well as any transitions occurring in ssDNA (data not shown). The data indicated that the ssWT, ssPdC, and sstC° undergo similar spectral changes with temperature, indicating no perturbations to the ss-structures after replacement of cytidine with PdC or tC°. The absorbance for all ssDNA increases by ~10% from 25°C to 40°C, signifying reorientation of the strands during melting. Thus, introduction of either PdC or tC° did not lead to the formation of any additional secondary conformation in ss-oligos as predicted based on the previous literature [17, 19].

No perturbation of dsDNA by PdC or tC° was observed. All duplexes melted at $45 \pm 3^\circ\text{C}$ in TE buffer. This indicated dsDNA containing the fluorescent bases were readily formed and stable at 25°C. Hence, PdC and/or tC° could be used to investigate single- and double-stranded regions of DNA functional structures.

Circular dichroism (CD) spectroscopy was also used to investigate whether the overall helical structural conformation of the DNA might have been affected by the fluorophores (Figure 2). Our results, in agreement with melting

data from absorbance and the literature, indicated that the fluorophores did not affect the structure of the duplexes or the manner in which they melted [17, 23, 24]. There were only negligibly small changes in the CD spectrum with temperature for ss-oligos, which indicated little structure at any temperature as compared to dsDNA. The dsDNA oligos showed considerable CD, consistent with the common global spectral features typical of B-DNA [25]. Major changes in the CD signal of dsDNA with respect to the temperature were observed at 245 nm, with relatively smaller changes that occurred at 280 nm. These CD spectra and their temperature dependence illustrated that the DNA is not affected by replacing cytidine with PdC or tC°.

3.2. The Local Helical Structure of DNA Can Be Probed by Induced FDCD of tC°. Circular dichroism has been used for detecting the local environmental changes around PdC integrated into DNA [26]. However, the direct detection of CD by PdC or other base analogs is restricted to analogs that either have an absorbance far enough removed from the DNA λ_{max} of 260 nm to prevent spectral overlap or are incorporated into short DNA sequences. In both cases, significant amounts of DNA are required to obtain good signal-to-noise. An alternative approach to direct CD observation is to use FDCD to study the local environment of the fluorescent base analog. Since our previous data has sufficiently demonstrated that the replacement of deoxycytidine with either PdC or tC° does not affect the original DNA structure, we analyzed the fluorophores by FDCD to explore the local structural changes occurring around the base. In our studies, PdC-containing DNA at concentrations up to 8.5 μM did not show any FDCD signal over the temperature range of 25°C to 90°C, suggesting that PdC’s low quantum yield is insufficient for FDCD measurements at reasonable DNA concentrations with our instrument. However, tC° showed a strong FDCD signal for both ss- and dsDNA between 300 nm and 420 nm even at submicromolar concentrations (0.5 μM ; Figure 3). The FDCD signal for duplex DNA was significantly larger than that for the ss oligo. In addition it had a positive value throughout, indicating the base was stacked parallel to the bases in the vicinity. The lower intensity of the peak in ss-oligos is attributed to the residual structure at room temperature. Thus, tC° is a very effective probe for monitoring the localized changes occurring in DNA structure in response to variations in structural parameters.

FDCD data obtained can be used for predicting the molecular details of higher-order secondary structures. For example, in case of loop regions of i-motif and cruciform structures, tC° will exhibit lower FDCD signal since these regions are less structured. However, tC° substituting for dC in the regions where base-pairing occurs should show higher FDCD signals. These structural details could then be used to assess the stabilities of functional DNA structures after binding of transcription factors or drugs.

3.3. Fluorescence Properties of the Base Analogs Reveal Their Pairing State. Fluorescence intensities of PdC and tC° depend on their hydrogen bonding state, and decrease if the

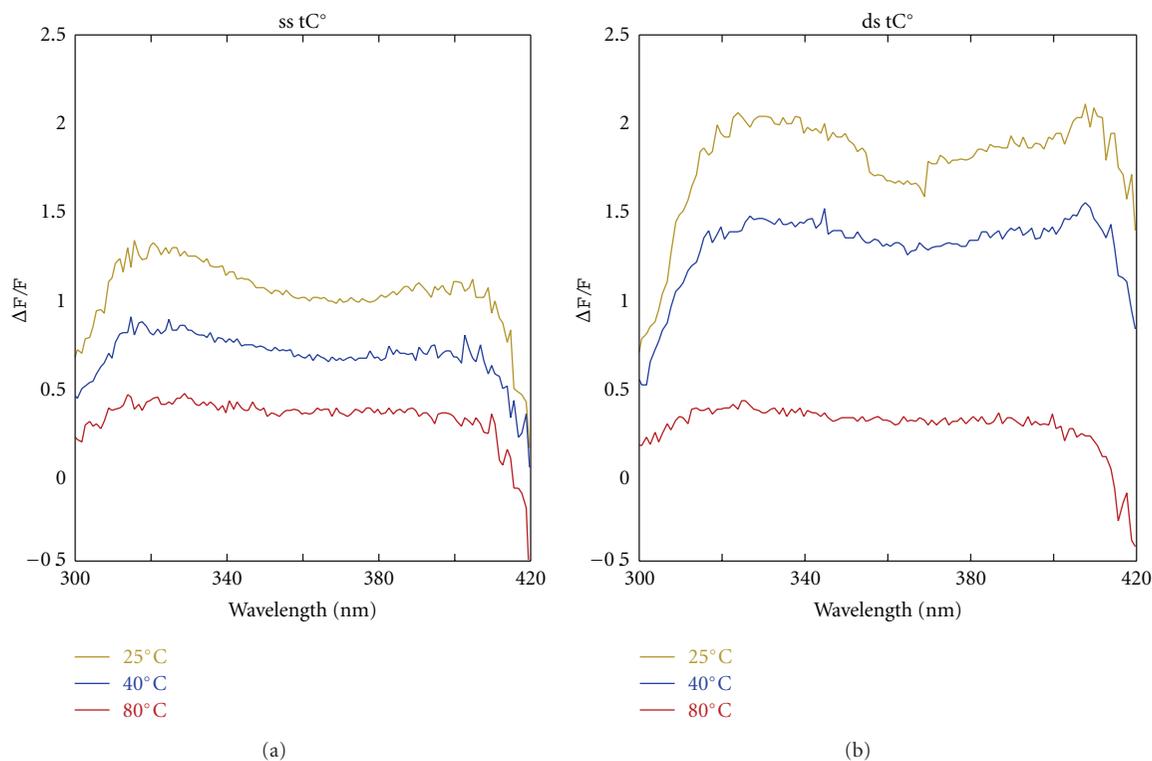


FIGURE 3: FDCD spectra with respect to temperature for PdC and tC°. The low quantum yield of PdC did not allow for detection of FDCD at this concentration. $\Delta F/F$ is the difference in the fluorescence caused by left- and right-handed polarized light over total fluorescence of the fluorophore.

fluorophores are base-paired [17, 19]. Figure 4 clearly demonstrate this quenching of fluorescence for both PdC and tC° in duplex dsDNA versus unpaired ssDNA. Fluorescence intensities from ssDNA to dsDNA are lowered by approximately 60% and 40% for PdC and tC°, respectively. Figure 4 also shows the difference in the quantum yield of PdC and tC°, with PdC exhibiting significantly lower fluorescence. Both the fluorophores had an excitation maximum at 350 nm and an emission maximum of 450 nm that did not change in ssDNA versus dsDNA. These decreases in the fluorescence and the location of the maxima are consistent with the known literature [8, 19].

Like steady-state intensities, fluorescence lifetimes of PdC and tC° decrease when they are paired to their complementary base [17, 19]. Since fluorescence lifetimes are independent of concentration, their decrease has the potential to be used for differentiating paired bases from unpaired, thus they can be used for deductively mapping functional DNA structures. Lifetime data was obtained at room temperature for both ss- and ds-oligos. Two lifetimes (τ_1 , τ_2) were fit to the phase-modulation data. The weighted average of the two lifetimes, τ_1 and τ_2 , were calculated and reported as $\langle \tau \rangle$ in Table 2. These weighted averages indicate that the lifetimes for tC° decreased slightly (0.3 ns) with respect to the strandedness. In contrast, PdC showed appreciable difference (1.9 ns) between single and duplex oligos. These lifetime values are similar to those previously reported [17, 19]. To unambiguously differentiate between the ss- and ds regions

of the secondary structures, greater differences in lifetimes of base-paired and unpaired fluorophores are required. Thus, the lifetime data indicates that PdC is a better choice than tC° for assessing the strand base-pairing in DNA via this methodology. Similar to FDCD, lifetime measurements can facilitate the studies to observe differences between ss loop regions and ds regions in functional DNA structures like i-motif and cruciforms, and hence aid the studies to investigate formation and stability of secondary structures.

4. Conclusions

In this report, spectroscopic techniques were used to determine the appropriate fluorescent deoxycytidine analog for probing the transition of duplex DNA into topologically distinct functional nucleic acid structures. Our conclusion is that neither analog tested is a perfect reporter. Rather, we suggest the two should be used in conjunction to obtain an overall description of changes to nucleic acid structure. While tC° has a very good quantum yield that allows for FDCD analysis, it exhibits only minor changes in fluorescence lifetime between ss- or dsDNA. PdC, on the other hand, undergoes a substantial decrease in the fluorescence lifetime when base-paired, but shows no FDCD signal at concentrations appropriate for most biochemical studies. Hence, we conclude that to map the transition of duplex DNA to other functional forms, both PdC and tC° at concentrations up to 8.5 μM can be used, depending on

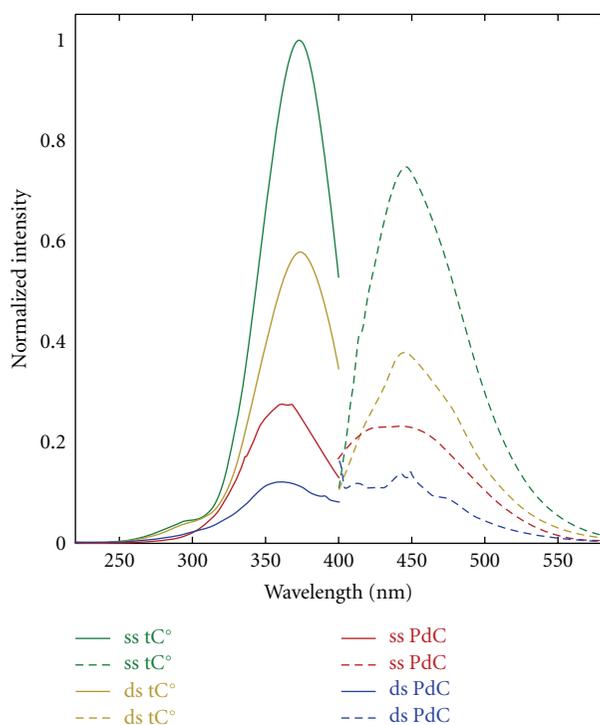


FIGURE 4: Excitation and emission spectra of PdC and tC° incorporated in ss- and dsDNA. The solid and dotted curves stand for excitation and emission spectra, respectively. All samples were at 0.35 μM concentration.

the structural information desired. We are now investigating spectral changes incurred by these probes, when they are contained within well-known functional DNA structures, such as quadruplex and i-motif DNA.

Acknowledgment

This paper has been supported by the US National Science Foundation EPSCoR Grant EPS-0903787.

References

- [1] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, and S. Neidle, "Quadruplex DNA: sequence, topology and structure," *Nucleic Acids Research*, vol. 34, no. 19, pp. 5402–5415, 2006.
- [2] M. D. Frank-Kamenetskii and S. M. Mirkin, "Triplex DNA structures," *Annual Review of Biochemistry*, vol. 64, pp. 65–95, 1995.
- [3] M. S. Han, A. K. R. Lytton-Jean, and C. A. Mirkin, "A gold nanoparticle based approach for screening triplex DNA binders," *Journal of the American Chemical Society*, vol. 128, no. 15, pp. 4954–4955, 2006.
- [4] R. M. Wadkins, "Targeting DNA secondary structures," *Current Medicinal Chemistry*, vol. 7, no. 1, pp. 1–15, 2000.
- [5] C. T. McMurray, "DNA secondary structure: a common and causative factor for expansion in human disease," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 5, pp. 1823–1825, 1999.
- [6] H. J. Lipps and D. Rhodes, "G-quadruplex structures: in vivo evidence and function," *Trends in Cell Biology*, vol. 19, no. 8, pp. 414–422, 2009.
- [7] H. Moore, P. W. Greenwell, C. P. Liu, N. Arnheim, and T. D. Petes, "Triplet repeats form secondary structures that escape DNA repair in yeast," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 4, pp. 1504–1509, 1999.
- [8] M. Nambiar, G. Goldsmith, B. T. Moorthy et al., "Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma," *Nucleic Acids Research*, vol. 39, pp. 936–948, 2011.
- [9] D. B. Davies, V. I. Pahomov, and A. N. Veselkov, "NMR determination of the conformational and drug binding properties of the DNA heptamer d(GpCpGpApApGpC) in aqueous solution," *Nucleic Acids Research*, vol. 25, no. 22, pp. 4523–4531, 1997.
- [10] W. Mäueler, G. Bassili, C. Epplen, H. G. Keyl, and J. T. Epplen, "Protein binding to simple repetitive sequences depends on DNA secondary structure(s)," *Chromosome Research*, vol. 7, no. 3, pp. 163–166, 1999.
- [11] T. A. Brooks and L. H. Hurley, "Targeting MYC expression through G-quadruplexes," *Genes Cancer*, vol. 1, pp. 641–649, 2010.
- [12] T. A. Brooks, S. Kendrick, and L. Hurley, "Making sense of G-quadruplex and i-motif functions in oncogene promoters," *FEBS Journal*, vol. 277, no. 17, pp. 3459–3469, 2010.
- [13] M. Kimoto, T. Mitsui, S. Yokoyama, and I. Hirao, "A unique fluorescent base analogue for the expansion of the genetic alphabet," *Journal of the American Chemical Society*, vol. 132, no. 14, pp. 4988–4989, 2010.
- [14] M. J. Rist and J. P. Marino, "Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions," *Current Organic Chemistry*, vol. 6, no. 9, pp. 775–793, 2002.
- [15] P. Sandin, L. M. Wilhelmsson, P. Lincoln, V. E. Powers, T. Brown, and B. Albinsson, "Fluorescent properties of DNA base analogue tC upon incorporation into DNA—negligible influence of neighbouring bases on fluorescence quantum yield," *Nucleic Acids Research*, vol. 33, no. 16, pp. 5019–5025, 2005.
- [16] S. J. O. Hardman and K. C. Thompson, "Influence of base stacking and hydrogen bonding on the fluorescence of 2-aminopurine and pyrrolocytosine in nucleic acids," *Biochemistry*, vol. 45, no. 30, pp. 9145–9155, 2006.
- [17] X. Zhang and R. M. Wadkins, "DNA hairpins containing the cytidine analog pyrrolo-dC: structural, thermodynamic, and spectroscopic studies," *Biophysical Journal*, vol. 96, no. 5, pp. 1884–1891, 2009.
- [18] S. Preus, K. Kilså, L. M. Wilhelmsson, and B. Albinsson, "Photophysical and structural properties of the fluorescent nucleobase analogues of the tricyclic cytosine (tC) family," *Physical Chemistry Chemical Physics*, vol. 12, no. 31, pp. 8881–8892, 2010.
- [19] P. Sandin, K. Börjesson, H. Li et al., "Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue," *Nucleic Acids Research*, vol. 36, no. 1, pp. 157–167, 2008.
- [20] L. M. Wilhelmsson, "Fluorescent nucleic acid base analogues," *Quarterly Reviews of Biophysics*, vol. 43, no. 2, pp. 159–183, 2010.

- [21] M. S. Z. Horwitz, "Transcription regulation in vitro by an E. coli promoter containing a DNA cruciform in the -35 region," *Nucleic Acids Research*, vol. 17, no. 14, pp. 5537–5545, 1989.
- [22] D. M. J. Lilley, D. Chen, and R. P. Bowater, "DNA supercoiling and transcription: topological coupling of promoters," *Quarterly Reviews of Biophysics*, vol. 29, no. 3, pp. 203–225, 1996.
- [23] N. P. Johnson, W. A. Baase, and P. H. von Hippel, "Low energy CD of RNA hairpin unveils a loop conformation required for λ N antitermination activity," *Journal of Biological Chemistry*, vol. 280, no. 37, pp. 32177–32183, 2005.
- [24] D. Jose, K. Datta, N. P. Johnson, and P. H. von Hippel, "Spectroscopic studies of position-specific DNA "breathing" fluctuations at replication forks and primer-template junctions," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 11, pp. 4231–4236, 2009.
- [25] J. Kypr, I. Kejnovská, D. Renčiuk, and M. Vorlíčková, "Circular dichroism and conformational polymorphism of DNA," *Nucleic Acids Research*, vol. 37, no. 6, pp. 1713–1725, 2009.
- [26] N. P. Johnson, W. A. Baase, and P. H. von Hippel, "Investigating local conformations of double-stranded DNA by low-energy circular dichroism of pyrrolo-cytosine," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 20, pp. 7169–7173, 2005.

Research Article

On-Chip DNA Methylation Analysis Using Osmium Complexation

Kaori Sugizaki,¹ Tadashi Umemoto,¹ and Akimitsu Okamoto^{1,2}

¹Nucleic Acid Chemistry Laboratory, RIKEN Advanced Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

²PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

Correspondence should be addressed to Akimitsu Okamoto, aki-okamoto@riken.jp

Received 28 February 2011; Accepted 1 April 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Kaori Sugizaki et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of a reaction for detecting the presence/absence of one methyl group in a long DNA strand is a chemically and biologically challenging research subject. A newly designed chemical assay on a chip for the typing of DNA methylation has been developed. A methylation-detection probe fixed at the bottom of microwells was crosslinked with methylated DNA mediated by osmium complexation and contributes to selective amplification of methylated DNA.

1. Introduction

Gene expression is well regulated by the epigenetic modification of DNA and histone tails independent of their primary sequences. In particular, cytosine methylation, in which the C5 position of the cytosine base is methylated enzymatically, plays a crucial role in the regulation of chromatin stability, gene regulation, parental imprinting, and X-chromosome inactivation in females [1–4]. Therefore, detection of cytosine methylation is very important, and much effort has gone into developing a simple reaction for 5-methylcytosine (^mC) detection.

For the evaluation of the methylation status of genes, several conventional methods have so far been used, such as a cleavage assay with methylation-insensitive restriction enzymes [5–7], hydrolysis and sequencing with a bisulfite salt [8–10], and immunofluorescence with anti-5-methylcytosine antibody [11, 12]. Although the conventional methods have many merits, there are many disadvantages, and methylation detection assays must be further improved through another approach. The existence of a more rapid and selective chemical reaction capable of distinguishing between methylcytosine and unmethylated cytosine on a chip has promise as a good method for efficiently analyzing the status of cytosine methylation at a specific site in a gene.

The sequence-selective DNA methylation-detection probe, ICON (interstrand complexes formed by osmium

and nucleic acids), may be effective for the development of an on-chip analysis of DNA methylation [13–15]. In the presence of osmium oxidants and a bipyridine ligand, 5-methylcytosine forms a stable osmium-centered complex, in contrast to unmethylated cytosine (Figure 1) [16–19]. ICON probes form a crosslink with a specific 5-methylcytosine in the probe-hybridizing DNA mediated by osmium-centered complex formation. This function will be effective for the capture of methylated DNA on a chip for sequence-selective methylation analysis.

In this paper, development of an on-chip analyzing method for typing of DNA methylation at a specific cytosine is reported. ICON probes fixed to the bottom of microwells assisted the on-chip detection of the methylation status of a specific cytosine in the target DNA.

2. Materials and Methods

2.1. Synthesis of an ICON Probe. Artificial DNA was synthesized by the conventional phosphoramidite method using an Applied Biosystems 392 DNA/RNA synthesizer or an NTS H-6 DNA/RNA synthesizer. The phosphoramidite form of bipyridine-modified adenine (B) was prepared according to the synthetic protocol described in a previous paper [13]. The 5'-amino end was attached using the phosphoramidite of 5'-aminomodifier C12 (Glen

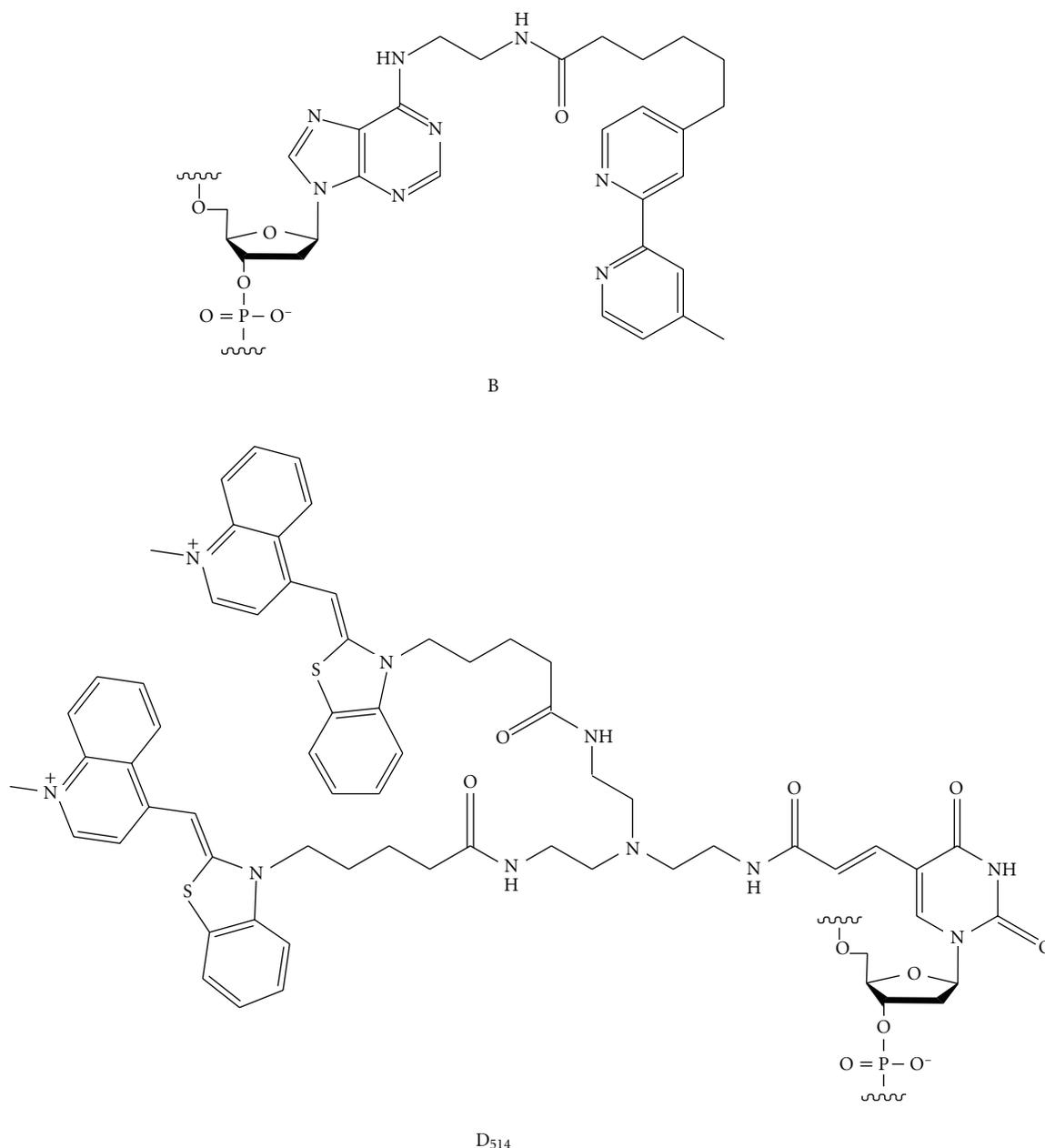


FIGURE 1: Structures of the “B” nucleotide of the ICON probe for 5-methylcytosine selective crosslink formation and the “D₅₁₄” nucleotide of the Exciton primer for real-time PCR monitoring.

Research (<http://www.glenresearch.com/index.php>). Synthesized DNA was purified by reverse phase HPLC on a 5-ODS-H column (10 mm × 150 mm, elution with a solvent mixture of 0.1 M triethylammonium acetate (TEAA), pH 7.0, linear gradient over 30 min from 5% to 20% acetonitrile at a flow rate of 3.0 mL/min). The DNA strand was characterized by MALDI-TOF MS. 5'-NH₂-(CH₂)₁₂-CCCCCCCCCACAACCTCCBTCATGTGCTGAA-3' ([M + H]⁺, calcd. C₃₄₃H₄₅₂N₁₁₈O₁₉₈P₃₃, 10418.0, found 10416.0).

2.2. *Preparation of Chips.* A 100 μL solution of synthetic DNA (100 nM) in TE buffer (pH 7.0) or deionized water

in the presence of 10 mM 1-ethyl-3-(3-dimethylamino-propyl)carbodiimide, and 10 mM 1-methylimidazole was put into each well of NucleoLink strips (Nalge Nunc (<http://www.nalgenunc.com/>)). After incubation at 50°C for 5 h, the reaction mixture was removed from the well, and the well was rinsed three times with a solution of 100 mM Tris-HCl (pH 7.5), 150 mM sodium chloride, and 0.1% Tween20, and then three times with deionized water.

2.3. *Osmium Treatment and DNA Amplification.* The target DNA sequence p53(N¹-N²) was 5'-TGT GCA GCT GTG G GTT GAT T CGA CAC CCC C GCC CGG CAC C CGC

GTC CGC G CCA TGG CCA T CTA CAA GCA G TCA CAG CAC A TGA N¹GG AGG T TGT GAG G N²G C TGC CCC CAC C-3' (N¹, N² = C or ^mC). A 25 μ L solution of DNA (8 nM) in 50 mM Tris-HCl buffer (pH 7.7), 0.5 mM EDTA, and 1 M sodium chloride was added to each probe-attached well at 0°C. The reaction mixture was incubated at 0°C for 5 min, and then the solution was removed from the wells. A 25 μ L solution of 5 mM potassium osmate(VI) and 100 mM potassium hexacyanoferrate(III) in 50 mM Tris-HCl buffer (pH 7.7), 0.5 mM EDTA, and 1 M sodium chloride was incubated at 0°C for 5 min or at 25°C for 10 min. The wells were rinsed seven times with 0.4 M sodium hydroxide, 0.25% Tween20 (130 μ L/well). After further rinsing with deionized water twice, the wells were coated with 10 mg/mL BSA in 100 mM Tris-HCl (pH 7.5), 150 mM sodium chloride, and 0.1% Tween20. The process of PCR amplification was performed in a reaction solution of 1 U *TaKaRa Ex Taq* HS, 10 \times buffer, 2.5 mM dNTP mix, and 1 μ M primer mix (Forward, 5'-AGCTGD₅₁₄GGGTTGATTC-3' for the Exciton primer method, 5'-TGTGCAGCTGTGGGTTGATTC-3' for the SYBR Green I method; reverse, 5'-ACTGCTTGTAGATGGCCATG-3'; D₅₁₄ in an Exciton primer is a hybridization-sensitive fluorescent nucleotide (Figure 1)). In the case of using the stain method with SYBR Green I fluorescence for monitoring the amplification, SYBR Green I dye was also added to the reaction mixture in advance. Amplifications were performed in microwells as follows: after heating at 95°C for 60 s, 35 cycles of denaturation at 95°C for 5 s, annealing with fluorescence monitoring at 52°C for 20 s, and extension at 72°C for 20 s on the Corbett Rotor-Gene, the amplification process was monitored by the fluorescence of D₅₁₄ or SYBR Green I through an SYBR Green I (470 nm/510 nm) filter.

3. Results and Discussion

3.1. Preparation of Chips and Osmium Complexation. For the on-chip study, we adopted a NucleoLink strip, because it is a microwell strip in which the amino-modified ICON probe can be attached to well bottoms with covalent bonds. The target DNA was a fragment of the human p53 gene exon 5 including mutation hotspots at CpG dinucleotides [20]. The sequence of the ICON probe was designed to form a crosslink with one of the methylated cytosines in the target DNA. The ICON probe was modified with an amino end for attaching to the plate and an alkyl linker and a poly C sequence for introducing enough distance between the probe and the plate. This ICON probe was synthesized by the conventional phosphoramidite method by using a DNA autosynthesizer. The phosphoramidite form of bipyridine-modified adenine (B) was prepared according to the synthetic protocol described in a previous paper [13]. Synthetic DNA was fixed to the bottom of each well of NucleoLink strips in the presence of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and 1-methylimidazole (Figure 2). The wells were rinsed and used for the next reaction.

Oxidative osmium complexation using ICON probes is a rapid and mild reaction for detection of methylated DNA, and it is not accompanied with nonspecific strand damage, in contrast to conventional bisulfite methods. Therefore, this reaction would be effective for on-chip analysis of DNA methylation. The target DNA was put into the wells and hybridized with the ICON probes fixed in the wells. The DNA samples in the wells were incubated at 0°C for 5 min then at 25°C for 5 min in the presence of 5 mM potassium osmate(VI) and 100 mM potassium hexacyanoferrate(III) in 50 mM Tris-HCl buffer (pH 7.7), 0.5 mM EDTA, and 1 M sodium chloride. After reaction, the wells were rinsed with 4 M sodium hydroxide and then coated with BSA.

3.2. Exciton Primer and Real-Time PCR Monitoring. The fixed DNA was detected using PCR amplification of a part of the crosslinking DNA strand. The PCR primers were designed for the region without ICON binding. For one of the primers, a hybridization-sensitive fluorescent DNA was used, containing a fluorescent nucleotide D₅₁₄ (Figure 1) [21–24]. This fluorescent DNA shows very weak fluorescence in the unhybridized state, whereas it shows strong fluorescence after hybridization with the complementary nucleic acids. This fluorescence switching is controlled by an intramolecular excitonic interaction between dyes tethered to the DNA. This fluorescent DNA is useful as a PCR primer. This DNA shows weak fluorescence emission, whereas the PCR mixture emits strong fluorescence after PCR amplification. This system has been applied to the detection of single nucleotide polymorphisms in genome DNA samples [25]. We tested this “Exciton primer” for on-chip PCR. The process of PCR amplification was performed in a reaction solution of *TaKaRa Ex Taq* HS polymerase in the presence of a mixture of dNTP and primer mix. Amplifications were performed in microwells and the change in the fluorescence intensity monitored using a real-time PCR system. In the experiment for p53(^mC–C) and p53(C–C), the methylation of the target cytosine was determined from the increase in the fluorescence signal associated with the exponential growth of the PCR product. Amplification of the p53(^mC–C) started first, and then the amplification of p53(C–C) started several cycles later (Figure 3). Washing out of uncrosslinked sample DNA brought about this lag in the starting point of amplification. The amplification curve observed for p53(C–C) almost overlapped that for the osmium-untreated p53(^mC–C), suggesting that the amplification curve for unmethylated DNA is due to amplification of the DNA nonspecifically adsorbing to the well surface. On-chip capture of methylated DNA by the ICON probe at the methylation site made possible the sequence specific detection of methylation through PCR amplification.

3.3. Sequence-Specific Amplification. A prototype for ICON-based on-chip methylation analysis makes possible sequence-specific amplification. We prepared four DNA strands with different methylation sites, p53(C–C), p53(^mC–C), p53(C–^mC), and p53(^mC–^mC). The sample DNA was added to the wells, in which the ICON probe targeting only the ^mC of

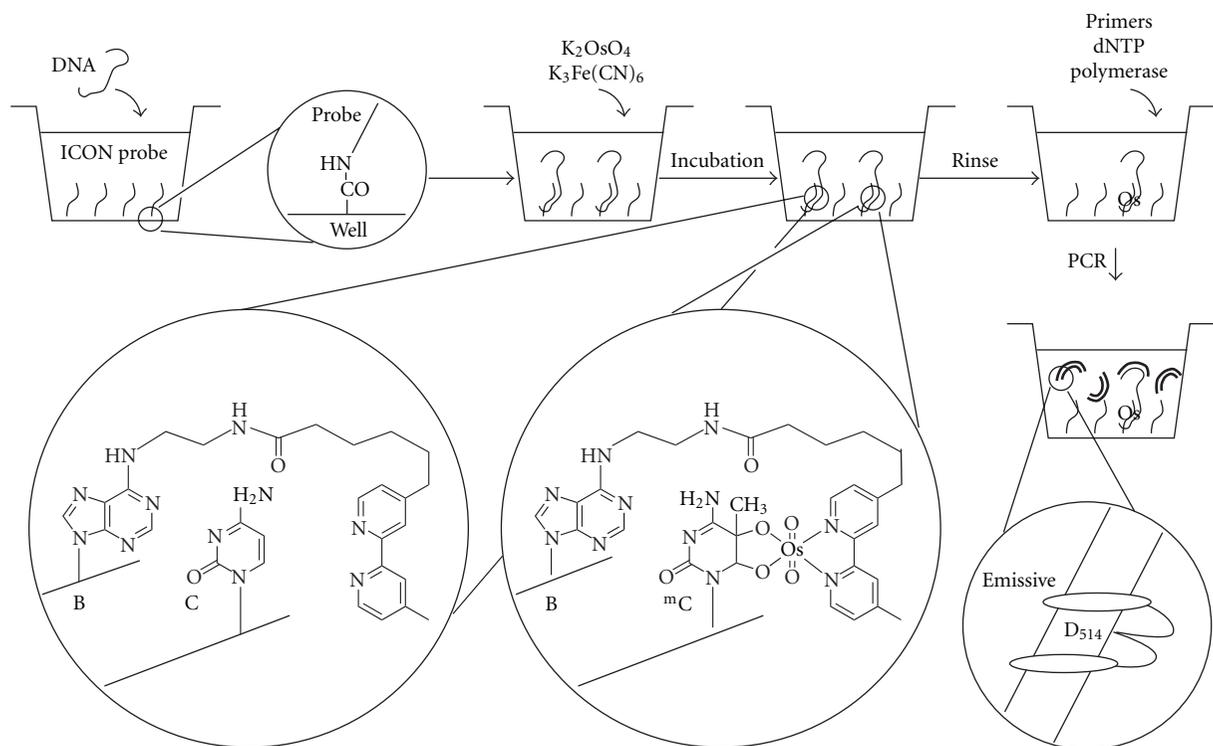


FIGURE 2: Schematic illustration of on-chip analysis of methylated DNA.

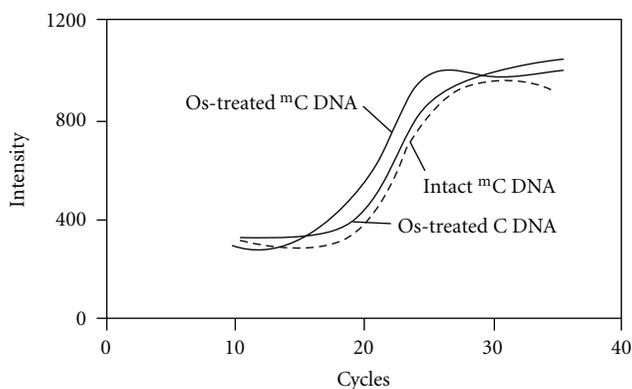


FIGURE 3: Profile of real-time PCR for p53(^mC-C) and p53(C-C) with/without osmium oxidation.

the sample DNA 5' side was fixed. After osmium treatment and BSA coating, the crosslinked DNA was amplified by PCR in the presence of unlabeled primers and SYBR Green I. After 15 cycles of the amplification reaction, the fluorescence intensity of SYBR Green I was quantified on a microplate reader (Figure 4). The wells containing p53(^mC-C) and p53(^mC-^mC) exhibited higher fluorescence intensities compared with those from wells containing p53(C-C) and p53(C-^mC). The ICON probe fixed on the well bottom distinguished 5'-^mC from 3'-^mC and detected only methylation of 5'-C regardless of methylation of 3'-C.

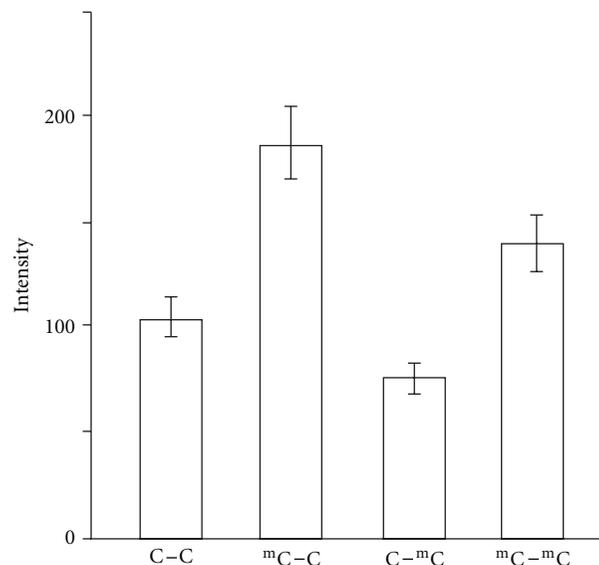


FIGURE 4: Sequence-specific detection of DNA methylation at N¹ in p53(N¹-N²) using an ICON probe.

4. Conclusions

We have described a new, high-value aspect of on-chip methylation analysis through osmium-DNA complexation. An ICON probe fixed onto a microwell formed a crosslink with the target 5-methylcytosine and assisted the detection

using PCR amplification. The crosslink was sequence-selective and completely independent of the other methylation site. Although there remain further aspects to be examined toward realizing an easier-to-use methylation analysis, such as optimization of PCR conditions suitable for ICON, this on-chip assay supported by the chemical basis could be an important component of the next generation of high-throughput methylation analyses.

Acknowledgment

The authors thank Dr. Takehiro Suzuki (RIKEN) for the MALDI-TOF mass spectrometry.

References

- [1] P. L. Jones and A. P. Wolffe, "Relationships between chromatin organization and DNA methylation in determining gene expression," *Seminars in Cancer Biology*, vol. 9, no. 5, pp. 339–347, 1999.
- [2] P. H. Tate and A. P. Bird, "Effects of DNA methylation on DNA-binding proteins and gene expression," *Current Opinion in Genetics and Development*, vol. 3, no. 2, pp. 226–231, 1993.
- [3] V. Colot and J. L. Rossignol, "Eukaryotic DNA methylation as an evolutionary device," *BioEssays*, vol. 21, no. 5, pp. 402–411, 1999.
- [4] R. Feil and S. Khosla, "Genomic imprinting in mammals: an interplay between chromatin and DNA methylation?" *Trends in Genetics*, vol. 15, no. 11, pp. 431–435, 1999.
- [5] M. F. Kane, M. Loda, G. M. Gaida et al., "Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines," *Cancer Research*, vol. 57, no. 5, pp. 808–811, 1997.
- [6] A. Okamoto, K. Tanabe, and I. Saito, "Site-specific discrimination of cytosine and 5-methylcytosine in duplex DNA by peptide nucleic acids (PNA)," *Journal of the American Chemical Society*, vol. 124, no. 35, pp. 10262–10263, 2002.
- [7] A. Okamoto, "Synthesis of highly functional nucleic acids and their application to DNA technology," *Bulletin of the Chemical Society of Japan*, vol. 78, no. 12, pp. 2083–2097, 2005.
- [8] H. Hayatsu, Y. Wataya, K. Kai, and S. Iida, "Reaction of sodium bisulfite with uracil, cytosine, and their derivatives," *Biochemistry*, vol. 9, no. 14, pp. 2858–2866, 1970.
- [9] M. Frommer, L. E. McDonald, D. S. Millar et al., "A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 5, pp. 1827–1831, 1992.
- [10] J. G. Herman, J. R. Graff, S. Myöhänen, B. D. Nelkin, and S. B. Baylin, "Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 18, pp. 9821–9826, 1996.
- [11] M. Mizugaki, K. Itoh, T. Yamaguchi et al., "Preparation of a monoclonal antibody specific for 5-methyl-2'-deoxycytidine and its application for the detection of DNA methylation levels in human peripheral blood cells," *Biological and Pharmaceutical Bulletin*, vol. 19, no. 12, pp. 1537–1540, 1996.
- [12] C. Montpellier, C. A. Bourgeois, N. Kokalj-Vokac et al., "Detection of methylcytosine-rich heterochromatin on banded chromosomes: application to cells with various status of DNA methylation," *Cancer Genetics and Cytogenetics*, vol. 78, no. 1, pp. 87–93, 1994.
- [13] K. Tanaka, K. Tainaka, T. Umemoto, A. Nomura, and A. Okamoto, "An osmium-DNA interstrand complex: application to facile DNA methylation analysis," *Journal of the American Chemical Society*, vol. 129, no. 46, pp. 14511–14517, 2007.
- [14] A. Nomura, K. Tainaka, and A. Okamoto, "Osmium complexation of mismatched DNA: effect of the bases adjacent to mismatched 5-methylcytosine," *Bioconjugate Chemistry*, vol. 20, no. 3, pp. 603–607, 2009.
- [15] A. Okamoto, "Chemical approach toward efficient DNA methylation analysis," *Organic and Biomolecular Chemistry*, vol. 7, no. 1, pp. 21–26, 2009.
- [16] A. Okamoto, K. Tainaka, and T. Kamei, "Sequence-selective osmium oxidation of DNA: efficient distinction between 5-methylcytosine and cytosine," *Organic and Biomolecular Chemistry*, vol. 4, no. 9, pp. 1638–1640, 2006.
- [17] T. Umemoto and A. Okamoto, "Synthesis and characterization of the 5-methyl-2'-deoxycytidine glycol-dioxosmium-bipyridine ternary complex in DNA," *Organic and Biomolecular Chemistry*, vol. 6, no. 2, pp. 269–271, 2008.
- [18] K. Tanaka, K. Tainaka, and A. Okamoto, "Methylcytosine-selective fluorescence quenching by osmium complexation," *Bioorganic and Medicinal Chemistry*, vol. 15, no. 4, pp. 1615–1621, 2007.
- [19] K. Tanaka, K. Tainaka, T. Kamei, and A. Okamoto, "Direct labeling of 5-methylcytosine and its applications," *Journal of the American Chemical Society*, vol. 129, no. 17, pp. 5612–5620, 2007.
- [20] S. Tornaletti and G. P. Pfeifer, "Complete and tissue-independent methylation of CpG sites in the p53 gene: implications for mutations in human cancers," *Oncogene*, vol. 10, no. 8, pp. 1493–1499, 1995.
- [21] S. Ikeda and A. Okamoto, "Hybridization-sensitive on-off DNA probe: application of the exciton coupling effect to effective fluorescence quenching," *Chemistry—An Asian Journal*, vol. 3, no. 6, pp. 958–968, 2008.
- [22] A. Okamoto, "Excitonic interaction: another photophysical process for fluorescence-controlled nucleic acid sensing," *Chemical Record*, vol. 10, no. 3, pp. 188–196, 2010.
- [23] T. Kubota, S. Ikeda, and A. Okamoto, "Doubly thiazole orange-labeled DNA for live cell RNA imaging," *Bulletin of the Chemical Society of Japan*, vol. 82, no. 1, pp. 110–117, 2009.
- [24] S. Ikeda, T. Kubota, M. Yuki, and A. Okamoto, "Exciton-controlled hybridization-sensitive fluorescent probes: multi-color detection of nucleic acids," *Angewandte Chemie International Edition*, vol. 48, no. 35, pp. 6480–6484, 2009.
- [25] A. Lezhava, T. Ishida, Y. Ishizu et al., "Exciton primer-mediated SNP detection in SmartAmp2 reactions," *Human Mutation*, vol. 31, no. 2, pp. 208–217, 2010.

Research Article

Influence of a Hairpin Loop on the Thermodynamic Stability of a DNA Oligomer

Jurij Lah, Mojca Seručnik, and Gorazd Vesnaver

Faculty of Chemistry and Chemical Technology, University of Ljubljana, Askerceva 5, 1000 Ljubljana, Slovenia

Correspondence should be addressed to Jurij Lah, jurij.lah@fkkt.uni-lj.si and Gorazd Vesnaver, gorazd.vesnaver@fkkt.uni-lj.si

Received 13 April 2011; Accepted 16 May 2011

Academic Editor: Luis A. Marky

Copyright © 2011 Jurij Lah et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

DSC was used to evaluate the mechanism of the thermally induced unfolding of the single-stranded hairpin $HP = 5'-CGGAATTCCTCCGGAATTCCTCCG-3'$ and its core duplex $D = (5'-CGGAATTCCTCCG-3')_2$. The DSC melting experiments performed at several salt concentrations were successfully described for HP and D in terms of a three-state transition model $HP \leftrightarrow I$ (intermediate state) $\leftrightarrow S$ (unfolded single-stranded state) and two state transition model $D \leftrightarrow 2S$, respectively. Comparison of the model-based thermodynamic parameters obtained for each HP and D transition shows that in unfolding of HP only the $HP \leftrightarrow I$ transition is affected by the TCTC loop. This observation suggests that in the intermediate state its TCTC loop part exhibits significantly more flexible structure than in the folded state while its duplex part remains pretty much unchanged.

1. Introduction

Hairpin loops are a common form of nucleic acid secondary structure and are crucial for tertiary structure and function [1]. They are known to play a key role in a number of biological processes such as gene expressions, DNA recombination, and DNA transposition [2–4]. In RNA molecules hairpins act as nucleation sites for RNA folding into final conformations [5–7] and play a critical role in RNA-protein recognition and gene regulation [8, 9]. Furthermore, due to the specificity of probe/target hybridization determined as a match-versus-mismatch discrimination, hairpin DNA oligomer probes have become an important tool in modern biotechnology and diagnostics [10, 11]. The thermodynamics and kinetics of hairpin formation, hairpin binding to complementary nucleic acids, and hairpin-ligand associations have been studied extensively [12–21]. There is no doubt that studies of hairpin-to-coil transitions and hairpin-ligand binding affinity and specificity have greatly enhanced our understanding of structural features and function of the naturally occurring nucleic acids [22, 23]. However, despite extensive biophysical research on the systems involving hairpin structures that produced a number of high-quality explanations and evaluations on properties and behavior of nucleic acids containing hairpin formations, there are still many unresolved questions.

As pointed out by Marky et al. [24] the most suitable hairpin molecules for studying the thermodynamics of their conformational transitions and ligand binding are the single-stranded hairpin molecules. They form stable partially paired duplexes that tend to melt in simple monomolecular transitions. Furthermore, their conformational stability and ligand binding properties are easily compared with those of the corresponding core duplexes. In this way one can evaluate the contributions of the loops to the thermal stability of the hairpins. Despite the simple structure of single-stranded hairpins it is not clear whether their monomolecular folding/unfolding transitions occur in a two-state or multistate manner. Measurements of their thermally induced unfolding transitions followed by UV, CD, and/or fluorescence spectroscopy as a rule result in sigmoidal melting curves suggesting that they may be considered as two-state processes. The same conclusion can be reached also on the basis of DSC measurements performed on the same sample solutions in the older generation of less sensitive DSC instruments (e.g., Microcal MC-2) which resulted in single-peak DSC thermograms. Recent measurements of conformational transitions of DNA quadruplex structures have shown, however, that the sigmoidal shape of UV or CD melting curves may be misleading. Namely, the DSC measurements performed on samples for which sigmoidal UV and CD melting curves were



FIGURE 1: Schematic presentation of the model oligonucleotides: hairpin (HP), duplex (D).

observed using the DSC of the latest generation (CSC, Microcal) resulted in thermograms containing two or three well-distinguished peaks thus indicating that the observed DNA melting process occurs in a multi-state manner [25, 26]. Furthermore, recent T-jump experiments performed on small hairpin molecules have produced a direct evidence that their unfolding transitions involve intermediate structures and thus cannot be considered as two-state processes [19, 27–29].

In our DSC study of the unfolding mechanism and stability of the 5'-CGGAATTCCGTCTCCGGAATTCCG-3' hairpin we performed the DSC melting experiments on the hairpin and its core duplex, (5'-CGGAATTCCG-3')₂ (see Figure 1), at several salt concentrations using an extremely sensitive microcalorimeter (CSC). To see to what extent the TCTC loop affects the hairpin unfolding process we attempted to describe for each oligonucleotide the measured DSC thermograms in terms of the simplest possible unfolding model. We derived the corresponding model functions and by fitting them to the experimental data we tested the appropriateness of the suggested models and obtained for each transition the characteristic thermodynamic quantities of transition $\Delta G_{(T)}^0$, $\Delta H_{(T)}^0$, $\Delta S_{(T)}^0$, and Δc_p^0 . By comparing these values determined for the hairpin and the core duplex we tried to estimate the contribution of the TCTC loop to the stability of the hairpin.

2. Materials and Methods

2.1. Materials. Self-complementary oligonucleotide 5'-CGGAATTCCG-3' and oligonucleotide 5'-CGGAATTCCGTCTCCGGAATTCCG-3' that in solution at room temperature form a duplex (D) and a single-stranded hairpin structure (HP), respectively, were purchased HPLC pure from Invitrogen Co., Germany and used without any further purification. Their concentrations in buffer solution (10 mM phosphate buffer and 1 mM Na₂EDTA adjusted to pH = 7.0) in the presence of 100 mM NaCl were determined at 25°C spectrophotometrically in the Cary Bio 100 UV-spectrophotometer. The molar extinction coefficients were determined using the nearest neighbor data of Cantor et al. [30] for single-stranded DNA at 25°C and the absorbance at 260 nm of thermally unfolded oligonucleotide extrapolated back to 25°C ($\epsilon_{D260} = 84600 \text{ M}^{-1}\text{cm}^{-1}$, $\epsilon_{HP260} = 216000 \text{ M}^{-1}\text{cm}^{-1}$). The phosphate buffer solutions used in all experiment contained 0, 0.1, 0.3, or 1.0 M NaCl.

Differential scanning calorimetry (DSC). Thermally induced unfolding of duplex (D) and hairpin (H) in buffer solutions with different added NaCl concentrations was followed between 5 and 95°C in a Nano-II DSC calorimeter (CSC; UT) at the heating rate of 1°C/min and essentially

the same results were obtained from several test-experiments performed at the heating rate of 0.25°C/min. The thermally induced unfolding of both oligonucleotides was monitored in terms of $c_{P_{ex}} = \bar{c}_{p_2} - \bar{c}_{p_{D,F}}$ versus T thermograms in which the differences between the partial molar heat capacity of the measured oligonucleotide \bar{c}_{p_2} (raw signals corrected for the solvent contributions) and the partial molar heat capacities of the corresponding folded states extrapolated from low temperatures over the whole measured temperature interval, $\bar{c}_{p_{D,F}}$, are normalized for the duplex or hairpin concentration. The total enthalpy of unfolding, $\Delta H_{(T)}^{\text{cal}}$, was obtained from the measured thermograms as the area under the $c_{P_{ex}}(T)$ versus T curve.

2.2. Analysis of the DSC Thermograms. The thermally induced conformational transitions can be experimentally followed in a model-independent way only by DSC. At relatively low concentrations used in DSC experiments the measured solute-normalized heat capacity of the sample solution, $c_{P(T)}$, with the subtracted baseline may be equalized with the oligonucleotide partial molar heat capacity, $\bar{c}_{p_2(T)}$. Thus, the overall heat effect that accompanies the measured conformational transition from its initial folded state at the temperature T_1 to its final unfolded state at T_2 can be expressed as

$$\Delta H_{(T_1 \rightarrow T_2)} = \int_{T_1}^{T_2} \bar{c}_{p_2(T)} dT. \quad (1)$$

Since the enthalpy is the state function, the enthalpy change $\Delta H_{(T_1 \rightarrow T_2)}$ may be expressed also as

$$\begin{aligned}
 \Delta H_{(T_1 \rightarrow T_2)} = & \int_{T_1}^{T_{\text{ref}}} (\bar{c}_{p_2(T)})_F dT + \Delta H_{(T_{\text{ref}})} \\
 & + \int_{T_{\text{ref}}}^{T_2} (\bar{c}_{p_2(T)})_U dT,
 \end{aligned} \quad (2)$$

where $(\bar{c}_{p_2(T)})_F$ and $(\bar{c}_{p_2(T)})_U$ are the partial molar heat capacities of the folded and unfolded DNA conformation, respectively, $\Delta H_{(T_{\text{ref}})}$ is the enthalpy of unfolding at T_{ref} which can be any temperature between T_1 and T_2 . By choosing $T_{\text{ref}} = T_{1/2}$ where $T_{1/2}$ is the melting temperature at which a half of oligonucleotide molecules undergo unfolding transition (2) transforms into

$$\begin{aligned}
 \Delta H_{T_{(1/2)}}^{\text{cal}} = & \int_{T_1}^{T_{1/2}} [\bar{c}_{p_2(T)} - (\bar{c}_{p_2(T)})_F] dT \\
 & + \int_{T_{1/2}}^{T_2} [\bar{c}_{p_2(T)} - (\bar{c}_{p_2(T)})_U] dT,
 \end{aligned} \quad (3)$$

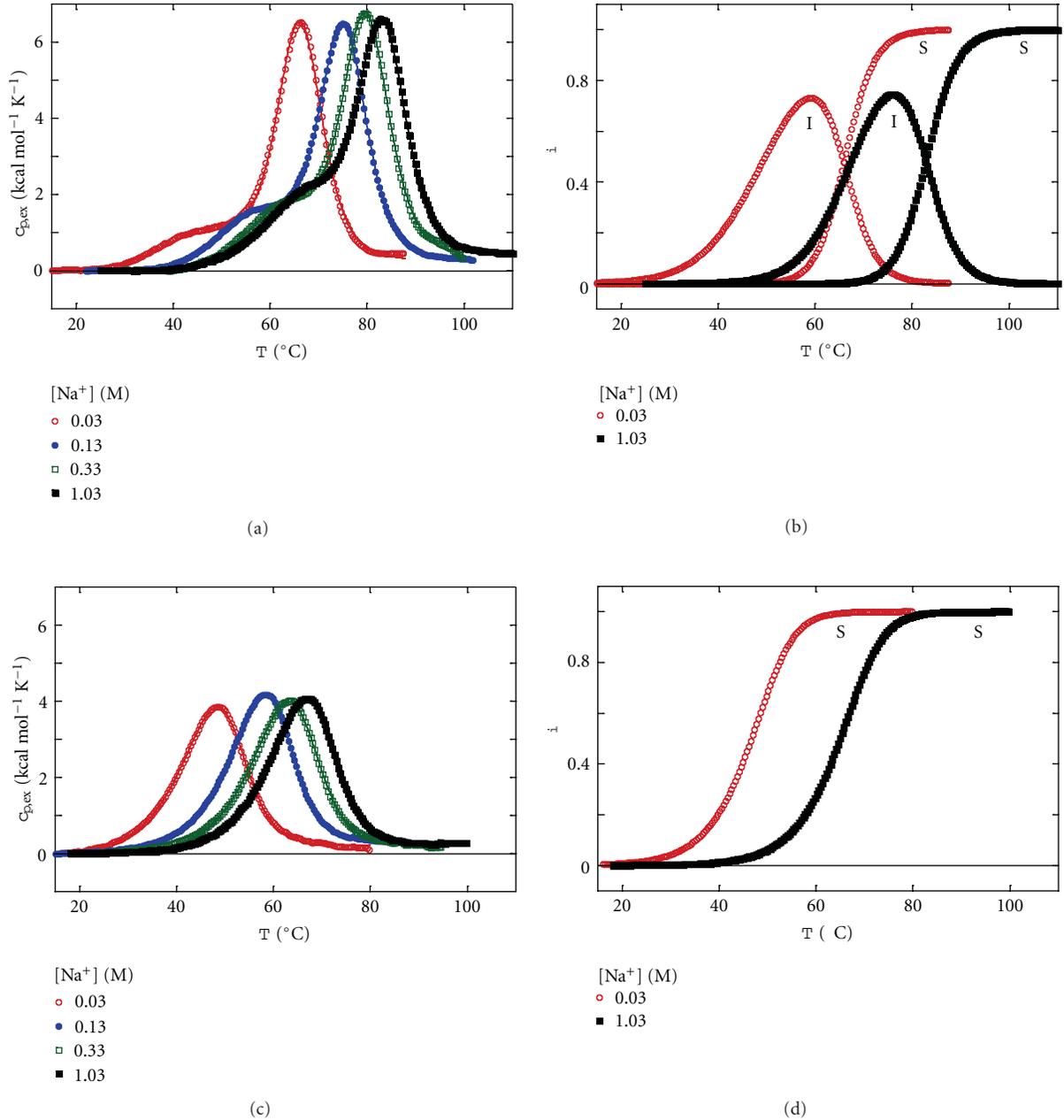
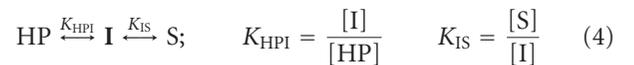


FIGURE 2: DSC thermograms and their model analysis: hairpin (HP) unfolding characterized in terms of a three-state model $HP \leftrightarrow I \leftrightarrow S$ (a) and the corresponding fractions of species (b); duplex (D) unfolding characterized in terms of a two-state model $D \leftrightarrow 2S$ (c) and the corresponding fractions of species (d). In panels (a) and (c) symbols represent experimental data points while lines of the same color correspond to the best-fit model functions ((14) and (18)).

where $\Delta H_{T_{1/2}}^{\text{cal}}$ is a model-independent enthalpy of transition at $T_{1/2}$ that can be easily determined by the appropriate integration of the experimental $[\bar{c}_{p_2(T)} - (\bar{c}_{p_2(T)})_F]$ and $[\bar{c}_{p_2(T)} - (\bar{c}_{p_2(T)})_U]$ curves as presented in (3).

According to the DSC thermograms of the measured hairpin (HP), its thermally induced unfolding involves at least two conformational transitions (Figure 2). Thus, the simplest suggested model to describe the observed thermal behavior would consist of two consecutive monomolecular transitions: HP (hairpin) \leftrightarrow I (intermediate state) \leftrightarrow S

(unfolded single-stranded state). The enthalpy, H , of a solution containing an HP sample characterized by the suggested thermal unfolding



can be expressed at given P and T as

$$H = n_1 \bar{H}_1 + n_2 \bar{H}_2 = n_1 \bar{H}_1 + n_{\text{HP}} \bar{H}_{\text{HP}} + n_I \bar{H}_I + n_S \bar{H}_S, \quad (5)$$

where K_{HP1} and K_{IS} are the corresponding equilibrium constants, the quantities in brackets are the equilibrium molar concentrations of HP, I, and S, n_1 is the number of moles of solvent, and n_2 is the number of moles of solute (oligonucleotide) that can be further expressed as:

$$n_2 = n_{\text{HP}} + n_1 + n_{\text{S}}, \quad (6)$$

n_{HP} in (6) represents the number of moles of the oligonucleotide in the folded hairpin state, n_1 is the number of moles in the intermediate state, n_{S} is the number of moles in the unfolded single-stranded state and \bar{H}_1 , \bar{H}_2 , \bar{H}_{HP} , \bar{H}_I , and \bar{H}_S are the corresponding partial molar enthalpies of the solvent, solute and folded, intermediate and unfolded oligonucleotide, respectively. By defining the molar fraction, α_i , of the solute species, i , as $\alpha_i = n_i/n_2$ one obtains from (5) that

$$\bar{H}_2 = \alpha_{\text{HP}}\bar{H}_{\text{HP}} + \alpha_I\bar{H}_I + \alpha_S\bar{H}_S. \quad (7)$$

Finally, by introducing $\alpha_{\text{HP}} = 1 - \alpha_I - \alpha_S$ into (7) and taking the temperature derivative of the modified (7) one obtains the model function for the measured DSC signal, $c_{P,\text{ex}}$, expressed as

$$\begin{aligned} c_{P,\text{ex}} &= \bar{c}_{P,2} - \bar{c}_{P,\text{HP}} \\ &= \alpha_I \Delta c_{P,\text{HP1}} + \frac{d\alpha_I}{dT} \Delta H_{\text{HP1}} + \alpha_S (\Delta c_{P,\text{HP1}} + \Delta c_{P,\text{IS}}) \\ &\quad + \frac{d\alpha_S}{dT} (\Delta H_{\text{HP1}} + \Delta H_{\text{IS}}), \end{aligned} \quad (8)$$

in which at any temperature $\bar{c}_{P,2}$ is the measured c_P (with subtracted baseline), $\bar{c}_{P,\text{HP}}$ is the partial molar heat capacity of HP extrapolated from low-temperature region over the entire measured temperature interval, $\Delta H_{\text{HP1}} = \bar{H}_1 - \bar{H}_{\text{HP}}$ (enthalpy of the hairpin to intermediate state transition), $\Delta H_{\text{IS}} = \bar{H}_S - \bar{H}_I$ (enthalpy of the intermediate state to unfolded single stranded state transition), $\Delta c_{P,\text{HP1}} = \bar{c}_{P,I} - \bar{c}_{P,\text{HP}}$ and $\Delta c_{P,\text{IS}} = \bar{c}_{P,S} - \bar{c}_{P,I}$.

$c_{P,\text{ex}}$ can be obtained experimentally simply by subtracting the hairpin $\bar{c}_{P,\text{HP}}$ versus T curve extrapolated from low- T region over the entire measured temperature interval from the corresponding measured $\bar{c}_{P,2}$ versus T curve. The model-based $c_{P,\text{ex}}$, however, can be calculated from the right-hand side term of (8). According to the suggested model (4) the total solute molar concentration, c_T , and the fractions α_i of the solute species present in the solution can be expressed as $c_T = [\text{HP}] + [\text{I}] + [\text{S}]$ and $\alpha_{\text{HP}} = [\text{HP}]/c_T$, $\alpha_I = [\text{I}]/c_T$ and $\alpha_S = [\text{S}]/c_T$. Since $\alpha_{\text{HP}} + \alpha_I + \alpha_S = 1$ one obtains from (4) that

$$\alpha_S = \frac{1}{(K_{\text{HP1}}K_{\text{IS}})^{-1} + K_{\text{S}}^{-1} + 1}, \quad \alpha_I = \frac{\alpha_S}{K_{\text{IS}}}. \quad (9)$$

For the description of the DSC experiment with the model function (8) one needs also the temperature derivatives of α_S and α_I . By using for each transition, i , the van't Hoff relation

$$\frac{d \ln K_i}{dT} = \frac{\Delta H_i^0}{RT^2}, \quad (10)$$

one obtains

$$\begin{aligned} \frac{d\alpha_S}{dT} &= \alpha_S^2 \left[K_{\text{HP1}}^{-1} K_{\text{IS}}^{-1} \frac{(\Delta H_{\text{HP1}}^0 + \Delta H_{\text{IS}}^0)}{RT^2} + K_{\text{IS}}^{-1} \frac{\Delta H_{\text{IS}}^0}{RT^2} \right], \\ \frac{d\alpha_I}{dT} &= K_{\text{IS}}^{-1} \left(\frac{d\alpha_S}{dT} - \alpha_S \frac{\Delta H_{\text{IS}}^0}{RT^2} \right). \end{aligned} \quad (11)$$

Assuming that for each transition, i , the corresponding $\Delta c_{P,i}^0$ does not depend on T the standard free energy of that transition, $\Delta G_{i(T)}^0$, can be obtained at any T from the integrated form of the Gibbs-Helmholtz relation as

$$\begin{aligned} \Delta G_{i(T)}^0 &= T \left[\frac{\Delta G_{i(T_{i,1/2})}^0}{T_{i,1/2}} + \Delta H_{i(T_{i,1/2})}^0 \left(\frac{1}{T} - \frac{1}{T_{i,1/2}} \right) \right. \\ &\quad \left. + \Delta c_{P,i}^0 \left(1 - \frac{T_{i,1/2}}{T} - \ln \frac{T}{T_{i,1/2}} \right) \right], \end{aligned} \quad (12)$$

where $T_{i,1/2}$ is the temperature at which the α_i values of species participating in transition i are the same. The corresponding equilibrium constant, K_i , is related to $\Delta G_{i(T)}^0$ as

$$\Delta G_{i(T)}^0 = -RT \ln K_i, \quad (13)$$

and for the suggested mechanism of the hairpin unfolding (4) it can be easily seen that for each suggested monomolecular transition $\Delta G_{i(T_{i,1/2})}^0 = 0$. Finally, according to the DSC experiments performed at different oligonucleotide concentrations the $\Delta H_{i(T)}$ values appear to be concentration independent thus indicating that one may assume for each transition that $\Delta H_{i(T)} = \Delta H_{i(T_{i,1/2})}^0$ and $\Delta c_{P,i} = \Delta c_{P,i}^0$. Using these assumptions and (8)–(13) one can express the model function (14)

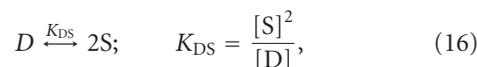
$$\begin{aligned} c_{P,\text{ex}} &= \alpha_I \Delta c_{P,\text{HP1}}^0 + \frac{d\alpha_I}{dT} \Delta H_{\text{HP1}}^0 + \alpha_S (\Delta c_{P,\text{HP1}}^0 + \Delta c_{P,\text{IS}}^0) \\ &\quad + \frac{d\alpha_S}{dT} (\Delta H_{\text{HP1}}^0 + \Delta H_{\text{IS}}^0), \end{aligned} \quad (14)$$

only in terms of parameters $T_{i,1/2}$, $\Delta H_{i(T_{i,1/2})}^0$, and $\Delta c_{P,i}^0$, characteristic for each of the suggested transitions. Their “best fit” values are obtained by fitting the model function (14) to the experimental $c_{P,\text{ex}}$ versus T curves. Furthermore, since for each transition, i , the corresponding $\Delta H_{i(T)}^0$ and $\Delta S_{i(T)}^0$ quantities can be expressed as

$$\begin{aligned} \Delta H_{i(T)}^0 &= \Delta H_{i(T_{i,1/2})}^0 + \Delta c_{P,i}^0 (T - T_{i,1/2}), \\ T \Delta S_{i(T)}^0 &= \Delta H_{i(T)}^0 - \Delta G_{i(T)}^0, \end{aligned} \quad (15)$$

the “best fit” parameters $T_{i,1/2}$, $\Delta H_{i(T_{i,1/2})}^0$ and $\Delta c_{P,i}^0$ can be used also to obtain the $\Delta H_{i(T)}^0$ and $\Delta S_{i(T)}^0$ values at any T .

In contrast to HP unfolding, the measured thermally induced duplex (D) to single strand (S) transition appears to be a simpler, all-or-none process



that can be described in terms of the total oligonucleotide concentration, c_T , the concentrations of the duplex form [D] and the single strands [S], the fraction of duplex molecules that undergo the unfolding transition at a given temperature, α_S , and the equilibrium constant K_{DS} interrelated as

$$c_T = [D] + \frac{[S]}{2}; \quad \alpha_S = \frac{[D]}{c_T}; \quad K_{DS} = \frac{4\alpha_S^2 c_T}{1 - \alpha_S}. \quad (17)$$

A similar, though much simpler derivation of the model function than the one presented for unfolding of the hairpin structure (14) leads for the suggested $D \leftrightarrow 2S$ transition to

$$c_{P,ex} = \bar{c}_{P,2} - \bar{c}_{P,D} = \alpha_S \Delta c_{P,DS}^0 + \frac{d\alpha_S}{dT} \Delta H_{DS}^0, \quad (18)$$

where $\bar{c}_{P,2}$ is the measured c_P of the sample solution with subtracted baseline, $\bar{c}_{P,D}$ is the heat capacity of the duplex form extrapolated from the low- T region over the whole measured temperature interval, $\Delta c_{P,DS}^0 = \Delta c_{P,DS} = 2\bar{c}_{P,S} - \bar{c}_{P,D}$ and $\Delta H_{DS}^0 = \Delta H_{DS} = 2\bar{H}_S - \bar{H}_D$. From the suggested model (16) and (17) it follows that $\Delta G_{DS(T_{1/2})}^0 = -RT \ln(2c_T)$ and

$$\alpha_S = \frac{1}{2} \left[-\frac{K_{DS}}{4c_T} + \sqrt{\left(\frac{K_{DS}}{4c_T}\right)^2 + \frac{K_{DS}}{c_T}} \right], \quad (19)$$

$$\frac{\partial \alpha_S}{\partial T} = \frac{\Delta H_{DS}^0}{RT^2} \frac{\alpha_S(1 - \alpha_S)}{2 - \alpha_S}.$$

The corresponding expressions for $\Delta G_{DS(T)}^0$, $\Delta H_{DS(T)}^0$ and $T\Delta S_{DS(T)}^0$ are the same as those shown for each transition in the suggested hairpin unfolding mechanism (12), (13) and (15). Similarly, in deriving (18) the $\Delta H_{DS(T)}$ and $\Delta c_{P,DS}$ are assumed to be independent on the oligonucleotide concentration and thus equal to $\Delta H_{DS(T)}^0$, and $\Delta c_{P,DS}^0$.

Inspection of (18) and (19) shows that the model function (18) is expressed in terms of adjustable parameters, $T_{1/2}$, $\Delta H_{DS(T_{1/2})}^0$ and $\Delta c_{P,DS}^0$ that can be determined by fitting the model function to the experimental $c_{P,ex}$ versus T curve and further used to determine the $\Delta G_{DS(T)}^0$, $\Delta H_{DS(T)}^0$ and $\Delta S_{DS(T)}^0$ values at any T .

To obtain a set of the “best fit” adjustable parameters $T_{i,1/2}$, $\Delta H_{i(T_{i,1/2})}^0$ and $\Delta c_{P,i}^0$ describing the hairpin and duplex thermal unfolding at each of the added salt concentrations the iterative nonlinear Levenberg-Marquardt χ^2 regression procedure was used [31]. Furthermore, assuming that for the observed transitions the accompanying $\Delta c_{P,i}^0$ quantities do not depend on the added NaCl concentration their values may be determined also from the slopes of the $\Delta H_{i(T_{i,1/2})}^0$ versus $T_{i,1/2}$ curves constructed from the “best fit” $\Delta H_{i(T_{i,1/2})}^0$ and $T_{i,1/2}$ parameters determined at different added salt concentrations [32]. These data can be also used to construct the corresponding $T_{i,1/2}$ versus $\ln[\text{Na}^+]$ plots from which the amount of the Na^+ ions released upon thermal unfolding of the hairpin or duplex structure can be estimated (see discussion, (20)).

3. Results and Discussion

According to the measured DSC thermograms presented in Figure 2 the thermally induced unfolding of the hairpin HP consists of at least two conformational transitions while the one observed for the duplex D occurs in a simpler “all or none” manner. In addition, UV melting experiments (not shown) resulted for HP in biphasic melting curves that exhibit transitions independent on HP concentration (monomolecular transitions) while for D monophasic melting curves dependent on D concentration (nonmonomolecular transition) were observed. Moreover, excellent repeatability of the consecutive measured heating and cooling c_P versus T curves and the observed independence of the measured DSC peaks on the applied heating rate (several test experiments) clearly shows that the studied thermal unfolding events may be considered as reversible processes. Model analysis of the measured thermograms shows that the hairpin thermal unfolding can be well described in terms of a three state model involving H (hairpin) \leftrightarrow I (intermediate state) \leftrightarrow S (unfolded single-stranded structure) transitions and the corresponding model function (14) characterized for each of the suggested transitions with the corresponding “best fit” adjustable parameters $T_{i,1/2}$, $\Delta H_{i(T_{i,1/2})}^0$, and $\Delta c_{P,i}^0$ (Table 1). However, analysis of the applied fitting procedure indicates that the parameter $\Delta c_{P,IS}^0$ is highly correlated to some other adjustable parameters. Thus, to obtain safe estimate of $\Delta c_{P,IS}^0$ another method of its determination has to be used. Assuming that it does not depend on the simple salt concentration $\Delta c_{P,IS}^0$ was estimated as a slope of the $\Delta H_{IS(T_{1/2})}^0$ versus $T_{IS,1/2}$ plot (Figure 3(a)) constructed from the “best fit” parameters determined at different NaCl concentrations (Table 1). This method of determining $\Delta c_{P,i}^0$ was justified by a good agreement between the $\Delta c_{P,i}^0$ values for other transitions obtained by the described fitting procedure and the $\Delta c_{P,i}^0$ values determined as the slopes of the corresponding $\Delta H_{i(T_{i,1/2})}^0$ versus $T_{i,1/2}$ plots (Table 1). By using the parameters presented in Table 1 one can calculate for the duplex and hairpin solutions the relative populations of the model-predicted species in the measured temperature interval and at all added salt concentrations (Figure 2). Evidently, the thermal stability of the folded state of the measured duplex and the hairpin is substantially enhanced by increasing the added salt concentration. At low salt concentrations, however, a small fraction of the hairpin molecules undergoes transition into the intermediate state already at physiological temperatures.

A standard way of testing the quality of a suggested model is to compare the enthalpy of unfolding determined at a given temperature directly by an appropriate integration of the experimental $c_{P,ex(T)}$ versus T curve ($\Delta H_{HS}^{\text{cal}}$, see (3)) with the corresponding model-based value ΔH_{HS}^0 calculated at the same temperature using the reported “best fit” adjustable parameters. As shown in Table 1 a good agreement was obtained which clearly supports the appropriateness of the suggested $H \leftrightarrow I \leftrightarrow S$ unfolding model.

It is well known that DNA unfolding is accompanied by release of counterions. The number of the released Na^+ ions,

TABLE 1: Thermodynamic parameters^a obtained from fitting the model functions ((14) and (18)) to the duplex (D) and hairpin (HP) DSC thermograms presented in Figure 2.

Transition	$T_{1/2}$	$\Delta H_{(T_{1/2})}^0$	$\Delta H_{(T_{1/2})}^{\text{cal}}$	Δc_p^0 ^b	Δc_p^0 ^c	Δn_{Na^+}
HP → I	59.2	37			0.34	0.9
I → S	75.0	74		0.37	0.25	1.5
HP → S		117 ^d	111		0.59	2.4
D → 2S	56.5	71	71	0.25	0.30	1.7
Error	±0.2	±2	±2	±0.05	±0.05	±0.2

^a Units: °C($T_{1/2}$), kcal mol⁻¹($\Delta H_{(T_{1/2})}^0$, $\Delta H_{(T_{1/2})}^{\text{cal}}$), kcal mol⁻¹ K⁻¹ (Δc_p^0); unless stated otherwise the values are those obtained at $[\text{Na}^+] = 0.13\text{M}$;

^b obtained from fitting the model function;

^c obtained as the slope of $\Delta H_{(T_{1/2})}^0$ versus $T_{1/2}$ curves (Figure 3(a));

^d the total enthalpy of the of the I → S transition was calculated as $\Delta H_{\text{HS}(T_{1/2,\text{IS}})}^0 = \Delta H_{\text{HI}(T_{1/2,\text{HI}})}^0 + \Delta c_{\text{PHI}}^0 (T_{1/2,\text{IS}} - T_{1/2,\text{HI}}) + \Delta H_{\text{IS}(T_{1/2,\text{IS}})}^0$ where $T_{1/2,\text{IS}}$ and $T_{1/2,\text{HI}}$ are the melting temperatures of the I → S and H → I transitions and $\Delta H_{\text{IS}(T_{1/2,\text{IS}})}^0$ and $\Delta H_{\text{HI}(T_{1/2,\text{HI}})}^0$ are the corresponding enthalpies of transition.

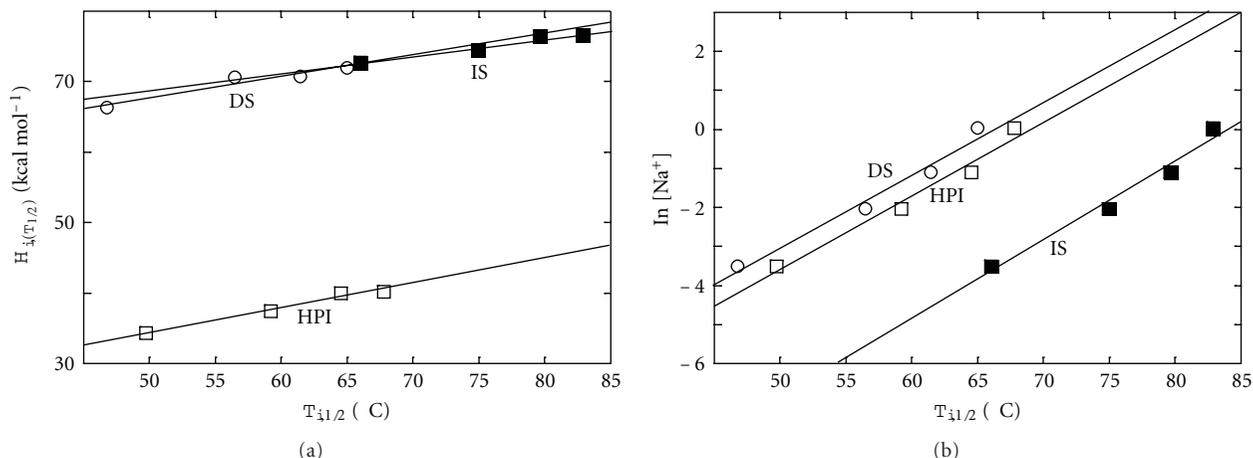


FIGURE 3: Estimation of heat capacity changes and number of released Na⁺ ions: (a) $\Delta c_{p,i}^0$ was determined for each transition i as the slope of the $\Delta H_{i(T_{i,1/2})}^0$ versus $T_{i,1/2}$ plot constructed from the model based $\Delta H_{i(T_{i,1/2})}^0$ and $T_{i,1/2}$ values determined at different salt concentrations; (b) the corresponding $\ln[\text{Na}^+]$ versus $T_{i,1/2}$ plots from which the $\Delta n_{\text{Na}^+,i}$ values were determined according to (20).

$\Delta n_{\text{Na}^+,i}$, upon each HP and D transition, expressed per mole of oligonucleotide, may be estimated from [33]

$$\frac{dT_{i,1/2}}{d\ln[\text{Na}^+]} = \frac{RT_{i,1/2}^2}{\Delta H_{i(T_{i,1/2})}^0} \Delta n_{\text{Na}^+,i}, \quad (20)$$

in which $T_{i,1/2}$ is the melting temperature at a given Na⁺ concentration, $[\text{Na}^+]$, $\Delta H_{i(T_{i,1/2})}^0$ is the corresponding enthalpy of transition at $T_{i,1/2}$. The $\Delta n_{\text{Na}^+,i}$ values presented in Table 1 were determined from the slopes of the $\ln[\text{Na}^+]$ versus $T_{i,1/2}$ plots (Figure 3(b)).

At any T in the measured range of physiological temperatures the difference between the given property characterizing the total unfolding of the hairpin H and the duplex D (for ex. $\Delta\Delta H_{(T)}^0 = \Delta H_{\text{HS}(T)}^0 - \Delta H_{\text{DS}(T)}^0$) reflects the contribution of the TCTC loop to that property relative to the core duplex (Figure 4, Table 2). Thus, the observed $\Delta\Delta H_{(T)}^0 > 0$ indicates a favorable energy contribution of the TCTC loop to the stability of the hairpin that results, very likely, from the increased number of stacking interactions (in the first place from those occurring at the core duplex-loop connections) [34]. The corresponding $\Delta\Delta S_{(T)}^0 > 0$ is

consistent with the highly positive $\Delta\Delta H_{(T)}^0$ indicating that the unfavorable entropy contribution of the TCTC loop to the hairpin stability arises largely from a substantial disruption of the loop structure accompanying the unfolding of the hairpin. The observed $\Delta\Delta c_p^0 > 0$ and $\Delta\Delta n_{\text{Na}^+} > 0$ show, however, that the loop contributions to $\Delta\Delta H_{(T)}^0$ and $\Delta\Delta S_{(T)}^0$ may be, to a certain extent, determined also by hydration [35] and electrostatic interactions. The $\Delta\Delta c_p^0 > 0$ suggests that within the folded hairpin conformation, not only the core duplex but also the TCTC loop are less exposed to water than in the unfolded state. In addition, the observed $\Delta\Delta n_{\text{Na}^+} > 0$ may be ascribed to a decrease in the surface charge density accompanying the unfolding of the oligonucleotide which is significantly more pronounced in the case of hairpin unfolding. Comparison of the $\Delta\Delta$ values for ΔG_T^0 , ΔH_T^0 , ΔS_T^0 , Δc_p^0 , and Δn_{Na^+} quantities determined for the HP → I and I → S transitions of the hairpin with the corresponding $\Delta\Delta$ values determined for the D → 2S transition of the duplex shows that the $\Delta\Delta$ values for the I → S and D → 2S transitions are very close (Figure 4, Table 2). Evidently, one may speculate that in the hairpin structure only the HP → I transition is influenced by the TCTC loop. In other words, it seems

TABLE 2: Difference thermodynamic stability parameters^a at 25°C exhibiting the influence of the TCTC loop to the unfolding features of the hairpin forming oligonucleotide (HP).

Transition	$\Delta\Delta G^0$	$\Delta\Delta H^0$	$T\Delta\Delta S^0$	$\Delta\Delta c_p^0$	$\Delta\Delta n_{Na^+}$
(HP → S)-(D → 2S)	2.4	26.6	24.3	0.29	0.7
(I → S)-(D → 2S)	-0.9	0.8	1.7	-0.05	-0.2
Error	±3	±3	±3	±0.07	±0.3

^aUnits: kcal mol⁻¹ ($\Delta\Delta G^0$, $\Delta\Delta H^0$, $T\Delta\Delta S^0$), kcal mol⁻¹ K⁻¹ ($\Delta\Delta c_p^0$); for temperature dependence of ΔG^0 , ΔH^0 , and $T\Delta S^0$ see Figure 4.

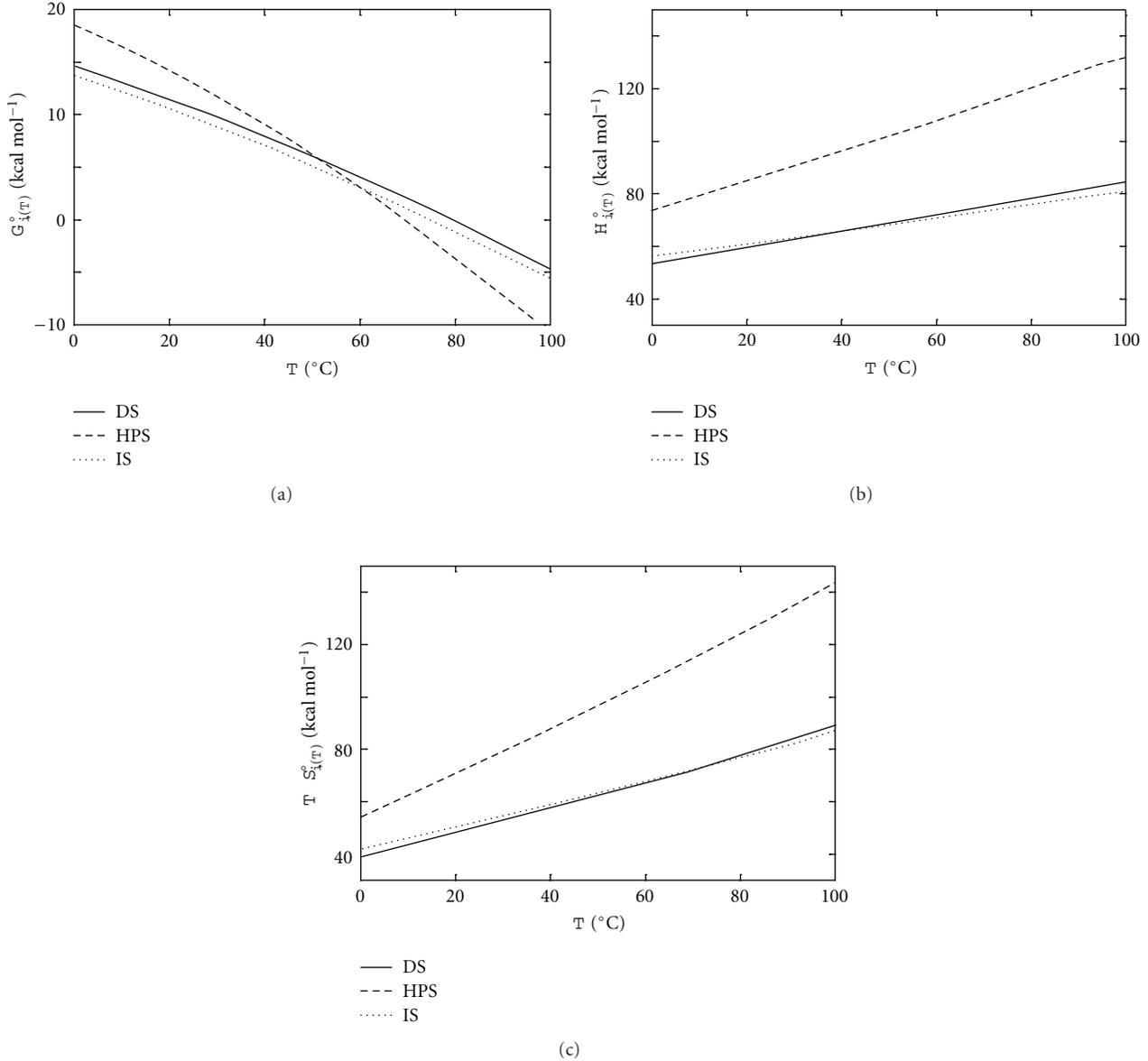


FIGURE 4: Thermodynamics of hairpin and duplex unfolding: standard Gibbs free energy (a), standard enthalpy (b), and the corresponding entropy contribution (c) presented for each model predicted hairpin (I → S, HP → S) and duplex (D → 2S) transitions as functions of temperature at $[Na^+] = 0.13M$.

that the observed HP → I transition reflects mainly the changes in the TCTC conformation. Thus, the intermediate state I may be considered as a state in which the core duplex remains more or less unchanged while the TCTC loop occurs as a more flexible structure characterized by the additional

stacking interactions and the freedom of the neighboring water molecules and ions similar to the one in the unfolded state.

To the best of our knowledge this is the first time that a three-state unfolding of a simple hairpin structure, observed

by DSC, has been reported and characterized thermodynamically. We believe that the main reason for this is that in most studies of thermal unfolding of hairpins too high starting temperatures have been chosen and therefore the low-temperature transitions have been overlooked.

Acknowledgments

This paper was supported by the Slovenian Research Agency through Grant no. P1-0201 and by the COST action MP0802.

References

- [1] P. Brion and E. Westhof, "Hierarchy and dynamics of RNA folding," *Annual Review of Biophysics and Biomolecular Structure*, vol. 26, pp. 113–137, 1997.
- [2] D. B. Roth, J. P. Menetski, P. B. Nakajima, M. J. Bosma, and M. Gellert, "V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes," *Cell*, vol. 70, no. 6, pp. 983–991, 1992.
- [3] A. Bhasin, I. Y. Goryshin, and W. S. Reznikoff, "Hairpin formation in Tn5 transposition," *Journal of Biological Chemistry*, vol. 274, no. 52, pp. 37021–37029, 1999.
- [4] A. K. Kennedy, A. Guhathakurta, N. Kleckner, and D. B. Hanford, "Tn10 transposition via a DNA hairpin intermediate," *Cell*, vol. 95, no. 1, pp. 125–134, 1998.
- [5] O. C. Uhlenbeck, "Nucleic-acid structure—tetraloops and RNA folding," *Nature*, vol. 346, no. 6285, pp. 613–614, 1990.
- [6] C. Cheong, G. Varani, and I. Tinoco, "Solution structure of an unusually stable RNA heparin—GGAC(UUCG)GUCC," *Nature*, vol. 346, no. 6285, pp. 680–682, 1990.
- [7] H. A. Heus and A. Pardi, "Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops," *Science*, vol. 253, no. 5016, pp. 191–194, 1991.
- [8] D. E. Draper, "Themes in RNA-protein recognition," *Journal of Molecular Biology*, vol. 293, no. 2, pp. 255–270, 1999.
- [9] P. Svoboda and A. Di Cara, "Hairpin RNA: a secondary structure of primary importance," *Cellular and Molecular Life Sciences*, vol. 63, no. 7-8, pp. 901–918, 2006.
- [10] G. Bonnet, S. Tyagi, A. Libchaber, and F. R. Kramer, "Thermodynamic basis of the enhanced specificity of structured DNA probes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 11, pp. 6171–6176, 1999.
- [11] Y. Kim, D. Sohn, and W. Tan, "Molecular beacons in biomedical detection and clinical diagnosis," *International Journal of Clinical and Experimental Pathology*, vol. 1, pp. 105–116, 2008.
- [12] I. E. Scheffler, E. L. Elson, and R. L. Baldwin, "Helix formation by d(TA) oligomers. II. Analysis of the helix-coil transitions of linear and circular oligomers," *Journal of Molecular Biology*, vol. 48, no. 1, pp. 145–171, 1970.
- [13] J. Gralla and D. M. Crothers, "Free energy of imperfect nucleic acid helices. II. Small Hairpin Loops," *Journal of Molecular Biology*, vol. 73, no. 4, pp. 497–511, 1973.
- [14] C. A. Haasnoot, C. W. Hilbers, G. A. van der Marel et al., "On loop folding in nucleic acid hairpin-type structures," *Journal of Biomolecular Structure & Dynamics*, vol. 3, no. 5, pp. 843–857, 1986.
- [15] M. M. Senior, R. A. Jones, and K. J. Breslauer, "Influence of loop residues on the relative stabilities of DNA hairpin structures," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 85, no. 17, pp. 6242–6246, 1988.
- [16] M. Zuker, "On findings all suboptimal foldings of an RNA molecule," *Science*, vol. 244, no. 4900, pp. 48–52, 1989.
- [17] D. Rentzeperis, K. Alessi, and L. A. Marky, "Thermodynamics of DNA hairpins: contribution of loop size to hairpin stability and ethidium binding," *Nucleic Acids Research*, vol. 21, no. 11, pp. 2683–2689, 1993.
- [18] M. W. Freyer, R. Buscaglia, B. Nguyen, W. David Wilson, and E. A. Lewis, "Binding of netropsin and 4,6-diamidino-2-phenylindole to an A2T2 DNA hairpin: a comparison of biophysical techniques," *Analytical Biochemistry*, vol. 355, no. 2, pp. 259–266, 2006.
- [19] S. V. Kuznetsov, C. C. Ren, S. A. Woodson, and A. Ansari, "Loop dependence of the stability and dynamics of nucleic acid hairpins," *Nucleic Acids Research*, vol. 36, no. 4, pp. 1098–1112, 2008.
- [20] J. Lah, I. Drobnak, M. Dolinar, and G. Vesnaver, "What drives the binding of minor groove-directed ligands to DNA hairpins?" *Nucleic Acids Research*, vol. 36, no. 3, pp. 897–904, 2008.
- [21] J. Lah, N. Carl, I. Drobnak, B. Šumiga, and G. Vesnaver, "Competition of some minor groove binders for a single DNA binding site," *Acta Chimica Slovenica*, vol. 53, no. 3, pp. 284–291, 2006.
- [22] P. C. Bevilacqua and J. M. Blose, "Structures, kinetics, thermodynamics, and biological functions of RNA hairpins," *Annual Review of Physical Chemistry*, vol. 59, pp. 79–103, 2008.
- [23] D. Bikard, C. Loot, Z. Baharoglu, and D. Mazel, "Folded DNA in action: hairpin formation and biological functions in prokaryotes," *Microbiology and Molecular Biology Reviews*, vol. 74, no. 4, pp. 570–588, 2010.
- [24] L. A. Marky, D. Rentzeperis, and J. Ho, "Contribution of loops and nicks to the formation of DNA dumbbells: melting behavior and ligand binding," *Biochemistry*, vol. 32, no. 10, pp. 2564–2572, 1993.
- [25] I. Prislán, J. Lah, and G. Vesnaver, "Diverse polymorphism of G-quadruplexes as a kinetic phenomenon," *Journal of the American Chemical Society*, vol. 130, no. 43, pp. 14161–14169, 2008.
- [26] I. Prislán, J. Lah, M. Milanic et al., "Kinetically governed polymorphism of d(G4T4G3) quadruplexes in K⁺ solutions," *Nucleic Acids Research*, vol. 39, pp. 1933–1942, 2011.
- [27] A. Ansari, S. V. Kuznetsov, and Y. Shen, "Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 14, pp. 7771–7776, 2001.
- [28] Y. Shen, S. V. Kuznetsov, and A. Ansari, "Loop dependence of the dynamics of DNA hairpins," *Journal of Physical Chemistry B*, vol. 105, no. 48, pp. 12202–12211, 2001.
- [29] H. Ma, D. J. Proctor, E. Kierzek, R. Kierzek, P. C. Bevilacqua, and M. Gruebele, "Exploring the energy landscape of a small RNA hairpin," *Journal of the American Chemical Society*, vol. 128, no. 5, pp. 1523–1530, 2006.
- [30] C. R. Cantor, M. M. Warshaw, and H. Shapiro, "Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides," *Biopolymers*, vol. 9, no. 9, pp. 1059–1077, 1970.
- [31] W. H. Press, B. P. Flannery, S. A. Teukolsky et al., *Numerical Recipes*, Cambridge University Press, Oxford, UK, 2007.
- [32] I. Drobnak, M. Seručnik, J. Lah, and G. Vesnaver, "Stability of a short DNA duplex as a function of temperature: the effect of ΔC_p and added salt concentration," *Acta Chimica Slovenica*, vol. 54, no. 3, pp. 445–451, 2007.

- [33] M. T. Record, C. F. Anderson, and T. M. Lohman, "Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity," *Quarterly Reviews of Biophysics*, vol. 11, no. 2, pp. 103–178, 1978.
- [34] C. W. Hilbers, C. A. G. Haasnoot, S. H. de Bruin, J. J. M. Joordens, G. A. Van Der Marel, and J. H. Van Boom, "Hairpin formation in synthetic oligonucleotides," *Biochimie*, vol. 67, no. 7-8, pp. 685–695, 1985.
- [35] P. M. Vallone and A. S. Benight, "Melting studies of short DNA hairpins containing the universal base 5-nitroindole," *Nucleic Acids Research*, vol. 27, pp. 3589–3596, 1999.

Research Article

A Structurally Variable Hinged Tetrahedron Framework from DNA Origami

**David M. Smith,^{1,2} Verena Schüller,¹ Carsten Forthmann,^{3,4} Robert Schreiber,¹
Philip Tinnefeld,^{2,4} and Tim Liedl^{1,2}**

¹Physik Weicher Materie und Biophysik, Ludwig Maximilian University, 80539 Munich, Germany

²Center for NanoScience, Ludwig Maximilian University, 80799 Munich, Germany

³Angewandte Physik-Biophysik, Ludwig Maximilian University, 80539 Munich, Germany

⁴Physikalische und Theoretische Chemie-NanoBioScience, Braunschweig University of Technology, 38106 Braunschweig, Germany

Correspondence should be addressed to David M. Smith, david.smith@physik.uni-muenchen.de

Received 15 May 2011; Accepted 27 June 2011

Academic Editor: F. C. Simmel

Copyright © 2011 David M. Smith et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nanometer-sized polyhedral wire-frame objects hold a wide range of potential applications both as structural scaffolds as well as a basis for synthetic nanocontainers. The utilization of DNA as basic building blocks for such structures allows the exploitation of bottom-up self-assembly in order to achieve molecular programmability through the pairing of complementary bases. In this work, we report on a hollow but rigid tetrahedron framework of 75 nm strut length constructed with the DNA origami method. Flexible hinges at each of their four joints provide a means for structural variability of the object. Through the opening of gaps along the struts, four variants can be created as confirmed by both gel electrophoresis and direct imaging techniques. The intrinsic site addressability provided by this technique allows the unique targeted attachment of dye and/or linker molecules at any point on the structure's surface, which we prove through the superresolution fluorescence microscopy technique DNA PAINT.

1. Introduction

The design and self-assembly of DNA strands into precisely defined objects on the nanometer scale has emerged as a promising technique in the field of nanotechnology. Stemming from the initial idea of generating periodic lattices from DNA [1], the concurrent exploitation of (i) complimentary base pairing between short strands, (ii) branch-like Holliday junctions, and (iii) the inherent helical twist of double-stranded DNA complexes has allowed for the assembly of small, identical motifs which constitute the repeating unit cells of periodic two-dimensional sheets or three-dimensional crystal structures extending nearly to the millimeter scale [2–5]. The development of techniques to build rigid, three-dimensional DNA-based structures is, however, an important aspect to the future utilization of this methodology in nanofabrication [6–11]. While many attempts to construct simple three-dimensional polyhedra have been fraught with problems of instability, unwanted by-products,

low-yield, or overly complex synthesis strategies, the recent utilization of hierarchical assembly schemes [12] and the DNA origami technique [13, 14] has provided a path towards the relatively simple generation of uniform populations.

The DNA origami technique is based on the use of a long, usually circular “scaffold” strand, which is folded and clamped into a desired shape by hybridization with hundreds of shorter “staple” oligonucleotides [13]. In contrast to earlier schemes for generating nanostructures from synthesized oligonucleotides, the utilization of a viral scaffold, which is typically the 7 kb circular single-stranded M13mp18 bacteriophage genome or alternatively PCR-templated products [15, 16], allows for the construction of far larger structures often extending to several hundreds of nanometers [17]. The heterogeneous sequence of the underlying scaffold strand forms the basis for the unique strength of this method, by which each staple strand has the potential to act as a “handle” for the placement of accessory molecules at virtually any site in the structure with nanometer precision

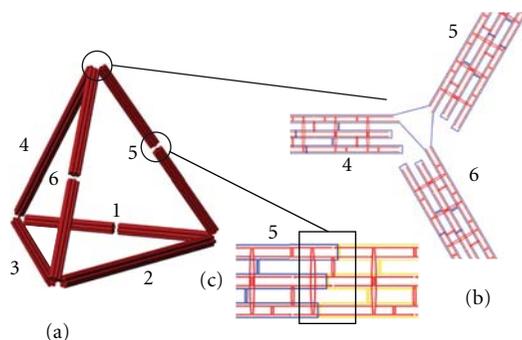


FIGURE 1: Schematic view of DNA origami tetrahedron. (a) In the three-dimensional representation, each double-helical section is represented by a single red cylinder. The six struts are labeled 1 through 6, and each consist of a bundle of six parallel connected double helices. (b) A two-dimensional diagram of the hinged vertex formed by struts 4, 5, and 6 shows the path of the m13MP18-based scaffold (blue) forming the single-stranded connections between bundles and the pattern of staple oligonucleotides (red) around the hinge. (c) Local structure of the scaffold path within strut 5 shows three closely aligned scaffold crossovers stabilized by short staple sections forming the weakened gap. The scaffold path on opposite sides of the gap is marked in either yellow or blue to emphasize the discontinuity.

[18, 19]. Through standard purification techniques, relatively uniform populations of folded structures of the desired conformation can be isolated. This method is suited not only for compact, structurally rigid objects, but also for those which incorporate single-stranded sections amongst rigidly extended helical bundles to generate objects displaying flexible hinges and inherent tensile strain [20].

Hollow, load-bearing frameworks represent one class of structures which holds a broad relevance in applied nanoscience as potential container systems. The use of programmed DNA self-assembly offers a highly suitable alternative pathway in their construction, as has been demonstrated in several recent studies. Two in particular have succeeded in generating highly uniform populations of rigid nanostructures. Small tetrahedral structures measuring less than 10 nanometers per strut have been rapidly assembled from four distinct oligonucleotides and have shown a significant mechanical stability, withstanding compression forces in excess of 100 pN before exhibiting buckling [10]. It was later demonstrated that a hierarchical self-assembly concept utilizing simple repeating motifs from a relatively small number of oligonucleotides [12] or DNA origami-based subunits [14] can be used to generate larger polyhedra suitable as containers or for the orientation of functional units in space with nanometer precision. While the above strategies based upon hybridization of shorter oligonucleotides produce high yields of correctly assembled products, the former is significantly size-limited based on its dependence on a simple annealing scheme, while the ability to selectively attach molecules at distinct points on the polyhedra constructed by hierarchical means suffers due to the redundancy of strands or units within the repeated motifs. The use of subunits constructed from DNA origami [14] can solve this issue of redundancy via a simple internal shifting of the scaffold strand; however, it does require the additional and often intricate control over correct three-dimensional orientation and self-assembly of the multiple segments. Consequently, the fabrication of enclosed hollow frames from a single DNA scaffold via the origami method,

as previously demonstrated in a spherical architecture [21], offers the potential for the straightforward generation of a wide variety of fully and uniquely addressable structures.

As shown in the present study, utilization of the DNA origami method to construct a rigid, enclosed tetrahedral framework, illustrated in Figure 1, provides the ability to generate structures reaching a size of 75 nanometers along each edge while retaining the full addressability afforded by the heterogeneous sequence of the scaffold strand. Single stranded “hinges” at each of the four vertices of the tetrahedron allow for a structural variability. By selective staple oligonucleotide omission, four unique two-dimensional structures varying in length and width between 75 nanometers and 300 nanometers are generated. Even though the full three-dimensional structure displays mechanical rupture likely due to electrostatic-mediated surface binding and flow-induced shearing, proper annealing can nevertheless be confirmed through gel electrophoresis and examination of the assorted open frames. Additionally, binding of fluorescent molecules at precisely determined points on the structure emphasizes its potential to act as a three-dimensional ruler for emerging superresolution microscopy methods such as DNA-PAINT and Blink [22–24].

2. Materials and Methods

2.1. Preparation of DNA Structures. Reverse-phase cartridge-purified unmodified staple oligonucleotides were shipped dry from Bioneer Corporation (Alameda, Calif, USA) and resuspended to a concentration of 100 μ M in H₂O. HPLC-purified dye- and biotin-modified oligonucleotides for PAINT analysis were purchased from IBA GmbH (Göttingen, Germany) at a concentration of 50 μ M in H₂O. The ATTO655-labeled oligonucleotide contained the sequence 5'-GGT GAA GA_{ATTO655}-3' and the biotin-modified strand 5'-GGT AGT AAT AGG AGA ATG_{bt}-3'. The 8634-nucleotide-long M13mp18-phage-based scaffold strand was prepared and isolated as previously described [14].

As a general annealing condition, 10 nM of scaffold strand and 100 nM of each staple were mixed in TE buffer (10 mM Tris-HCl + 1 mM EDTA, pH 8.0 at 20°C) containing 14 mM MgCl₂, heated to 80°C for 5 minutes, and quickly cooled to 60°C over 20 minutes, before being slowly cooled to 24°C in steps of 0.5°C over approximately 40 hours. When applicable for dye or biotin attachment, staple oligonucleotides with docking handle extensions replaced unextended staples in the solution and biotin- or dye-modified oligonucleotides were included at 500 nM per staple during annealing. For PAINT analysis, a total of six biotin attachment sites (two per vertex) and three ATTO655 sites (one per strut) were used to modify struts I, II, and III, as shown in Figure 4(a) and in the Supplementary Materials in Figure S1.

Gel electrophoresis was performed with an agarose concentration of either 2.0% (for band analysis) or 0.7% (for structure purification). Agarose was first heated to boiling temperature in 45 mM Tris borate + 1 mM EDTA (pH 8.2 at 20°C) then cooled to 60°C in an ice bath, at which point MgCl₂ was added to a final concentration of 11 mM, before being cooled to room temperature in a gel cask. For sample preparation, 5 μL of DNA origami structures were mixed with 3 μL of 6x agarose gel loading buffer (30% glycerol weight-to-volume in water, 0.025% xylene cyanol, 0.025% bromophenol blue) and brought to a total volume of 18 μL in H₂O, for roughly $5 \cdot 10^{-14}$ moles of structures per gel pocket. Similarly, bare scaffold strands were mixed with loading buffer and H₂O to give a total amount of $1 \cdot 10^{-13}$ moles per sample. The different gel pockets were filled with the origami and scaffold samples along with a 1 kb DNA ladder and run for approximately 3 hours at constant 70 V over a cathode-anode distance of 22 cm. The entire apparatus was cooled in an ice water bath while running to avoid unwanted structure denaturation. Afterwards, the gel was stained with an ethidium bromide solution at 0.5 μg/mL and imaged under 302 nm UV excitation. Following electrophoresis, the primary band was excised from the gel and centrifuged at 5000 ×g for 7 minutes in spin columns (Freeze'n Squeeze Spin Columns, Biorad) to isolate structures. Purified origami samples were usually estimated to contain approximately 1 nM of annealed structures.

2.2. Visualization of Structures (TEM and AFM)

TEM Imaging. Formvar-supported carbon coated TEM grids were purchased from Plano GmbH (Wetzlar, Germany). Grids were first hydrophilicized in a Diener Electronic Femto plasma cleaner. 2 μL of purified origami samples was added to the grids, and structures were allowed to bind for 60 seconds. Excess sample was then quickly removed from the grids by absorption with filter paper. Grids were quickly washed with 1% uranyl acetate in H₂O, then stained for 15 seconds with the same and allowed to dry completely. The samples were then imaged with a JEM-1011 transmission electron microscope (JEOL) operated at 100 kV.

AFM Imaging. For viewing with atomic force microscopy, 5 μL of purified origami sample placed onto a freshly cleaved

mica surface (Plano GmbH, Wetzlar, Germany) which had been attached by hot glue to a 15 mm metal specimen disc (Ted Pella, Inc., Redding, Calif, USA). Structures were allowed to bind for 60 seconds before being washed twice with 30 μL of TE buffer solution containing 12.5 mM MgCl₂ to remove unbound origami objects and other debris. Samples were imaged in Tapping Mode in the previously mentioned TE/Mg⁺⁺ buffer conditions using a NanoScope III Multimode AFM from Digital Instruments (Veeco Instruments Ltd., Plainview, Tex, USA) with a silicon-nitride tip with a spring constant of $k = 0.24$ N/m (Veeco).

2.3. Shape Analysis by DNA PAINT

Sample Preparation. Chambered Cover Glass Slides (LabTek, NUNC, Langensfeld, Germany) were prepared for Total Internal Reflection (TIR) Microscopy in the following manner: glass surfaces were first cleaned with 0.1 M HF for 30 seconds and washed with PBS. They were then passivated with a mixture of 5 mg/mL bovine serum albumin (BSA) and 1 mg/mL biotin-conjugated BSA for 16 hours. Following a second washing with PBS to remove unbound proteins, the surface was incubated for 15 minutes with 0.01 mg/mL streptavidin (IBA GmbH) then again washed with PBS to remove any excess, unbound streptavidin. 50 μL of the sample solution which contained 90% of a 1 M MgCl₂ buffer and 10% purified origami solution (for a final structure concentration of approximately 0.1 nM) was added to the chamber and allowed to bind to the surface. Binding was facilitated by biotin linkers on the origami structures, included at each vertex of the plane formed by struts I, II, and III. Chambers were again washed with PBS to remove unbound structures, and a 50 nM solution of ATTO655-labeled oligonucleotides (PAINT-DNA) in 500 mM NaCl was added.

Microscopy Setup. An inverted Olympus IX-71 in objective-type TIRF configuration outfitted with a UPlanSApo objective (100x, NA 1.40, oil immersion) was used for monitoring. Dyes were excited by a single-mode diode laser (XTL, Toptica Photonics, Grafelfing, Germany) operated at 100 mW with a wavelength of $\lambda = 650$ nm. Fluorescence was imaged on an EMCCD camera (Ixon DU-897, Andor Technologies, Belfast, Northern Ireland). Image size was 128 × 128 pixels, with each pixel representing a length of approximately 90 nm. Each acquired sequence of images contained approximately 2000 frames taken at a rate of 20 Hz.

3. Results and Discussion

3.1. Tetrahedron Design. The tetrahedral design was implemented using the caDNano [25] software package, which was developed to facilitate the layout of both three-dimensional and flat DNA origami structures. A circular scaffold strand of 8634 nucleotides was used to allow for a maximal size, which was folded into its final configuration by a total of 211 staple strands. Each of the six struts in the tetrahedron is based on a bundle of six parallel double helices [14, 26], 227 paired bases in length,

with each double helix represented by a cylinder in the three-dimensional schematic (Figure 1(a)). The bundles are connected at each of the four hinged vertices to the two neighboring bundles by a four-base single-stranded section of the circular scaffold extending between their termini, as indicated in the inset in Figure 1(b). At each vertex, adjacent double helices within a single six-helix bundle were selected as connecting points to each of the two neighboring struts so as to minimize possible strains resulting from sterically induced stretching of the roughly 4 nm single-stranded connecting section. Additionally, connection points on opposite ends of a single strut extend from the same pair of double-stranded helical sections, which minimizes any possible unfavorable twist-strain within the bundle [21, 27]. We have adopted the nomenclature as indicated in Figure 1(a) whereby each strut is labeled 1 through 6. Due to circular continuity of the scaffold strand and the periodicity of scaffold crossovers between neighboring helices within a single six-helix bundle, struts 1, 5, and 6 contain points where three scaffold crossovers are contained within a seven-base section along the double helix. As highlighted in Figure 1(c) for strut 5, the three nearly aligned gaps within the strut are stabilized only by staple oligonucleotides hybridized to scaffold sections on opposing sides of each opening. Lacking a direct connection from the internal scaffold path running throughout the structure, these junctions, referred to as 1, 5, and 6 in coordination with the previous nomenclature, must be viewed as weak points susceptible to rupture; however, they provide the overall tetrahedron with its structural variability as will be discussed in later sections. A full schematic of both scaffold and staple arrangement within the overall structure can be seen in the Supplementary Materials in Figure S1.

Specific binding of tetrahedra to surfaces was facilitated by the addition of an 18-base sequence on the 3' end of six staples, each terminating at the end of a bundle lying on the plane formed by struts 1, 2, and 3, to which a complementary oligonucleotide containing a biotin molecule conjugated to its 3' end could be hybridized. This allowed stable attachment of structures to surfaces coated with biotin-binding proteins such as streptavidin. The positioning of the biotin on the 3' end of the modified strand directly abutting the structure imposes the closest possible binding between the object and surface. In a similar manner, one staple positioned along the length of each of struts 1, 2, and 3 with their 3' terminus on the outer surface of the bundle was selected as a handle for complimentary oligonucleotides with an ATTO655 dye conjugated to the 3' end of the strand in DNA PAINT experiments. In both cases, any combination of biotin or dye sites could be included within the structure by the replacement of normal staple strands with those extended by the complimentary "handle" or docking sequence.

3.2. Full Tetrahedron Structural Properties. The structures were annealed in a one-pot reaction in TE buffer in the presence of 14 mM MgCl₂. As analyzed by agarose gel electrophoresis after annealing, well-defined bands were observed indicating a population of uniformly annealed structures, along with signs of additional slower migrating

aggregate products, as can be seen in the Supplementary Materials in Figure S2. Little change in band structure was observed by varying the annealing time from 18 to 170 hours, although a lengthening of the single-stranded scaffold sections between struts caused a higher migration speed and a less-defined band structure (see Figure S2 in the Supplementary Materials). From analysis of relative band intensity in gels, the yield of primary annealed product was approximately 20%. The majority of scaffold-containing material remained in the gel pockets, where presumably aggregated structures are unable to effectively enter or migrate through the gel matrix.

Annealed objects were negatively stained with uranyl acetate and analyzed via transmission electron microscopy (TEM) for structural characteristics. Despite the defined structure of gel bands, TEM observations yielded an absence of intact tetrahedra, with a large portion of the structures displaying apparent breaks along one or more of the individual struts. Two such examples are shown in the left panel of Figure 2; the individual six-helix bundle struts appear to be well formed and connectivity within most vertices is maintained; however, the structures display a deformed or flattened configuration appearing to arise from a small number of ruptures along the length of one or two edges. Measured lengths of individual struts varied from 70 to 90 nm, close to the estimated value of 75 nm based on a 0.34 nm axial pitch between adjacent bases in the double helix.

3.3. Piecewise Verification of Proper Annealing. The pervasive existence of malformed structures under TEM observation points to one of two possible causes; either a systematic error in annealing occurs due to unexpected internal strains or kinetic traps, or properly annealed tetrahedra are denatured as a consequence of mechanical stresses endured in the process of purifying, fixing, and staining the objects. The lack of a clear predominance of any single "broken" tetrahedron configuration as viewed under TEM despite well-defined band structure in electrophoresis strongly indicates that the second consideration is the more likely explanation for the aforementioned observations. As previously emphasized, the three gap areas within each of struts 1, 5, and 6 formed by the close alignment of scaffold crossovers must be considered as weakened points along the bundles. A closer examination of the underlying staple connections across the gaps as pictured in Figure 1(c) for strut 5 shows that the hybridized sections stabilizing gaps along a single helical axis are in several cases as short as two nucleotides in overlap before the termination of the particular staple strand. While some degree of stabilization stemming from stacking interactions could occur for some abutting helices [28, 29], these points are nevertheless more susceptible to rupture under strong shear or compression forces than other points along continuous sections of a strut or at vertices. This can be understood through a simple comparison of the two different bond types responsible for stabilizing scaffolded DNA structures: hydrogen bonding of complimentary base pairing and the covalently connected polymer backbone comprising the main DNA scaffold through the structure.

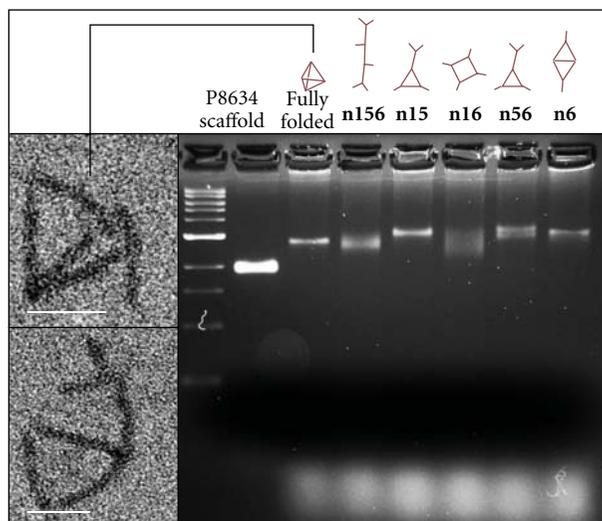


FIGURE 2: (Left) Transmission electron microscopy (TEM) imaging of tetrahedron structure containing all staple nucleotides. Typically, tetrahedrons displayed one or more ruptured struts, generating the widely observed flattened configuration. Scale bars: 60 nm. (Right) Agarose gel electrophoresis separation of the different configurations formed by targeted gap staple omission in struts 1, 5, and 6. Left to right: lanes 1-2 contain a 1 kb ladder and the circular p8634 scaffold, respectively. Lanes 3-8 contain the full tetrahedron, and the open configurations **n156**, **n15**, **n16**, **n56**, and **n6**, as indicated by the sketched representations above each lane.

In the case of the former, it has been demonstrated that forces on the order of 50 piconewtons are sufficient to induce shear-oriented rupture of hybridized DNA strands [30, 31]. Conversely, earlier work has shown that covalent bonds can generally withstand forces up to the nanonewton scale before experiencing rupture [32]. Concerning the overall stability of the tetrahedron presented in this work, the relatively weak hybridized staple oligonucleotides spanning the gap regions in struts 1, 5, and 6 would be far more susceptible to rupture under external stress than any of the other areas whose stabilities are in addition to base pairing supported by the continuous, covalently-linked scaffold backbone.

Standard techniques for examining DNA-based nanostructures such as TEM or AFM depend on a strong electrostatic binding of the negatively charged structures to layers such as mica or the carbon coating on grids used for TEM, and staining protocols include washing and drying steps with rapid, high-shear addition and removal of fluids. While this has proven to cause only nominal damage to more compactly rigid structures, the weakened struts are particularly susceptible to rupture due to electrostatic-driven compression or deforming forces containing a shear component and are likely candidates for the observed open configurations compressed on the surface.

With force-induced rupturing of weakened points formed by gaps within the struts identified as a likely candidate for the observed deformation of tetrahedra, we sought to investigate the various configurations resulting from the targeted opening of those gaps in different combinations. Due to their design and the local stapling paths around these junctions, these two-dimensional conformations could be generated by the omission of staple groups spanning each of the three gaps during annealing. For this purpose, an open configuration formed by a particular combination of

omitted gap staples is referred to by the number assigned to the gap or gaps presumably left open. By this nomenclature, the configuration generated by the omission of all gap staples is referred to as **n156**; that resulting from the opening of the gap along strut 6 is **n6**, and so on. While there are a total of seven possible combinations of gap staple groups which can be left out, structural redundancies mean that a total of only four possible flat configurations exist. These are roughly sketched above their corresponding lane in the gel shown in the right panel of Figure 2, with **n6** being the representative structure for the three identical configurations containing a single omission.

The various flat structures were annealed and analyzed as the full tetrahedron. As can be seen in the gel in Figure 2, differences in migration and band structure amongst the variants were evident after a three-hour separation at 70 V over 22 cm in 2% agarose. Bands of the different flat two-dimensional assemblies displayed a predictably lower resolution than their fully annealed counterpart, presumably due to their increased degree of configurational flexibility around open, flexible joints. Two of these structures, the linearized **n156** form and the rectangular **n16**, have bands with a significant population of products which migrate faster than the full tetrahedron structure containing all staples. This can be explained by the lack of structurally rigid triangular substructures within the two frames, which can lead to an elongated conformation able to effectively reptate through the gel pores. Furthermore, the geometrically redundant **n15** and **n56** as well as the **n6** configurations all have primary populations which display a slower migration character visibly absent from the fully annealed version. In both cases, this is a strong indication that none of those are preferentially occurring in any detectable amount during annealing of the full tetrahedron and that weakened

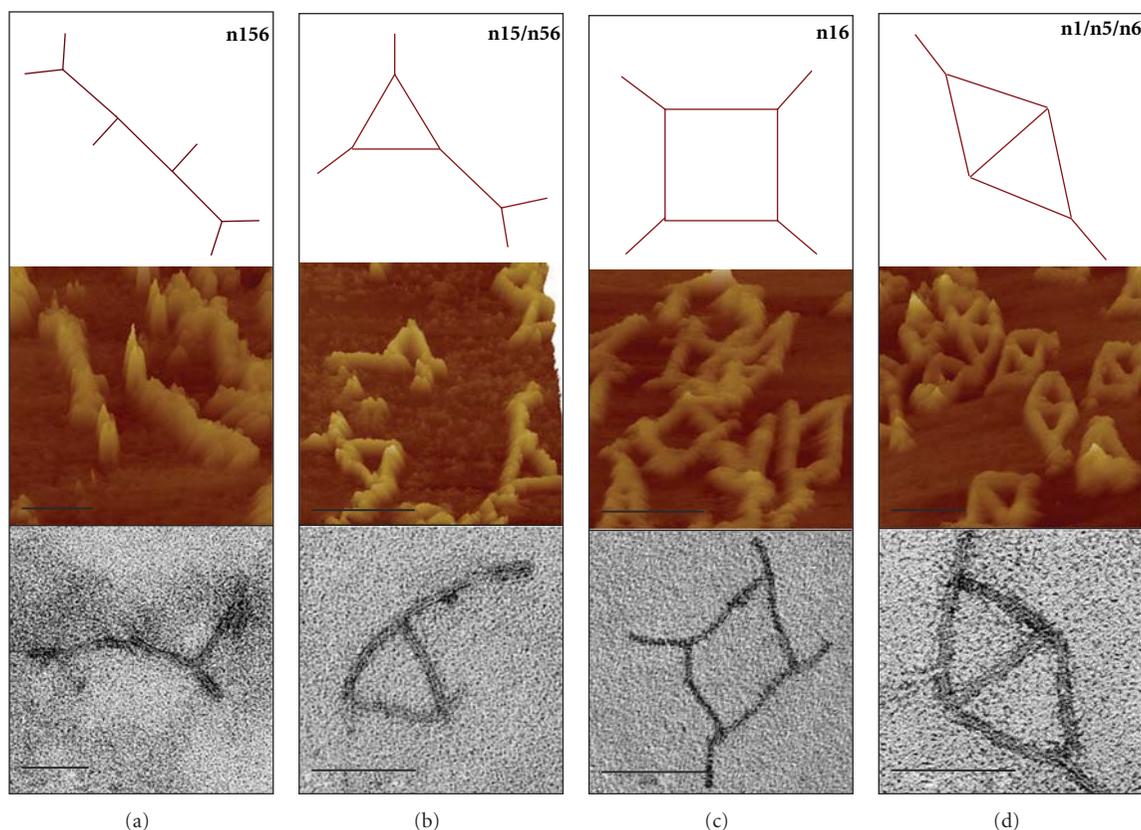


FIGURE 3: Schematic (upper row), atomic force microscopy (center), and TEM imaging (bottom) of flattened structures produced by targeted gap staple omission. (a) The chain-like **n156** configuration is formed when all gaps are left open. (b) Removal of staple groups connecting gaps in struts 1 and 5 or alternatively 5 and 6 form the geometrically redundant **n15** and **n56** projections. (c) Jointed square structures **n16** are formed by the opening of gaps in struts 1 and 6, displaying a degree of transverse flexibility due to possessing four flexible hinges. Omission of a single gap staple group as in (d) has a three-fold redundancy of structures with geometry indicated by the **n6** object shown. Scale bars: 150 nm (AFM) and 75 nm (TEM).

points in the struts are not ruptured within the gel. This absence of any primary product resulting from the variable open configurations which displays the same migration speed as the full tetrahedron as well as the narrow band structure of the closed tetrahedron indicate that it adopts its final, defined conformation unique from those resulting from some combination of ruptured struts. Approximate yields of primary annealed products were estimated through comparison of relative band intensities. For all opened variants, this was found to be approximately 20–25%, in a similar range as the fully annealed tetrahedron, although a lesser degree of band clarity does make any precise determination in this manner unreliable. Highest yields were found with the more flexible and elongated **n156** and **n16** variants, likely due to the greater amount of overall material able to enter the gel. Conversely, the most rigid of the variants containing only one gap opening (represented as **n6** in Figure 2) showed the smallest yields of all products, usually under 20%; however, with the greatest clarity in band structure and highest amount of aggregation near the gel pockets.

Prominent bands of each flattened conformation were excised from the gel, and structures were isolated via spin

filtration before being imaged by TEM and AFM, as shown in Figure 3. In each case, most structures were found in the configuration suggested by the pattern of openings along the struts. Vertices forming connections between three partial or complete struts appeared to be well formed, and the length of spanning segments was in the range of the ideally expected 75 nm. In some cases, partial bundles resulting from an intentionally opened gap were not clearly distinguishable, likely due to the small size of these parts and possible overlapping with other struts. Populations containing fewer deformed or aggregated objects were generally seen under observation via AFM, likely resulting from the lack of potentially harsh drying steps which are typical in uranyl acetate negative staining used for TEM observations.

Taken in concert, the different two-dimensional frames demonstrate, on an individual basis, the ability for each vertex and strut, whether continuous or containing a gap region, to correctly anneal as a part of a particular flat variant of the tetrahedron. This does not fully eliminate the possibility that unintended internal stresses resulting from closure of the full structure lead to some degree of deformation. However, the single primary band migrating with the speed different to those of all open conformations

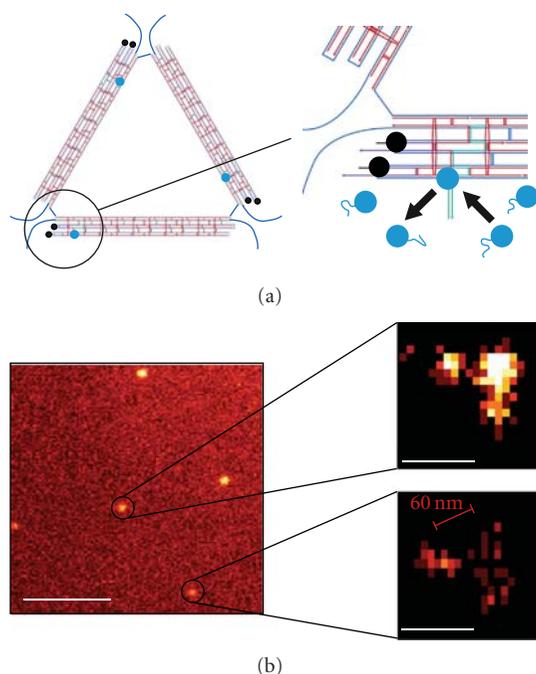


FIGURE 4: DNA PAINT measurements on the triangular base of the tetrahedron. (a) Design schematic of the base showing docking sites for complimentary biotin- and ATTO655-conjugated oligonucleotides. The scaffold path is shown in blue, and staples lacking extended docking sequences are in red. Only basal struts 1, 2, and 3 are shown in the full schematic (left), and local detail at docking sites on strut 1 is given (right). The six extended staples containing docking sites for complimentary biotin-nucleotides are shown in purple, with biotin and the affixed nucleotide strand in black. Likewise, the three docking staples for ATTO655 dye are shown in green, while the ATTO655 dye and attached 8-base nucleotide which transiently hybridize to the structures are in light blue. (b) Resolution of triangular structure via DNA PAINT. (Left) Sample frame and (right) superresolution 2D reconstruction of local binding events. In the color scheme on the reconstructions, lighter colors correspond to a higher number of binding events, showing the underlying triangular form matching the schematic arrangement of dyes. For a comparison, the approximate expected distance between docking sites of 60 nm is indicated in the image. Scale bars: 4 μm (left) and 100 nm (right).

is strong evidence for the successful annealing of the full tetrahedron. From this, we can conclude that rupturing occurs during fixing and staining procedures onto carbon-coated TEM grids due to strong surface interactions from surface adsorption or transverse shear forces incurred in washing and drying steps.

3.4. Superresolution Analysis via DNA PAINT. Single-layer DNA origami structures have recently been shown to function as ideal substrates for superresolution microscopy techniques such as DNA PAINT [24] (points accumulation for imaging in nanoscale topography) or Blink microscopy [19] due to the ability to place fluorescent dyes at prescribed locations on the structure's surface with nanometer precision. The localization of fluorescent molecules beyond the

normal diffraction limit of visible light is predicated by the principle that only a single molecule within the detection area is active, that is, emitting fluorescent light, at any given time [22, 23]. Switching between ON and OFF states of fluorescent molecules is generally controlled by either chemical means or by utilizing intrinsic dark states [33]. In contrast, for systems involving interactions on surface-bound DNA-based nanostructures due to hybridization-based binding and unbinding between dye-oligonucleotide-conjugate detector strands and complimentary docking strands on the structures, switching occurs as a natural consequence of two factors: (i) freely diffusing strands within solution cause minimal background due to TIR excitation of only dyes near to the surface by the evanescent wave and (ii) a 70% increase of fluorescence is attained upon hybridization due to a significant reduction of intrastrand guanine quenching [24]. This results in the relatively straightforward detection of single hybridization events occurring between single, labeled detector oligonucleotides and complimentary docking strands extending outward from the surface of the origami object. Furthermore, control over rates of binding/unbinding events can be achieved by altering the length of the hybridized segment and concentration of the detector strand within the surrounding solution, respectively [24].

Binding of tetrahedra to a streptavidin-coated surface was accomplished by the inclusion of six handle sites for complimentary biotin-oligonucleotide hybridization, two each extending from the ends of struts 1, 2, and 3, as indicated in Figure 4(a) and in Figure S1 of the Supplementary materials. Passivation of the glass surface with a protein layer containing BSA and streptavidin has the additional advantage of shielding the negatively charged structures from the strong electrostatic binding which proves destructive to their three-dimensional morphology as seen with the carbon or mica surfaces described in the previous sections. In addition, one staple with an extended handle sequence extending from the 3' end for ATTO655-oligonucleotide docking was incorporated into each of struts 1, 2, and 3, as outlined in detail for strut 1 in Figure 4(a) and shown for all struts in the Supplementary materials in Figure S1. The positioning of each dye site along its respective strut is shown in Figure 4(a), which together form a triangular arrangement with approximately 60 nm dimension. An 8-base docking length for the detector oligonucleotide was selected to give sufficiently long binding times for accurate position determination, and a high detector concentration of 50 nM was chosen in order to have high enough association rates [24].

Binding/unbinding events were recorded over 100 s with a capture rate of 20 Hz during TIR excitation from a single-mode diode laser with wavelength of 650 nm. Recorded movies were analyzed in a custom software package programmed in LabVIEW [34]. Spots where binding events in each frame occurred were fitted with a 2D Gaussian, and the peak coordinates of each spot were entered into a 2D histogram with a binning of 10 nm. According to the color scheme used, brighter spots represent a greater frequency of binding events at that particular position.

The left frame in Figure 4(b) shows a single snapshot of the $11.5\ \mu\text{m} \times 11.5\ \mu\text{m}$ area monitored, and the two frames on the right show the superresolved 2D histograms of two different localized areas of binding/unbinding events. In contrast to raw data, the superresolved histograms of local spots indicate a localization of three peak event occurrences. The three apparent binding locations on the origami object display the correct distance scale and geometrical arrangement for spacing suggested by docking site placement along the base of the tetrahedron. Through the analysis of six docking site pairs, we found an average distance between localized histogram peaks of (68 ± 12) nm. For a comparison, the approximate expected distance of 60 nm is indicated in the image. Our data generally demonstrates the ability for DNA PAINT to resolve subdiffraction limit structures in two dimensions as well as the potential of such a DNA origami framework for superresolution techniques.

4. Conclusion

The construction of wire-frame, cage-like structures of nanometer size is an endeavor which not only holds great potential for studies requiring mechanically stable nanoscale spacers and scaffoldings but also could prove crucial for the development of containers suitable for targeted drug delivery. In this work, we have presented a DNA origami-based design for the simplest architecture of this kind; a four-faced tetrahedron consisting of six flexibly jointed struts. Analysis of annealed objects via gel electrophoresis strongly indicates the presence of a population of uniformly annealed tetrahedra comprising approximately 20% of the total number of scaffolded objects, although structures appear to suffer ruptures at select weakened points along the struts likely due to forces occurring during electrostatic adsorption to surfaces and staining procedures. Intentional, targeted opening of these weak “gaps” in the struts via selected staple oligonucleotide omission led to the formation of a set of four different open configurations, again at a yield of approximately 20%, which were analyzed by gel electrophoresis and visualized via atomic force microscopy and transmission electron microscopy. We found populations of structures that displayed the expected open morphologies when bound to surfaces, which demonstrated proper assembly of the individual vertices and full struts, further supporting the suggestion of rupture of full three-dimensional frameworks from external forces. By exploiting the nanometer-precise positioning of dyes on the tetrahedra, a key feature provided by the DNA origami technique, its triangular base was used as a stage for visualization in two dimensions with the superresolution DNA PAINT microscopy method. Analysis confirmed the triangular arrangement of attached dyes and similar spacing as predicted by the structural schematic. As a consequence, we expect such structures to be suitable test objects for 3D superresolution microscopy.

While the observed breakage of the tetrahedra at weakened points along the struts is proposed to arise from external stresses during fixation and staining, this apparent

structural instability under stress is nevertheless a concern for future implementation of such objects as stable load-bearing scaffolds or cargo-bearing containers. This can potentially be resolved by longer overlaps between the connecting staples spanning the individual helices or by the incorporation of a linear scaffold strand into future designs to reduce the weakened gap regions along the struts imposed by the necessity of maintaining circular scaffold continuity. Enzymatic ligation of abutting 3' and 5' ends of staples within origami structures has been posited as a potential means to increase the overall stability of single-layer architectures [35] and could serve to further strengthen three-dimensional frameworks as presented in this work. Additionally, recently developed techniques for metallization of DNA nanostructures could be applied as a potential means to mechanically stabilize and electrically functionalize the tetrahedron and its two-dimensional variants [36–38].

Acknowledgments

This work was financially supported by funding from NIM, CeNS, and DFG (LI1743/2-1, TI329/5-1). The authors would like to additionally acknowledge Dr. Susanne Kempter for assistance with transmission electron microscopy imaging and both Philipp Nickels and Dr. Ralf Jungmann for their help with atomic force microscopy visualization.

References

- [1] N. C. Seeman, “Nucleic acid junctions and lattices,” *Journal of Theoretical Biology*, vol. 99, no. 2, pp. 237–247, 1982.
- [2] E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman, “Design and self-assembly of two-dimensional DNA crystals,” *Nature*, vol. 394, no. 6693, pp. 539–544, 1998.
- [3] J. Malo, J. C. Mitchell, C. Vénien-Bryan et al., “Engineering a 2D protein-DNA crystal,” *Angewandte Chemie—International Edition*, vol. 44, no. 20, pp. 3057–3061, 2005.
- [4] Y. He, Y. Chen, H. Liu, A. E. Ribbe, and C. Mao, “Self-assembly of hexagonal DNA two-dimensional (2D) arrays,” *Journal of the American Chemical Society*, vol. 127, no. 35, pp. 12202–12203, 2005.
- [5] J. Zheng, J. J. Birktoft, Y. Chen et al., “From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal,” *Nature*, vol. 461, no. 7260, pp. 74–77, 2009.
- [6] J. Chen and N. C. Seeman, “Synthesis from DNA of a molecule with the connectivity of a cube,” *Nature*, vol. 350, no. 6319, pp. 631–633, 1991.
- [7] Y. Zhang and N. C. Seeman, “Construction of a DNA-truncated octahedron,” *Journal of the American Chemical Society*, vol. 116, no. 5, pp. 1661–1669, 1994.
- [8] W. M. Shih, J. D. Quispe, and G. F. Joyce, “A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron,” *Nature*, vol. 427, no. 6975, pp. 618–621, 2004.
- [9] J. Zimmermann, M. P. J. Cebulla, S. Mönninghoff, and G. von Kiedrowski, “Self-assembly of a DNA dodecahedron from 20 trisigonucleotides with C_{3h} linkers,” *Angewandte Chemie—International Edition*, vol. 47, no. 19, pp. 3626–3630, 2008.
- [10] R. P. Goodman, I. A. T. Schaap, C. F. Tardin et al., “Chemistry: rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication,” *Science*, vol. 310, no. 5754, pp. 1661–1665, 2005.

- [11] Z. Li, B. Wei, J. Nangreave et al., "A replicable tetrahedral nanostructure self-assembled from a single DNA strand," *Journal of the American Chemical Society*, vol. 131, no. 36, pp. 13093–13098, 2009.
- [12] Y. He, T. Ye, M. Su et al., "Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra," *Nature*, vol. 452, no. 7184, pp. 198–201, 2008.
- [13] P. W. K. Rothmund, "Folding DNA to create nanoscale shapes and patterns," *Nature*, vol. 440, no. 7082, pp. 297–302, 2006.
- [14] S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, and W. M. Shih, "Self-assembly of DNA into nanoscale three-dimensional shapes," *Nature*, vol. 459, no. 7245, pp. 414–418, 2009.
- [15] B. Högberg, T. Liedl, and W. M. Shih, "Folding DNA origami from a double-stranded source of scaffold," *Journal of the American Chemical Society*, vol. 131, no. 26, pp. 9154–9155, 2009.
- [16] E. Pound, J. R. Ashton, H. A. Becerril, and A. T. Woolley, "Polymerase chain reaction based scaffold preparation for the production of thin, branched DNA origami nanostructures of arbitrary sizes," *Nano Letters*, vol. 9, no. 12, pp. 4302–4305, 2009.
- [17] S. M. Douglas, J. J. Chou, and W. M. Shih, "DNA-nanotube-induced alignment of membrane proteins for NMR structure determination," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 16, pp. 6644–6648, 2007.
- [18] I. H. Stein, V. Schüller, P. Böhm, P. Tinnefeld, and T. Liedl, "Single-molecule FRET ruler based on rigid DNA origami blocks," *ChemPhysChem*, vol. 12, no. 3, pp. 689–695, 2011.
- [19] C. Steinhauer, R. Jungmann, T. L. Sobey, F. C. Simmel, and P. Tinnefeld, "DNA origami as a nanoscopic ruler for super-resolution microscopy," *Angewandte Chemie—International Edition*, vol. 48, no. 47, pp. 8870–8873, 2009.
- [20] T. Liedl, B. Högberg, J. Tytell, D. E. Ingber, and W. M. Shih, "Self-assembly of three-dimensional prestressed tensegrity structures from DNA," *Nature Nanotechnology*, vol. 5, no. 7, pp. 520–524, 2010.
- [21] H. Dietz, S. M. Douglas, and W. M. Shih, "Folding DNA into twisted and curved nanoscale shapes," *Science*, vol. 325, no. 5941, pp. 725–730, 2009.
- [22] B. Huang, M. Bates, and X. Zhuang, "Super-resolution fluorescence microscopy," *Annual Review of Biochemistry*, vol. 78, pp. 993–1016, 2009.
- [23] S. W. Hell, "Microscopy and its focal switch," *Nature Methods*, vol. 6, no. 1, pp. 24–32, 2009.
- [24] R. Jungmann, C. Steinhauer, M. Scheible, A. Kuzyk, P. Tinnefeld, and F. C. Simmel, "Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami," *Nano Letters*, vol. 10, no. 11, pp. 4756–4761, 2010.
- [25] S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, and W. M. Shih, "Rapid prototyping of 3D DNA-origami shapes with caDNAo," *Nucleic Acids Research*, vol. 37, no. 15, pp. 5001–5006, 2009.
- [26] R. Jungmann, T. Liedl, T. L. Sobey, W. Shih, and F. C. Simmel, "Isothermal assembly of DNA origami structures using denaturing agents," *Journal of the American Chemical Society*, vol. 130, no. 31, pp. 10062–10063, 2008.
- [27] C. E. Castro, F. Kilchherr, D.-N. Kim et al., "A primer to scaffolded DNA origami," *Nature Methods*, vol. 8, no. 3, pp. 221–229, 2011.
- [28] A. Rajendran, M. Endo, Y. Katsuda, K. Hidaka, and H. Sugiyama, "Programmed two-dimensional self-assembly of multiple DNA origami jigsaw pieces," *ACS Nano*, vol. 5, no. 1, pp. 665–671, 2011.
- [29] K. N. Kim, K. Sarveswaran, L. Mark, and M. Lieberman, "Comparison of methods for orienting and aligning DNA origami," *Soft Matter*, vol. 7, no. 10, pp. 4636–4643, 2011.
- [30] M. Rief, H. Clausen-Schaumann, and H. E. Gaub, "Sequence-dependent mechanics of single DNA molecules," *Nature Structural Biology*, vol. 6, no. 4, pp. 346–349, 1999.
- [31] T. Strunz, K. Oroszlan, R. Schäfer, and H. J. Güntherodt, "Dynamic force spectroscopy of single DNA molecules," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 20, pp. 11277–11282, 1999.
- [32] M. Grandbois, M. Beyer, M. Rief, H. Clausen-Schaumann, and H. E. Gaub, "How strong is a covalent bond?" *Science*, vol. 283, no. 5408, pp. 1727–1730, 1999.
- [33] C. Steinhauer, C. Forthmann, J. Vogelsang, and P. Tinnefeld, "Superresolution microscopy on the basis of engineered dark states," *Journal of the American Chemical Society*, vol. 130, no. 50, pp. 16840–16841, 2008.
- [34] Software for PAINT analysis, <http://www.e14.ph.tum.de>.
- [35] A.-P. Eskelinen, A. Kuzyk, T. K. Kaltiaisenaho et al., "Assembly of single-walled carbon nanotubes on DNA-origami templates through streptavidin-biotin interaction," *Small*, vol. 7, no. 6, pp. 746–750, 2011.
- [36] R. Schreiber, S. Kempter, S. Holler et al., "DNA origami templated growth of arbitrarily shaped metal nanoparticles," *Small*, vol. 7, no. 13, pp. 1795–1799, 2011.
- [37] K. Keren, M. Krueger, R. Gilad, G. Ben-Yoseph, U. Sivan, and E. Braun, "Sequence-specific molecular lithography on single DNA molecules," *Science*, vol. 297, no. 5578, pp. 72–75, 2002.
- [38] J. Liu, Y. Geng, E. Pound et al., "Metallization of branched DNA origami for nanoelectronic circuit fabrication," *ACS Nano*, vol. 5, no. 3, pp. 2240–2247, 2011.

Research Article

Photoswitching of Site-Selective RNA Scission by Sequential Incorporation of Azobenzene and Acridine Residues in a DNA Oligomer

Akinori Kuzuya,^{1,2} Keita Tanaka,¹ and Makoto Komiyama¹

¹Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan

²Department of Chemistry and Materials Engineering, Kansai University, 3-3-35 Kamiyamate-cho, Suita, Osaka 564-8680, Japan

Correspondence should be addressed to Akinori Kuzuya, kuzuya@kansai-u.ac.jp and Makoto Komiyama, komiyama@mkomi.rcast.u-tokyo.ac.jp

Received 10 May 2011; Accepted 6 July 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Akinori Kuzuya et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Photoresponsive systems for site-selective RNA scission have been prepared by combining Lu(III) ions with acridine/azobenzene dual-modified DNA. The modified DNA forms a heteroduplex with substrate RNA, and the target phosphodiester linkages in front of the acridine residue is selectively activated so that Lu(III) ion rapidly cleaves the linkage. Azobenzene residue introduced adjacent to the acridine residue acts as a photoresponsive switch, which triggers the site-selective scission upon UV irradiation. A *trans* isomer of azobenzene efficiently suppresses the scission, whereas the *cis* isomer formed by UV irradiation hardly affects the scission. As a result, 1.7–2.4-fold acceleration of the cleavage was achieved simply by irradiating UV for 3 min to the mixture prior to the reaction. Considering the yield of photoisomerization, the intrinsic activity of a *cis* isomer is up to 14.5-fold higher than that of the *trans* isomer.

1. Introduction

In this couple of decades, significant attention has been focused on site-selective RNA scission, since it is indispensable for future molecular biology and therapy [1–3]. Discovery of important roles of short RNA in living cells further promoted this [4]. We have recently developed efficient artificial systems for site-selective RNA scission by combining a metal ion (lanthanide ions or some transition metal ions) as molecular scissor and an acridine-modified DNA as a sequence selective RNA activator [5]. Either of the 5'- or 3'-phosphodiester linkage of the target nucleotide in front of the acridine moiety, which is in protonated form under neutral condition, is site-selectively activated through general acid catalysis [6]. When Lu(III) ion is used as the catalyst, the general acid catalysis preferentially promotes the cleavage at the 3'-side linkage. In addition, conformational change of RNA backbone caused by acridine intercalation is thought to be another important factor in

the activation. The other portions of RNA are protected from metal ion-induced hydrolysis by duplex formation with DNA additives. Sequence of the target site can be freely chosen, and the reaction is selective and efficient enough to achieve simultaneous tandem scission in close proximity as small as 10 nucleotides [7]. This technique has been applied to new genotyping methods for single-nucleotide (SNP) or insertion-deletion (indel) polymorphisms [8].

One significant advantage of this system is that any desired function can be added to it by additional modification to the acridine-modified DNA. One such example is the addition of a ligand to DNA to localize the catalyst near the target site [9]. Introduction of an iminodiacetate ligand to the point over the major groove results in more than 3-fold acceleration of the site-selective cleavage. Photoswitching of the cleavage is another attractive functionalization to the system, since light can trigger the reaction at any desired timing without changing chemical or physical conditions [10]. Azobenzene is one of the most popular photoresponsive

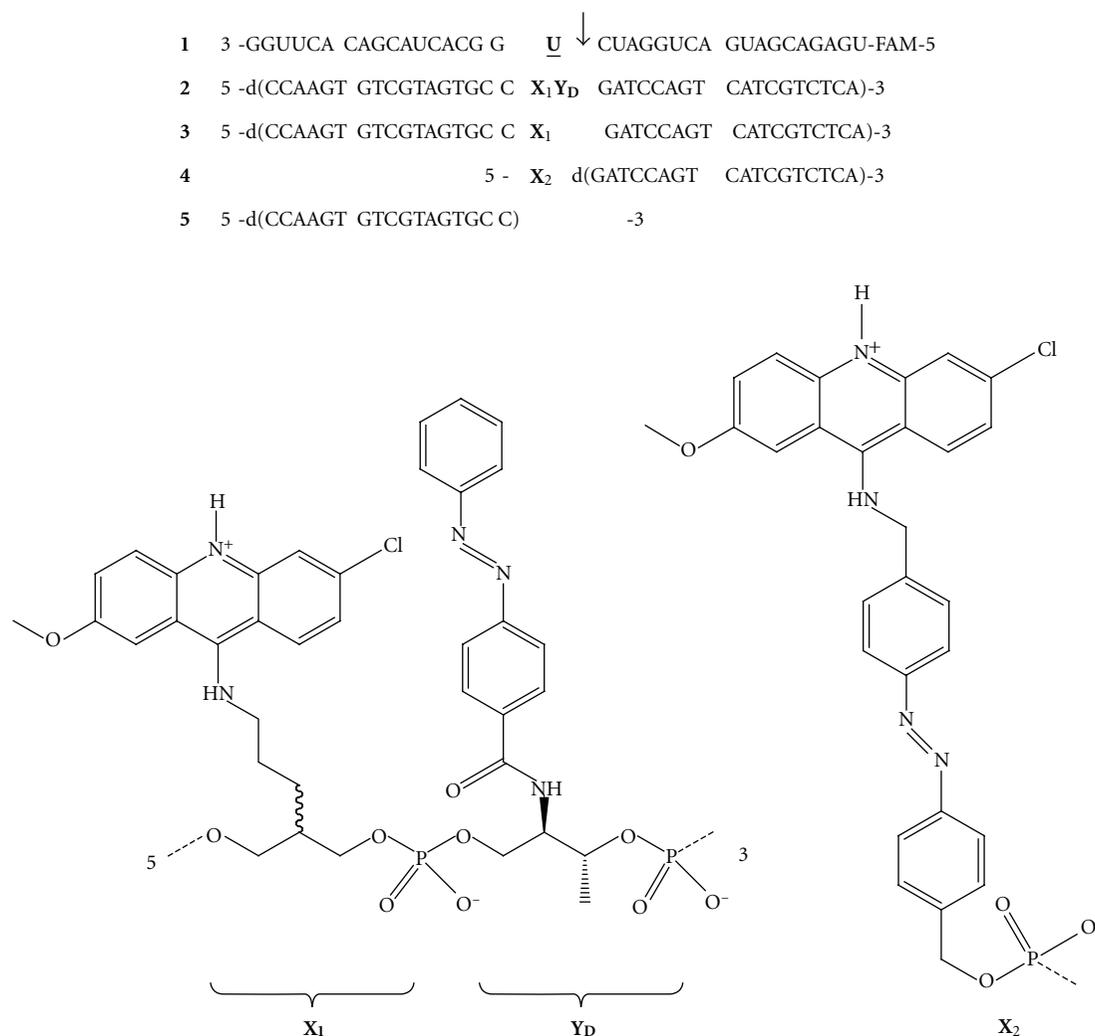


FIGURE 1: Structures of the substrate RNA (1) and the modified DNA (2–4). The target phosphodiester linkage (the 5'-phosphate of U-19) is indicated by the arrow.

molecules used in such studies. Azobenzene isomerizes from *trans* isomer, which is planer and rather hydrophobic, into bulky and polar *cis* isomer upon UV ($\lambda \approx 320$ nm) irradiation and goes back to *trans* isomer upon visible light ($\lambda \approx 450$ nm) irradiation.

We have reported the first photoresponsive system of site-selective RNA scission by using acridine-modified DNA and free Mn(II) ion as a cleaving catalyst [11]. There, an azobenzene group was inserted between the acridine ring and the backbone of DNA as a linker (see X₂ in Figure 1). Although phototriggered acceleration of site-selective RNA scission was achieved in the system, the cleavage activity was quite marginal (the yield of selective cleavage after 18 h was 4%), and the magnitude of acceleration was only 40% after UV irradiation. Accordingly, more active and clear-cut photocontrol of site-selective RNA scission using free Lu(III) ion as the catalyst has been desired. In this study, new photocontrollable site-selective RNA activator has been prepared by introducing an independent azobenzene residue to DNA oligomer in combination with an acridine residue

(Figure 1). Such azobenzene residues in *trans* form stack on the acridine in the adjacent residue and efficiently inhibit the site-selective RNA activation. When the solution is irradiated with UV light and the azobenzene isomerizes into *cis* isomer, the acridine is released from stacking, and the cleavage reaction is significantly accelerated. Clear-cut photo-control of efficient site-selective RNA scission has been successfully accomplished.

2. Materials and Methods

2.1. Spectroscopy. ¹H NMR spectra were obtained on a 500 MHz NMR spectrometer (Bruker Biospin). For electronic spray ionization (ESI) mass spectroscopy, an HITACHI M-8000 LC/3DQMS mass spectrometer was used. MALDI-TOF/MS was measured on a Bruker Daltonics Auto-FLEX mass spectrometer.

2.2. Preparation of Oligonucleotides. The phosphoramidite of X₃ was synthesized starting from the coupling between

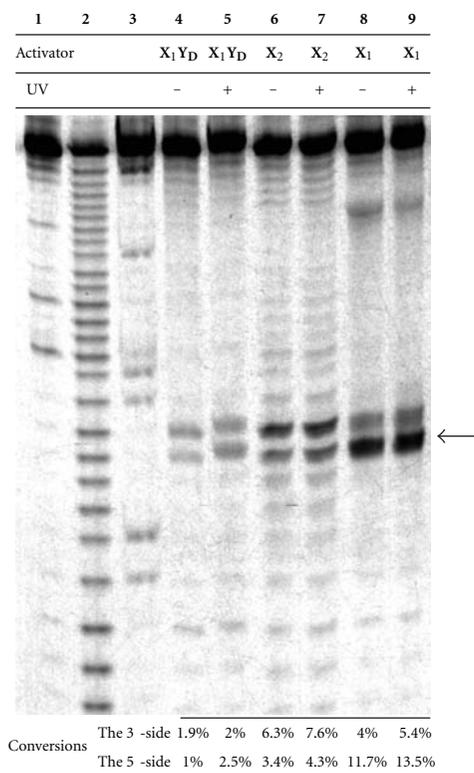


FIGURE 2: A denaturing 20% PAGE pattern for site-selective scission of **1** activated by **2** or (**4** + **5**) in the presence of Lu(III) with or without UV irradiation. Lane 1, control in a buffer solution; lane 2, treatment with Lu(III) alone; lane 3, RNase T1 digestion; lane 4, **2** and Lu(III) without UV irradiation; lane 5, **2** and Lu(III) with UV irradiation; lane 6, **4**, **5**, and Lu(III) without UV irradiation; lane 7, **4**, **5**, and Lu(III) with UV irradiation; lane 8, **3** and Lu(III) without UV irradiation; lane 9, **3** and Lu(III) with UV irradiation. The 5'-side cleavage of U-19 is indicated by the arrow. Conversions of each scission are presented in the bottom. Reaction conditions; [RNA] = 5 μ M, [DNA] = 10 μ M, [LuCl₃] = 100 μ M, [Tris] = 10 mM, [NaCl] = 200 mM, pH 7.5, 37°C, 4 h.

2-chloro-4-nitrobenzoic acid and 4-isopropoxyaniline as described in [6, 12]. The synthetic route for the phosphoramidite of X₂ is described in [11]. The phosphoramidites for Y_D and Y_L were synthesized according to [13]. All the oligonucleotides were synthesized on an ABI 3400 DNA Synthesizer in 1 μ mol scale. Reagents for automated DNA syntheses were purchased from Glen Research Co. (VA, USA). For the synthesis of modified DNA, an extended coupling time of 10 min was adopted for the coupling of azobenzene- and acridine-phosphoramidite monomers. The DMTr-off oligonucleotides were cleaved from the support and deprotected by a treatment with methanolic sodium hydroxide (1 mL, 0.4 M, methanol:water = 4:1) at room temperature for 16 h.

2.3. Purification and Characterization of the Modified Oligonucleotides. The oligonucleotides were first desalted by Poly-Pak II cartridges (Glen Research Co.), and the resulting crude products were then purified by denaturing 20% PAGE.

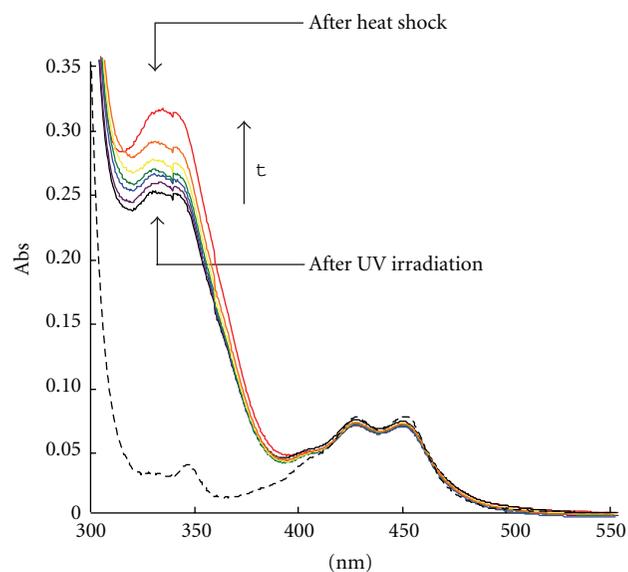


FIGURE 3: Time dependence of UV-VIS absorption of 1/2 complex. The spectra were measured 0 (black), 1 (purple), 6 (blue), 11 (green), 16 (yellow), and 21 h (orange) after UV irradiation at pH 7.5 and 37°C. The spectrum after heat shock (*trans*-dominant) is drawn in red. The dashed line represents the spectrum of 1/3 complex without azobenzene.

TABLE 1: MALDI-TOF/MS analyses of the conjugates.

	Calcd. <i>m/z</i>	
	[M + H]	Found
2	11511.1	11509.5
3	11136.8	11116.4
6	11592.8	11566.0
7	11592.8	11563.6
8	11592.8	11563.8
9	11592.8	11543.5
10	11218.5	11250.5

Final purification was carried out on a reversed phased HPLC equipped with an RP-C18 column (Cica-Merck LiChro-CART 125-4; a linear gradient of 0%–25% acetonitrile with ammonium formate (50 mM) over 25 min; flow rate 0.5 mL/min). They were fully characterized by MALDI-TOF/MS in the positive ion mode (Table 1). Concentration of the stock solution of each oligonucleotide was determined both with UV absorption of DNA at 260 nm and HPLC quantification of each nucleosides formed by digestion with snake venom phosphodiesterase and alkaline phosphatase.

2.4. RNA Cleavage Assay. A mixture of the substrate RNA (18 μ L, 5.5 μ M) and modified DNA (11 μ M) in pH 7.5 Tris buffer (11 mM) containing NaCl (220 mM) was prepared and divided into halves. Both portions were heated to 90°C for 1 min and slowly cooled to room temperature. One of the portions was irradiated with UV light from UV Spot Light Source (Hamamatsu Photonics) through a UV-D36C

6	5 -d(CCAAGT GTCGTAGTGC C	X ₃ Y _L	GATCCAGT	CATCGTCTCA)-3
7	5 -d(CCAAGT GTCGTAGTGC C	Y _L X ₃	GATCCAGT	CATCGTCTCA)-3
8	5 -d(CCAAGT GTCGTAGTGC C	X ₃ Y _D	GATCCAGT	CATCGTCTCA)-3
9	5 -d(CCAAGT GTCGTAGTGC C	Y _D X ₃	GATCCAGT	CATCGTCTCA)-3
10	5 -d(CCAAGT GTCGTAGTGC C	X ₃	GATCCAGT	CATCGTCTCA)-3

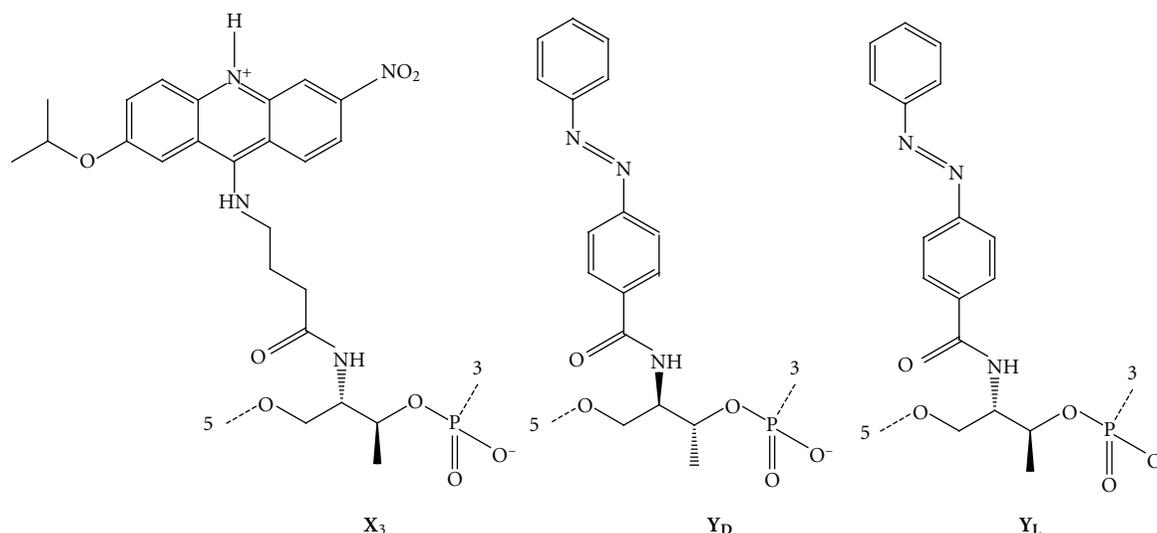


FIGURE 4: Structures of the modified DNA bearing more active acridine (X₃) and the two azobenzene residues (Y_D and Y_L).

(Asahi Technoglass) for 3 min to isomerize *trans*-azobenzene into its *cis* isomer. Selective cleavage reaction was initiated by adding 1 μ L of aqueous LuCl₃ solution (1 mM, final concentration was 100 μ M) to the mixture. After incubation for a predetermined reaction time at 37°C, the reaction was quenched by 1 μ L of EDTA·2Na solution (100 mM) and analyzed on 20% denaturing PAGE. All the reactions were carried out in black tubes so that the *cis* isomer did not isomerize into *trans* isomer by ambient visible light.

3. Results and Discussion

3.1. Photoresponsive Site-Selective RNA Scission by a Combination of Lu(III) Ion and Acridine-Azobenzene Dual-Modified DNA. To achieve efficient photo-control of site-selective RNA scission catalyzed by Lu(III) ion, new acridine/ azobenzene dual-modified DNA was synthesized (2 in Figure 1). This DNA has an acridine residue (X₁) as the 18th residue. The acridine ring in X₁ is 9-amino-6-chloro-2-methoxyacridine. This residue is the most popular acridine residue, which is often used as a duplex stabilizer or a fluor-escent label [14, 15]. The linker moiety in X₁ is a flexi-ble alkyl chain, and the stereochemistry of the branching point is not controlled. Adjacent to X₁, an azobenzene resi-due (Y_D) was also introduced as the 19th residue. This conjugate of D-threoninol and azobenzene is known to be a good photoswitch of duplex formation [13]. When the corresponding simple acridine-modified DNA 3 is

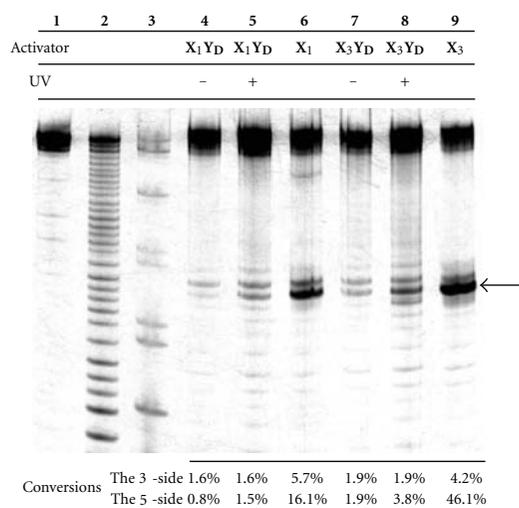


FIGURE 5: A denaturing 20% PAGE pattern for site-selective scission of 1 activated by 2, 3, 8, or 10. Lane 1, control reaction in a buffer solution; lane 2, treatment with Lu(III) alone; lane 3, RNase T1 digestion; lane 4, 2 and Lu(III) without UV irradiation; lane 5, 2 and Lu(III) with UV irradiation; lane 6, 3 and Lu(III) without UV irradiation; lane 7, 8 and Lu(III) without UV irradiation; lane 8, 8 and Lu(III) with UV irradiation; lane 9, 10 and Lu(III) without UV irradiation. The 5'-side cleavage of U-19 is indicated by the arrow. Conversions of each scission are presented in the bottom. Reaction conditions; [RNA] = 5 μ M, [DNA] = 10 μ M, [LuCl₃] = 100 μ M, [Tris] = 10 mM, [NaCl] = 200 mM, pH 7.5, 37°C, 4 h.

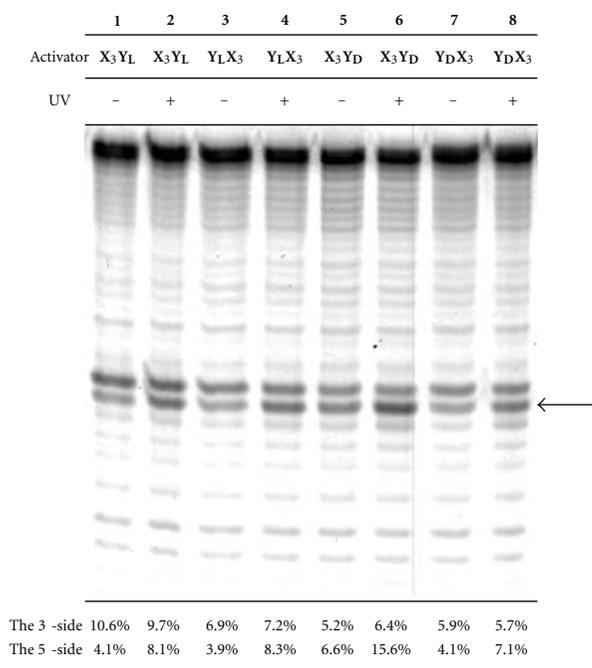


FIGURE 6: A denaturing 20% PAGE pattern for site-selective scission of **1** activated by **6–9**. Lane 1, **6** and Lu(III) without UV irradiation; lane 2, **6** and Lu(III) with UV irradiation; lane 3, **7** and Lu(III) without UV irradiation; lane 4, **7** and Lu(III) with UV irradiation; lane 5, **8** and Lu(III) without UV irradiation; lane 6, **8** and Lu(III) with UV irradiation; lane 7, **9** and Lu(III) without UV irradiation; lane 8, **9** and Lu(III) with UV irradiation. The 5'-side cleavage of U-19 is indicated by the arrow. Conversions of each scission are presented in the bottom. Reaction conditions; [RNA] = 5 μ M, [DNA] = 10 μ M, [LuCl₃] = 100 μ M, [Tris] = 10 mM, [NaCl] = 200 mM, pH 7.5, 37°C, 7 h.

hybridized to the complementary substrate RNA **1**, the target U-19 residue alone remains unpaired, and the 5'-phosphate linkage (indicated by the arrow in Figure 1) of the target is efficiently cleaved by Lu(III) ion in the aid of acid catalysis by the acridine. The 3'-phosphate linkage is also marginally cleaved, but it is simply caused by conformational changes of RNA backbone and is not much affected by the acridine. As a control, previous dual-modified DNA (**4**) that bears an azobenzene as a linker moiety was also used for the cleavage experiment in combination with an unmodified DNA (**5**) and Lu(III).

Figure 2 is a typical denaturing 20% polyacrylamide gel electrophoresis (PAGE) pattern of the cleavage reactions. When **1** was hybridized to **2** and Lu(III) ion was added to the solution under dark, the 3'-phosphodiester of U-19 was only marginally cleaved, and the target 5'-phosphodiester was almost intact after 4 h reaction at 37°C (the lower band in lane 4). The cleavage yield was 1.9% and 1.0%, respectively. Compared with 11.7% for the 5'-side cleavage using simple acridine-modified DNA **3** (lane 8), addition of *trans*-azobenzene adjacent to the acridine almost completely suppressed the selective cleavage. When the annealed mixture was irradiated with UV prior to Lu(III) addition and the

azobenzene was isomerized into *cis*-form, on the other hand, the cleavage at the 5'-phosphodiester significantly recovered (lane 5). The cleavage yield at the 5'-phosphodiester was 2.5% here (the yield at the 3'-phosphodiester was 2.0% again). Similar cleavage yields were achieved in two more repeats of the reaction with independently prepared reaction mixtures. More than two times acceleration of the cleavage was achieved selectively at the target linkage simply with 3-min UV irradiation. In contrast, little change was observed in the cleavage in the presence of previous **4 + 5** combination (lanes 6 and 7). This combination preferentially promotes the 3'-side cleavage unlike the other acridine-modified DNA, but the ratio of the 3'-side cleavage to the 5'-side cleavage was both 1.8 irrespective of the UV irradiation. Previous X₂ is not capable of photo-control of Lu(III) catalyzed RNA hydrolysis. As expected, UV irradiation did not affect the cleavage using **3**, which does not bear azobenzene, as well (lanes 8 and 9). Effective photoregulation of site-selective RNA scission catalyzed by Lu(III) is achievable only with the present system where the acridine and the azobenzene moieties are separately introduced to DNA as individual residues.

3.2. Estimation of the Yield of Photoisomerization in DNA/RNA Heteroduplex. According to the previous studies, photoisomerization of azobenzene introduced in DNA is somewhat suppressed when the DNA forms duplex with complementary DNA or RNA, probably because of stacking interaction with adjacent base pairs. The yield of *trans* to *cis* isomerization is reported to be only 20%–40% in DNA/DNA duplex [16]. The isomerization of azobenzene in **2** is also expected to be not perfect. In addition, it is known that *cis* to *trans* reverse isomerization is triggered not only by visible light irradiation, but thermal isomerization also occurs in a significant rate [17]. It is necessary to estimate these properties of Y_D introduced next to the acridine residue.

The population of each isomer can be estimated by monitoring the change of absorbance at 332 nm before and after UV irradiation. Figure 3 shows UV spectra of **1/2** complex before and after UV irradiation. After UV was irradiated to the solution, significant hypochromicity at 332 nm was observed. The absorbance gradually recovered thereafter, and the first-order rate constant of reverse isomerization was obtained from the time dependence of Abs₃₃₂ to be 0.032 h⁻¹. This number corresponds to the half-life of 16 h. It is fairly long enough for the reaction time employed in this study. Considering that $\epsilon_{332, trans}/\epsilon_{332, cis} = 9.15$ and the ratio of the population of *trans*- to *cis*-isomer is 94:6 after heat shock (90°C for 1 min) [18], the ratio just after UV irradiation is estimated to be 70:30. Almost no change was observed for the absorbance at 260 nm before and after UV irradiation (data not shown). Thus, isomerization of the azobenzene in such a long oligonucleotides does not affect the overall duplex formation under the conditions employed here.

3.3. Still Faster Photocontrolled Cleavage by Use of More Active Acridine Residue. The acridine residue X₁ used in the

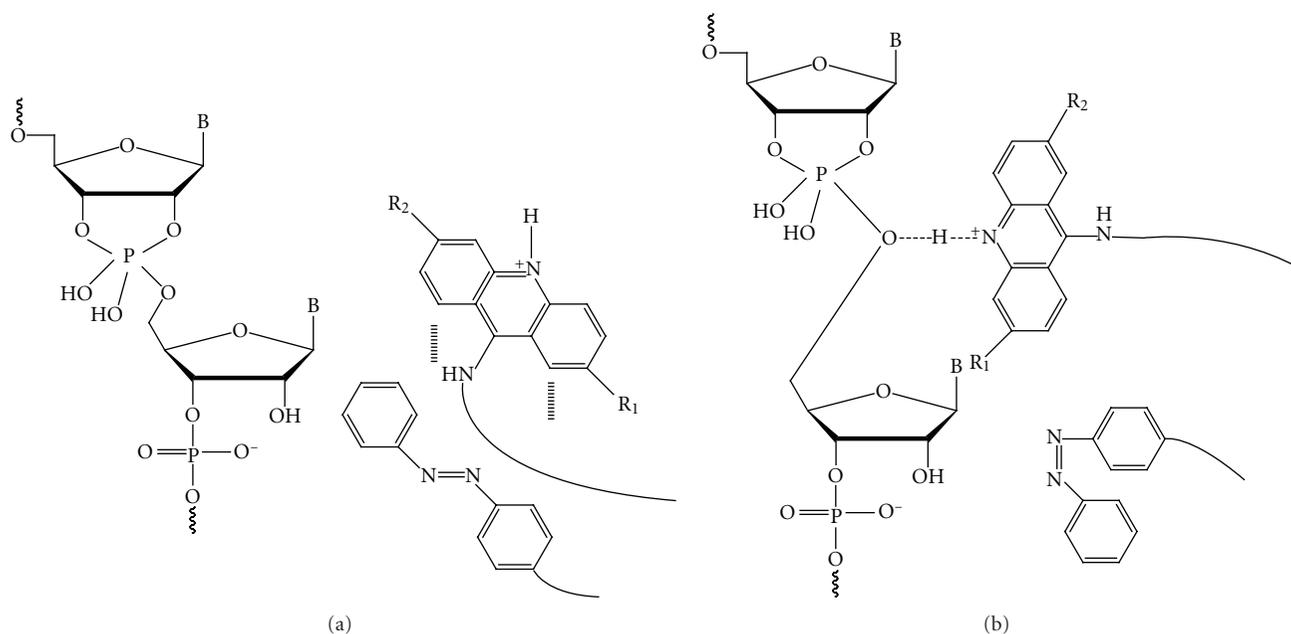


FIGURE 7: A proposed mechanism of the Photoswitching of site-selective RNA scission by azobenzene residues. (a) Azobenzene residues in *trans* form stack on the acridine in the adjacent residue and efficiently inhibit site-selective RNA activation. (b) When the solution is irradiated with UV light and the azobenzene isomerizes into *cis* isomer, the acridine is released from stacking and the cleavage reaction is significantly accelerated.

above experiments is commercially available and bears 9-amino-6-chloro-2-methoxyacridine. In our previous studies optimizing the structure of the acridine ring and the linker moiety, far more active acridine residue (**X₃** in Figure 4) has been developed [12, 19]. This **X₃** bears more acidic 9-amino-2-isopropoxy-6-nitroacridine via optically pure linker based on L-threoninol and 4-aminobutanoic acid and shows nearly three times higher RNA activation ability than **X₁**. By using **X₃**, four kinds of dual-modified oligonucleotides were prepared. The dual-modified oligonucleotide **8** is a counterpart of **2** and bears **X₃** in place of **X₁** in a combination with **Y_D** in the 3' side. The oligonucleotide **9** also bears **X₃** and **Y_D** in the middle, but **Y_D** is inserted in the 5' side of **X₃**. A previous study has shown that the steric effect caused by the isomerization of azobenzene strongly depends on the regiochemical conformation of the threoninol backbone [13]. For the purpose of comparison, **Y_L**, a diastereomer of **Y_D**, was prepared and incorporated into oligonucleotides **6** and **7**.

Figure 5 shows a comparison of site-selective RNA scission using oligonucleotides bearing 9-amino-6-chloro-2-methoxyacridine (**2** and **3**) and 9-amino-2-isopropoxy-6-nitroacridine (**8** and **10**). As in lane 9, the oligonucleotide bearing 9-amino-2-isopropoxy-6-nitroacridine via L-threoninol linker (**10**) also efficiently activates RNA and promotes site-selective scission catalyzed by Lu(III). The efficacy is nearly three times as high as that of **3**. When an azobenzene residue with D-threoninol linker (**Y_D**) in *trans* form was inserted to the 3' side of the acridine residue (**8**, lane 7), the site-selective cleavage was drastically suppressed just as in the system with 9-amino-6-chloro-2-methoxyacridine (**2**,

lane 4). However, the cleavage activity significantly recovered when UV was irradiated to the reaction mixture for 3 min prior to the initiation of cleavage reaction (lane 8). The yield of site-selective cleavage at the 5'-phosphate is again twice as high as that without UV irradiation.

To evaluate how the position of insertion and the regiochemical conformation of an azobenzene residue affect on its Photoswitching yield, site-selective RNA scission using **6–9** was performed (Figure 6). Note that the reaction time in this experiment was extended to 7 h to obtain more explicit difference in the cleavage yields. The recovery of the reaction activity after UV irradiation (=the yield of the 5'-phosphate cleavage with UV irradiation/that without UV irradiation) was 2.0-fold for **6**, 2.1-fold for **7**, 2.4-fold for **8**, and 1.7-fold for **9**. These results indicate that an azobenzene residue in *trans* form efficiently inhibits site-selective RNA activation irrespective of the position of its insertion and the regiochemical conformation of the linker moiety, and an azobenzene residue in *cis* form does not. In all of the cases, change in the 3'-side scission before and after UV irradiation was quite marginal. Considering that the general acid catalysis by protonated acridine has almost nothing to do with the 3'-side cleavage [6, 19], the present enhancement of the 5'-side cleavage may be mainly because of the recovery in general acid catalysis. It has been shown that an azobenzene in *trans* form is sufficiently planer and hydrophobic to stack on adjacent DNA base pairs when it is introduced into DNA duplex [20]. Strong stacking interaction in DNA duplex is also reported for various dyes bearing similar azobenzene rings [21]. The efficient inhibition of the function of acridines observed in

the present study may thus be attributed to the stacking interaction between *trans*-azobenzene and the acridines as well (Figure 7). This mechanism is also supported by an observation that incorporation of two acridine residues into the target site strongly decreases the cleavage yield [5]. The *trans* to *cis* isomerization triggered by UV irradiation releases the acridine from the stacking, and let it activates RNA. If this is the case, 1.7–2.4-fold recovery of the cleavage activity observed in Figure 6 after UV irradiation can be attributable mainly to the increase of the population of *cis* isomer from 6% to 30%. The ratio of the intrinsic activity of the *cis* isomer of **8** to the *trans* isomer can be calculated from

$$2.4(0.06x + 0.94y) = 0.30x + 0.70y, \quad (1)$$

where x is the intrinsic activity of the *cis* isomer and y is that of the *trans* isomer. Thus, the activity of 100% *cis* isomer calculated from these numbers is as much as 14.5 times higher than that of 100% *trans* isomer for **8** and 4.5 times higher for **9**.

4. Conclusions

Efficient photo-control of site-selective RNA scission have been accomplished by combining free Lu(III) ion with dual-modified oligonucleotides which bear photoresponsive azobenzene residue adjacent to the acridine residue. Intrinsic activity of *cis* isomers of the dual-modified oligonucleotides is estimated to be up to 14.5 times as high as that of *trans* isomers. The suppressed activities by the introduction of *trans* azobenzenes were recovered by a factor of 2–3-fold simply with 3-min UV irradiation. Since the isomerization of azobenzene is reversible and *cis*- to *trans*-isomerization is with visible light, switching-off of the selective cleavage is also quite feasible with the present design. The most important advantage of the present design is the ease of individual optimization of the acridine and the azobenzene residues, as been presented for the acridine residue in this study. Various azobenzene derivatives with additional substituents of improved properties such as higher photoisomerization yield or enhanced thermal stability have been developed and introduced to DNA to date [22, 23]. Perfect on-off switching of RNA cleavage may be feasible by further promoting the isomerization using such improved azobenzene monomers. Development of further versatile tools in biochemistry is expected.

Acknowledgments

The authors thank Professor Hiroyuki Asanuma of Nagoya University for helpful suggestions. This study was partially supported by a Grant-in-Aid for Specially Promoted Scientific Research (18001001) and Grant-in-Aid for Young Scientists (B) (20750126) from the Ministry of Education, Science, Sports, Culture, and Technology, Japan, and by the Global COE Program for Chemistry Innovation.

References

- [1] A. Kuzuya and M. Komiyama, "Site-selective artificial ribonucleases and their applications," *Current Organic Chemistry*, vol. 11, no. 16, pp. 1450–1459, 2007.
- [2] T. Niittymaki and H. Lonnberg, "Artificial ribonucleases," *Organic and Biomolecular Chemistry*, vol. 4, no. 1, pp. 15–25, 2006.
- [3] M. Murtola and R. Stromberg, "PNA based artificial nucleases displaying catalysis with turnover in the cleavage of a leukemia related RNA model," *Organic and Biomolecular Chemistry*, vol. 6, no. 20, pp. 3837–3842, 2008.
- [4] J. Kurreck, "RNA interference: from basic research to therapeutic applications," *Angewandte Chemie International Edition*, vol. 48, no. 8, pp. 1378–1398, 2009.
- [5] A. Kuzuya, R. Mizoguchi, F. Morisawa, K. Machida, and M. Komiyama, "Metal ion-induced site-selective RNA hydrolysis by use of acridine-bearing oligonucleotide as cofactor," *Journal of the American Chemical Society*, vol. 124, no. 24, pp. 6887–6894, 2002.
- [6] A. Kuzuya, K. Machida, and M. Komiyama, "A highly acidic acridine for efficient site-selective activation of RNA leading to an eminent ribozyme mimic," *Tetrahedron Letters*, vol. 43, no. 46, pp. 8249–8252, 2002.
- [7] A. Kuzuya, R. Mizoguchi, T. Sasayama, J. M. Zhou, and M. Komiyama, "Selective activation of two sites in RNA by acridine-bearing oligonucleotides for clipping of designated RNA fragments," *Journal of the American Chemical Society*, vol. 126, no. 5, pp. 1430–1436, 2004.
- [8] T. Sasayama, M. Kato, H. Aburatani, A. Kuzuya, and M. Komiyama, "Simultaneous genotyping of indels and SNPs by mass spectroscopy," *Journal of the American Society for Mass Spectrometry*, vol. 17, no. 1, pp. 3–8, 2006.
- [9] A. Kuzuya, Y. Shi, T. Sasayama, and M. Komiyama, "Cooperation of metal-ion fixation and target-site activation for efficient site-selective RNA scission," *Journal of Biological Inorganic Chemistry*, vol. 10, no. 3, pp. 270–274, 2005.
- [10] G. Mayer and A. Hechel, "Biologically active molecules with a "light switch"," *Angewandte Chemie International Edition*, vol. 45, no. 30, pp. 4900–4921, 2006.
- [11] K. Tanaka, Y. Yamamoto, A. Kuzuya, and M. Komiyama, "Synthesis of photo-responsive acridine-modified DNA and its application to site-selective RNA scission," *Nucleosides, Nucleotides and Nucleic Acids*, vol. 27, no. 10–11, pp. 1175–1185, 2008.
- [12] Y. Shi, K. Machida, A. Kuzuya, and M. Komiyama, "Design of phosphoramidite monomer for optimal incorporation of functional intercalator to main chain of oligonucleotide," *Bioconjugate Chemistry*, vol. 16, no. 2, pp. 306–311, 2005.
- [13] H. Asanuma, T. Takarada, T. Yoshida, D. Tamaru, X. G. Liang, and M. Komiyama, "Enantioselective incorporation of azobenzenes into oligodeoxyribonucleotide for effective photoregulation of duplex formation," *Angewandte Chemie International Edition*, vol. 40, no. 14, pp. 2671–2673, 2001.
- [14] P. S. Nelson, M. Kent, and S. Muthini, "Oligonucleotide labeling methods. 3. Direct labeling of oligonucleotides employing a novel, non-nucleosidic, 2-aminobutyl-1,3-propanediol backbone," *Nucleic Acids Research*, vol. 20, no. 23, pp. 6253–6259, 1992.
- [15] J. C. François and C. Helene, "Recognition of hairpin-containing single-stranded DNA by oligonucleotides containing internal acridine derivatives," *Bioconjugate Chemistry*, vol. 10, no. 3, pp. 439–446, 1999.

- [16] M. Liu, H. Asanuma, and M. Komiyama, "Azobenzene-tethered T7 promoter for efficient photoregulation of transcription," *Journal of the American Chemical Society*, vol. 128, no. 3, pp. 1009–1015, 2006.
- [17] T. Asano and T. Okada, "Further kinetic evidence for the competitive rotational and inversional Z-E isomerization of substituted azobenzenes," *Journal of Organic Chemistry*, vol. 51, no. 23, pp. 4454–4458, 1986.
- [18] M. Liu, *Photoregulation of Enzymatic Reaction by Azobenzene-Tethered DNA*, Ph.D. thesis, The University of Tokyo, Tokyo, Japan, 2006.
- [19] A. Kuzuya, Y. Shi, K. Tanaka, K. Machida, and M. Komiyama, "Efficient site-selective RNA activation and scission achieved by geometry control of acridine intercalation in RNA/DNA heteroduplex," *Chemistry Letters*, vol. 38, no. 5, pp. 432–433, 2009.
- [20] X. Liang, H. Asanuma, H. Kashida et al., "NMR study on the photoresponsive DNA tethering an azobenzene. Assignment of the absolute configuration of two diastereomers and structure determination of their duplexes in the trans-form," *Journal of the American Chemical Society*, vol. 125, no. 52, pp. 16408–16415, 2003.
- [21] H. Kashida, T. Fujii, and H. Asanuma, "Threoninol as a scaffold of dyes (threoninol-nucleotide) and their stable inter-strand clustering in duplexes," *Organic and Biomolecular Chemistry*, vol. 6, no. 16, pp. 2892–2899, 2008.
- [22] H. Nishioka, X. G. Liang, H. Kashida, and H. Asanuma, "2',6'-dimethylazobenzene as an efficient and thermo-stable photoregulator for the photoregulation of DNA hybridization," *Chemical Communications*, no. 42, pp. 4354–4356, 2007.
- [23] X. G. Liang, N. Takenaka, H. Nishioka, and H. Asanuma, "Molecular design for reversing the photoswitching mode of turning on and off DNA hybridization," *Chemistry—An Asian Journal*, vol. 3, no. 3, pp. 553–560, 2008.

Review Article

Use of Nucleic Acid Analogs for the Study of Nucleic Acid Interactions

Shu-ichi Nakano,^{1,2} Masayuki Fujii,^{3,4} and Naoki Sugimoto^{1,2}

¹ Faculty of Frontiers of Innovative Research in Science and Technology, Konan University, 7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan

² Frontier Institute for Biomolecular Engineering Research, Konan University, 7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan

³ Department of Environmental and Biological Chemistry, Kinki University, 11-6 Kayanomori, Iizuka, Fukuoka 820-8555, Japan

⁴ Molecular Engineering Institute, Kinki University, 11-6 Kayanomori, Iizuka, Fukuoka 820-8555, Japan

Correspondence should be addressed to Shu-ichi Nakano, shuichi@center.konan-u.ac.jp and Naoki Sugimoto, sugimoto@konan-u.ac.jp

Received 14 April 2011; Accepted 2 May 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Shu-ichi Nakano et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Unnatural nucleosides have been explored to expand the properties and the applications of oligonucleotides. This paper briefly summarizes nucleic acid analogs in which the base is modified or replaced by an unnatural stacking group for the study of nucleic acid interactions. We also describe the nucleoside analogs of a base pair-mimic structure that we have examined. Although the base pair-mimic nucleosides possess a simplified stacking moiety of a phenyl or naphthyl group, they can be used as a structural analog of Watson-Crick base pairs. Remarkably, they can adopt two different conformations responding to their interaction energies, and one of them is the stacking conformation of the nonpolar aromatic group causing the site-selective flipping of the opposite base in a DNA double helix. The base pair-mimic nucleosides can be used to study the mechanism responsible for the base stacking and the flipping of bases out of a nucleic acid duplex.

1. Introduction

Nucleic acids have many remarkable properties that other molecules do not possess. The most notable property is the ability of sequence-specific hybridization through Watson-Crick base pairing. Even a short oligonucleotide sequence, readily synthesized chemically and available on the market at a relatively low cost, can self-assemble into a defined structure and hybridize specifically to a target sequence in accordance with the base pair-rule of A/T and G/C. Importantly, the controls of the self-assembly and the hybridization are not difficult when one considers the interaction energy of nucleic acid reactions [1]. Additionally, it is possible to conjugate with other molecules, such as fluorescent dyes, amino acids, and nanoparticles. Thus, the methodologies that utilize DNA and RNA oligonucleotides as a tool for technology such as nanomaterial and medicinal and therapeutic usages have become of broader interest over the past decades.

The most common structure formed by base pairing is the right-handed double helix. The geometry of Watson-Crick base pairs mediated by hydrogen bonding is similar regardless of the nucleotide sequence, and this allows a double helical conformation virtually identical without disrupting coplanar stacking between adjacent base pairs. Interbase hydrogen bonding is responsible for the association of complementary bases, which is essential for the storage and retrieval of genetic information. Hydrogen donors and acceptors on the purine and pyrimidine bases direct the base pair partner by forming two hydrogen bonds in the A/T pair and three in the C/G pair (Figure 1(a)). According to the number of hydrogen bonds, the C/G pair appears more stable than the A/T pair. However, because base stacking is formed simultaneously with the hydrogen bonding, both interactions contribute to the integrity and the thermodynamic stability of base-paired structures. In contrast to hydrogen bonding, the base stacking does not demand a particular interaction

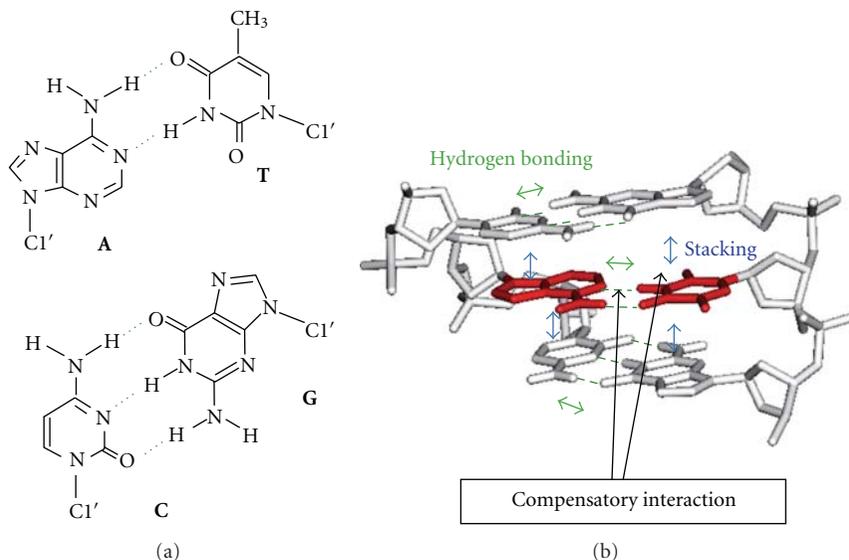


FIGURE 1: (a) Watson-Crick A/T and C/G base pairs. Cl' represents the 1' carbon atom of deoxyribose in DNA. (b) Interbase hydrogen bonding and stacking interactions formed in a DNA duplex. A compensatory relationship is suggested between the interaction energies of the hydrogen bonding and the base stacking.

partner, while the interaction energy between purine bases is usually greater than that between pyrimidine bases due to the larger overlapping area of purine bases. The strength of the stacking interaction has particular relevance to the conformation of unpaired nucleotides, for example, single-stranded overhangs and the helical junction containing a nick site, whether stacked or bent [2–5]. The degree of stacking is also important for the design of fluorescent dye molecules attached to an oligonucleotide [6]. It is an important feature in nucleic acids that the base pair is formed in concert with the binding of cations and water molecules. Because the base pairing brings the sugar-phosphate backbones close to each other which increases the electrostatic repulsion between the phosphate groups, counterions must bind to nucleic acids through Coulomb interaction [7]. Formation of the base pairs also accompanies rearrangements of the hydration layer surrounding nucleic acid chains, especially around the bases and within the helical grooves [8, 9].

The nearest-neighbor model is widely used to account for the thermodynamic behavior of Watson-Crick duplexes. The model assumes that the base pair formation is mostly affected by adjacent (nearest-neighbor) base pairs by taking into account the contributions from base stacking as well as interbase hydrogen bonding. Nearest-neighbor parameters for base pairing have been extensively investigated, and the Gibbs free energy at 37°C (ΔG_{37}°) that ranges from -0.2 to -3.4 kcal mol $^{-1}$ (1 kcal = 4.18 kJ) for each nearest-neighbor base pair is useful to predict the hybridization energy and folding structures of DNA and RNA [2, 10]. Although the energy data include contributions from the hydrogen bonding and the base stacking, the free-energy increments from each interaction have been estimated from the studies using unnatural nucleotides and dangling end residues and by

manipulating known loop interactions [11–13]. Interestingly, the quantitative data suggest that the free energies for forming a single hydrogen bond and the stacking interaction are comparable to each other, providing from -0.2 to -1.8 kcal mol $^{-1}$ in ΔG_{37}° under a competitive correlation (Figure 1(b)), where the base pairing with a lower hydrogen bond energy provides a greater stacking energy [11]. The phenomenon can be accounted for by assuming the interaction mechanism in which the geometry optimized for interbase hydrogen bonding is not suitable for base stacking and *vice versa*. On the other hand, investigations of the coaxial stacking of nicked and gapped sites suggest that base stacking is the major stabilizing factor in a double helical structure of DNA [3, 5]. Studies on the stacking interaction are important for understanding not only the fundamental aspects of nucleic acid interactions but also the biological processes involving base pair formation and strand opening, such as DNA replication and refolding of nucleic acid structures.

Many unnatural nucleosides have been explored according to various demands of researchers. They have been modified or replaced the nucleotide base (C5-modified uridine nucleosides, N3-modified cytidine nucleosides, nonpolar nucleosides replaced with an aromatic hydrocarbon group, etc.) or the sugar-phosphate backbone (2'-O-modified RNA, phosphorothioate DNA, morpholino oligonucleotide, peptide nucleic acid, locked nucleic acid, etc.), as introduced in preceding articles (e.g., [14–16]). In this, we briefly introduce the nucleic acid analogs possessing an unnatural stacking group. We also describe the nucleoside derivatives of a base pair-mimic structure that we have examined to understand the biochemical properties of nucleic acid interactions, for example, the mechanisms responsible for the nucleotide base stacking and the flipping of bases out of a nucleic acid duplex.

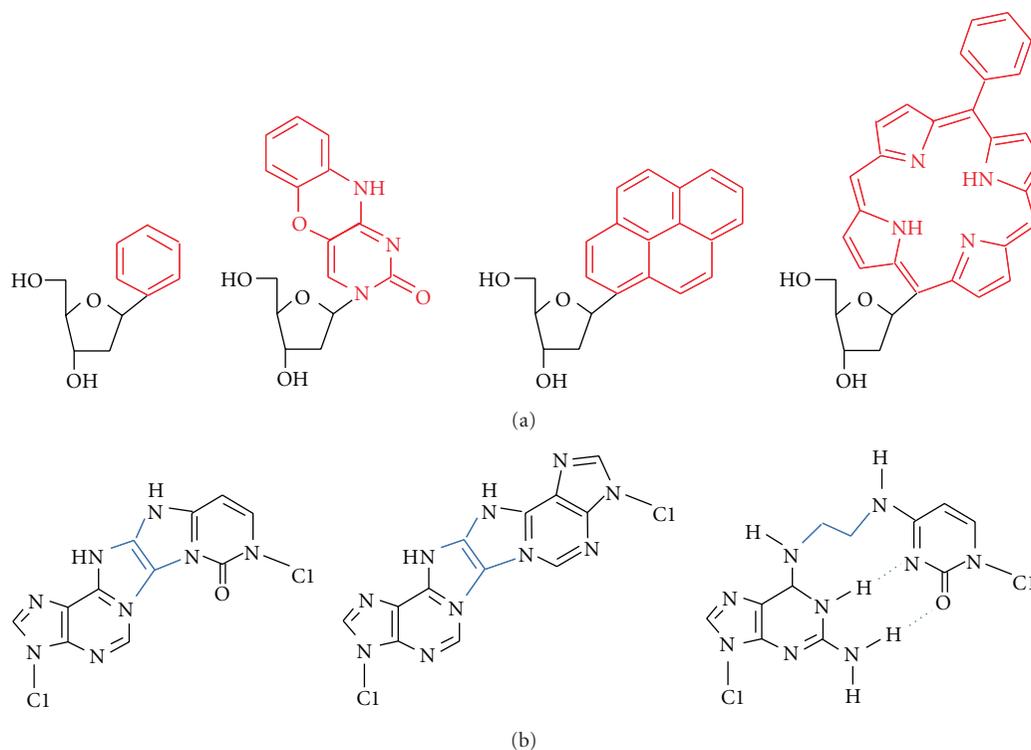


FIGURE 2: (a) Structures of unnatural nucleosides as a base analog with an aromatic hydrocarbon group in place of the purine and pyrimidine bases. (b) Structures of the base pair analogs that provide the interstrand crosslinking sites. The covalent bonds linking the nucleic acid bases are highlighted in blue.

2. Unnatural Nucleosides That Mimic Nucleotide Bases

There are many reports of unnatural nucleosides developed for various purposes. Some are aimed at enhancing the affinity and selectivity in targeting to DNA and RNA sequences by increasing the number of hydrogen bonding sites or by addition of extra aromatic rings to the pyrimidine base [14]. The DNA base analogs lacking particular hydrogen-bond donor and acceptor groups are also used to investigate the influences of the polar groups in DNA bases on the base pair stability [17]. On the other hand, many nonhydrogen-bonding analogs with an aromatic hydrocarbon group in place of the base have been explored (some examples are given in Figure 2(a)). Planar aromatic molecules of an expanded size are beneficial for increasing the interaction energy. If the aromatic group lacks the atoms involved in hydrogen bonding, they may pair with any of the natural bases with little discrimination [15, 18]. The nonpolar base mimics of an aromatic hydrocarbon group, such as benzene, naphthalene, and pyrene, attached to C1' of ribose in place of the purine and pyrimidine bases were incorporated at the end of and in the center of a DNA strand [19]. It was found from the research that a less-polar compound stacked more strongly when molecules of the same size were compared and that the pyrene stacking was the strongest among the tested aromatic groups. However, it is known that the interaction with a strong stacking group often disrupts the helical structure of

DNA. For example, planar polycyclic surrogates possessing fused 1–3 aromatic rings or more intercalate into a DNA duplex and perturbs the helix conformation [20–23]. The covalently appended quinoline residue at the terminal of an oligonucleotide also largely disrupts the DNA duplex structure [24]. The large aromatic groups of the pyrene-modified and porphyrin-modified nucleotides inserted into a DNA helix are found to interfere with the opposite base stacking and are forced to flip to an extrahelical position [25, 26]. The energy cost for the base flipping is quite high due to the loss of base stacking, but it can be compensated by intercalation of the large nonpolar aromatic group into the duplex.

Several types of compounds to introduce a covalently linked base pair portion have been developed to provide interstrand crosslinking sites in DNA strands (Figure 2(b)). In principle, the covalent bonds adducting between the probe strand and a target sequence are not dissociable, so that they are assumed to be useful for applications in gene regulation. There are two strategies for incorporating covalently linked sites in a DNA duplex. One is to use a fused base pair analog consisting of purine and pyrimidine nucleosides linked by covalent bonds [27, 28]. An alternative strategy is to use an unnatural nucleoside bearing a reaction group for alkylation, Schiff base formation, or other types of covalent bond formation triggered by the addition of a reaction reagent or the exposure to a light [29, 30]. Formation of the covalent bonds between two bases is triggered by a sequence-specific hybridization with a target sequence, while particular metal

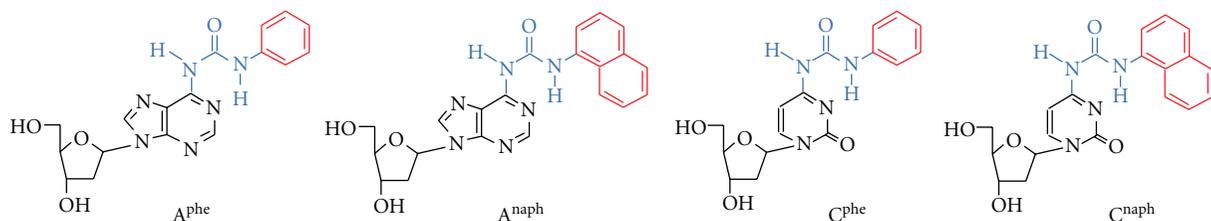


FIGURE 3: Structures of the base pair-mimic nucleosides of deoxyadenosine and deoxycytidine derivatives tethering the nonpolar aromatic group (colored in red) through an ureido linker (blue).

ions (e.g., Ag^+ , Hg^{2+} , and Cu^{2+}) can mediate covalent bonding with the use of natural bases as well as unnatural bases [31, 32]. Because the covalent bonds are formed only when a target site is located at a close distance, molecular design considering the distance between the crosslinking group and the target site is critical.

Chemical synthesis using the solid phase method is widely used for site-selective incorporations of unnatural nucleosides by preparing their phosphoramidite derivatives. On the other hand, DNA polymerase reaction can be applied, especially for incorporation at multiple sites and into a long DNA strand. Many pairs of base analogs that extend the genetic code have been reported [33, 34]. They are selectively incorporated into a DNA strand at desired positions using DNA polymerase, in accordance with their hydrogen-bond donor and acceptor sites and even through steric complementarity of the shape and size of the base analogs. Examinations of whether unnatural nucleosides can be used as a substrate for biological enzymes are important for applications as an anticancer drug and an agonist of receptors and enzymes [35, 36].

3. The Base Pair Analogs of a Base Pair-Mimic Structure

3.1. Design and Synthesis of the Base Pair Analogs Tethering a Nonpolar Stacking Group. Stacking interaction of the purine and pyrimidine bases is mediated by the combination of electrostatic, hydrophobic, and dispersive forces. Although the base pair interaction energy is well studied, the mechanism responsible for the base stacking is poorly understood. We are aiming to understand better the biochemical properties of nucleic acid interactions and the mechanisms behind the stacking interaction by using base pair analogs. For the interaction mechanism study, it is important to design nucleic acid analogs that are compatible with the interaction geometry of canonical base pairs in a double helical conformation. We had designed the compounds tethering a simple aromatic hydrocarbon group of a base pair-mimic structure, as shown in Figure 3: the deoxyadenosine derivatives containing the phenyl group A^{phe} (*N*6-(*N*'-phenylcarbamoyl)-2'-deoxyadenosine) or the naphthyl group A^{napH} (*N*6-(*N*'-naphthylcarbamoyl)-2'-deoxyadenosine) and the deoxycytidine derivatives containing the phenyl group C^{phe} (*N*6-(*N*'-phenylcarbamoyl)-2'-deoxycytidine) or the naphthyl group C^{napH} (*N*6-(*N*'-naphthylcarbamoyl)-2'-deoxycytidine). The base pair analogs of A^{X} and C^{X} , where X is phe or naph, have

a nonpolar base analog of the phenyl or naphthyl group attached to the amino group of deoxyadenosine or deoxycytidine by an ureido linker. Thus, the configuration of the ureido linker is associated with the orientation of the nonpolar aromatic group. The phenyl and naphthyl groups can stack with a nucleic acid duplex when adopting the base pair-mimic geometry, of which the nonpolar base analog occupies the Watson-Crick face of the adenine or cytidine moiety (Figure 4(a)). According to the molecular modeling study, the naphthyl group as well as the phenyl group can be accommodated in a DNA duplex without significant perturbation of the sugar-phosphate backbone conformation when the opposite nucleotide base is absent. On the other hand, the base pairing with a complementary base, $\text{A}^{\text{X}}/\text{T}$ and $\text{C}^{\text{X}}/\text{G}$, through intermolecular hydrogen bonding is allowed when the nonpolar aromatic group is located out of the helix (Figure 4(b)). The potential to adopt two different conformations is characteristic of the base pair-mimic nucleosides shown in Figure 3. It is important to note that the stacking mechanism between the natural bases and the nonpolar aromatic groups is different (Figure 4(c)). In general, stacking of a planar aromatic group can be mediated by the combination of electrostatic, hydrophobic, and dispersive forces. However, less contributions from the hydrophobic effects are suggested for the stacking of natural bases, while the hydrophobic effect and dispersion become more significant than electrostatic forces for the stacking of nonpolar groups [37–40].

Chemical synthesis and incorporation of the base pair-mimic nucleosides into a DNA strand are simple. Synthesis of the deoxyadenosine and deoxycytidine derivatives can be started with 2'-deoxyadenosine and 2'-deoxycytidine, respectively (see the supplemental data in [41]). The compounds are incorporated into an oligonucleotide at high efficiency using an automated synthesizer based on phosphoramidite chemistry. We have prepared the DNA oligonucleotides bearing A^{X} or C^{X} at the end of and in the middle of a sequence. The thermal melting curve was determined to obtain the thermodynamic parameters for DNA structure formations in the 1 M NaCl-phosphate buffer at pH 7.0, which is the condition widely used for determining the stability of oligonucleotide structures. The duplex conformation was investigated using circular dichroism (CD) spectra, a fluorescent base analog, and polyacrylamide gel electrophoresis.

3.2. Dangling End Stacking of the Base Pair-Mimic Nucleosides. According to the nearest-neighbor model, energy

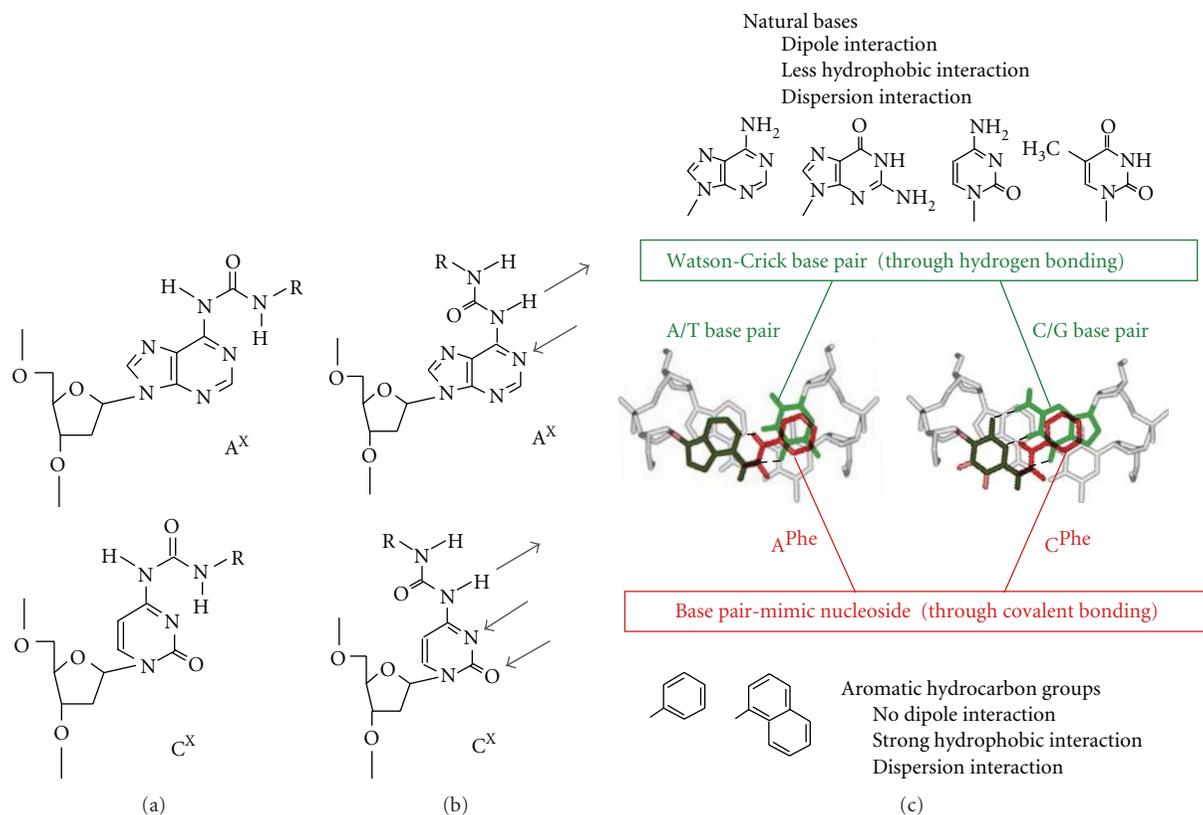


FIGURE 4: (a) and (b) Possible conformations of the deoxyadenosine and deoxycytidine derivatives, the nonpolar group stacking conformation (a) and the base pair conformation (b). indicates the phenyl or naphthyl group. The arrow indicates the site of hydrogen bonding with a complementary base. (c) Comparison of the major interaction forces for the stacking of the A/T and C/G base pairs and the stacking of the A^{Phe} and C^{Phe}. The nonpolar aromatic group in the base pair-mimic nucleosides is indicated in red, and the complementary base is indicated in green.

contribution from the stacking interaction can be evaluated from the interaction energy between the unpaired dangling residue and the adjacent base pair at a helix terminus [42, 43]. Duplex stability increases when the dangling end stacking is significant. For natural DNAs, increments in the interaction energy by a single dangling end ranges from 0.48 to $-0.96 \text{ kcal mol}^{-1}$ in ΔG_{37}° , depending on the dangling end residue, its position at either 5' or 3' of the strand, and the adjacent base pair [42, 43]. Particularly, a dangling A increases the duplex stability by $0.1 \sim 1.0 \text{ kcal mol}^{-1}$, which is greater than that provided by a dangling C ($0.3 \sim -0.5 \text{ kcal mol}^{-1}$), indicating the greater stacking strength of adenine than of cytosine. In contrast, we found that both A^X and C^X provided much more interaction energy ($-0.1 \sim -1.8 \text{ kcal mol}^{-1}$) [44, 45], which was comparable to the formation of Watson-Crick A/T and C/G base pairs ($-0.5 \sim -1.8 \text{ kcal mol}^{-1}$) and the stability reported for the dangling pyrene-modified nucleotide ($-1.7 \text{ kcal mol}^{-1}$) [19, 38]. The large stabilization energy suggests that the nonpolar aromatic groups efficiently stack with the terminal base pair by adopting the base pair-mimic geometry as indicated in Figure 5(a), in which the ureido linker may interact with N1 of adenine or N3 of cytosine.

The dangling end study provides valuable insights into the stacking energy contributed from the nonpolar aromatic groups. The stabilization energies from the dangling A^{Phe} and A^{naph} were similar to each other, and those from C^{Phe} and C^{naph} were as well. The similarity in the energy contributions from the phenyl group and the naphthyl group suggests that the overlapping area of the stacking group, which is relevant to the dispersive and hydrophobic contributions, is not the major determinant for the stacking energy. It has been proposed that the dominant contribution to the stabilization from a dangling end nucleotide comes from the stacking conformation that covers the atoms participating in the hydrogen bonding of an adjacent base pair [4]. In fact, the hydrogen-bonding atoms of the terminal base pair are well covered with the stacked phenyl and naphthyl groups and the ureido linker (Figure 5(a)). It is an interesting finding that large stabilization energy was provided by C^X regardless of the low stacking energy from the cytosine alone. This observation suggests a positively cooperative interaction between the stacking of the base moiety and the stacking of the nonpolar aromatic group. The interplay in the interactions of two planar aligned stacking groups could also be inherent in Watson-Crick base pairs that are noncovalently linked to each other (Figure 5(b)).

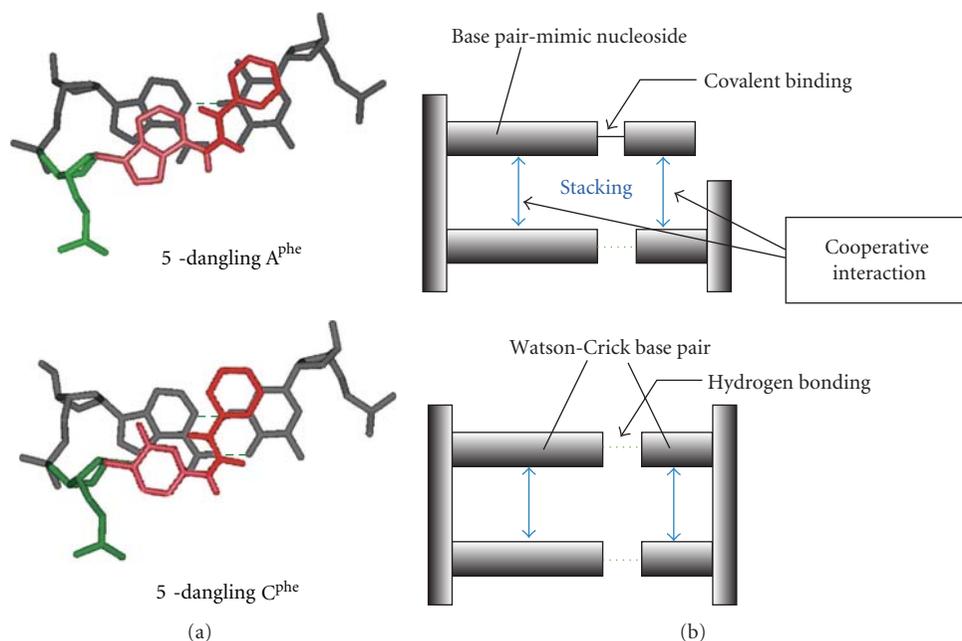


FIGURE 5: (a) Stacking conformations of the dangling A^{phe} and C^{phe} (colored in red) at the 5'-end of a DNA duplex. (b) Side view of the DNA double helix, representing the interaction mechanism of the dangling end stacking (upper) and the Watson-Crick base pairing (lower). Cooperative interaction in the base pair-mimic nucleoside is suggested between stacking of the base moiety and stacking of the nonpolar group.

3.3. Mechanism of the Base Stacking Interaction Revealed by the Base Pair-Mimic Nucleosides. The stacking circumstances between at the terminal and in the center of a DNA duplex are quite different. There is more susceptibility to base pair fraying and water accessibility at the terminal than at the center of a DNA strand. Additionally, there is no remarkable conformational restriction for the stacking at the end; thus the stacking interaction by a dangling end residue can be maximized while the stacking geometry in the center of a DNA duplex is highly restricted. For comparison with the dangling end stacking, we further investigated the DNA duplexes bearing the base pair-mimic nucleosides in a helix center. Because severe steric hindrance with the opposite nucleotide base was expected, the DNA duplexes bearing tetrahydrofuran as an abasic site analog were investigated (Figure 6(a)) [41]. Introduction of the abasic site in an 11-mer DNA duplex largely decreased the duplex stability (by $5.2 \text{ kcal mol}^{-1}$) due to losing the base stacking and providing additional flexibility to the helix. However, the duplex stability was markedly restored by the replacement of A by A^{phe} ($-3.0 \text{ kcal mol}^{-1}$, as the restored free energy) or A^{naph} ($-4.1 \text{ kcal mol}^{-1}$) opposite the abasic site and also by the displacement of C by C^{phe} ($-2.7 \text{ kcal mol}^{-1}$) or C^{naph} ($-3.6 \text{ kcal mol}^{-1}$). The thermodynamic data indicate intercalation of the nonpolar aromatic groups in the DNA duplex by adopting the base pair-mimic geometry presented in Figure 6(a). In contrast to the case of dangling end stacking, the interaction from the naphthyl group was stronger than that from the phenyl group, and their interaction energies were obviously lower than the formation of a Watson-Crick base pair.

It is a remarkable finding that, although the interaction mechanism differs between the base pair-mimic nucleoside and the Watson-Crick base pair, a linear free-energy correlation between them are exhibited: as the interaction free energy from the base pair analog increases, the interaction energy for the corresponding base pair formation (A^{X}/F for A/T base pair, and C^{X}/F for C/G base pair) increases [45]. A similar relationship was obtained with the dangling end data relative to the corresponding Watson-Crick base pairs (A^{X} for A/T base pair and C^{X} for C/G base pair). These observations suggest that the major interaction mechanism that determines the strength of DNA base stacking is maintained in the aromatic group stacking, even though the phenyl and naphthyl groups lack significant dipole moments and hydrogen bonding sites. This finding is useful to understand nucleic acid interactions and to design new unnatural nucleotides with aromatic groups for use in diverse applications.

Our base pair analogs are also useful for the study of DNA hydration. Hydration of a DNA duplex has been extensively studied from the structural point of view [46–50]. While the purine and pyrimidine bases have the hydration sites of the amino group and oxygen and nitrogen atoms, the nonpolar aromatic groups in A^{X} and C^{X} lack hydrogen donor and acceptor sites. Because perturbation of a DNA duplex structure by the nonpolar group stacking is minimized, it would be possible to evaluate the contributions of polar groups and the polarity of DNA bases to the water binding, which is in progress.

3.4. Site-Selective Base Flipping Using the Base Pair-Mimic Nucleosides. Remarkably, even when we did not insert an

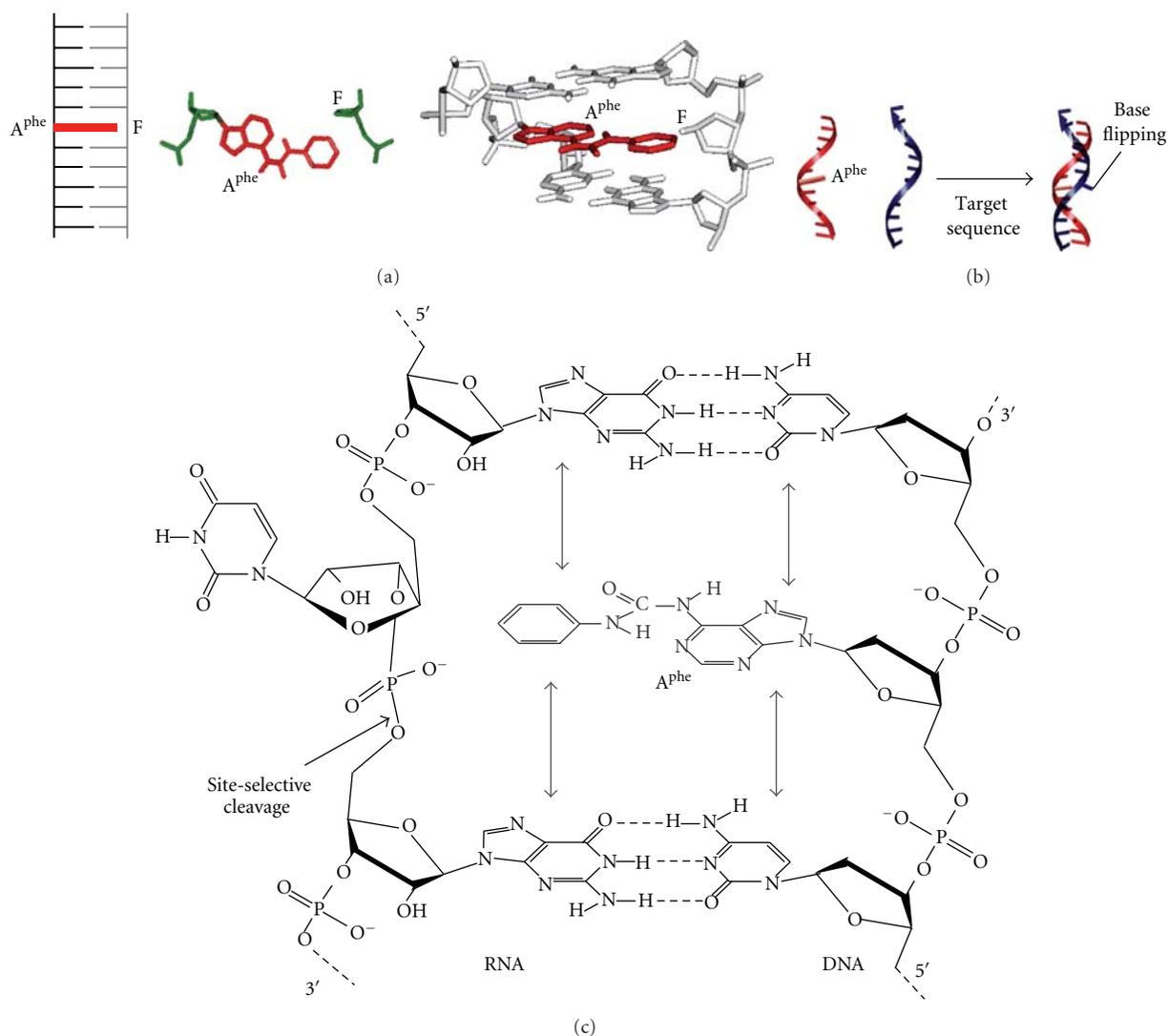


FIGURE 6: (a) The stacking conformation of A^{phe} opposite an abasic site F in the center of a DNA duplex. (b) Hybridization of the DNA strand bearing A^{phe} with a complementary target strand, followed by the formation of the flipped-out conformation. (c) The base flipping conformation induced by A^{phe} in an RNA/DNA duplex. The hybridized RNA strand is cleaved site-selectively at the 5'-side of the phosphodiester bond of the flipped-out ribonucleotide.

abasic site in the complementary DNA strand, the deoxyadenosine derivatives adopted the base pair-mimic geometry by intercalating the nonpolar aromatic group in a DNA duplex with minimized disruptions of the overall duplex structure. Strikingly, the ΔG_{37}° values for forming the DNA duplex containing A^X opposite any nucleotide component (A, G, C, and T) were similar to each other and even similar to when the abasic analog was applied. The stacking of the nonpolar aromatic group causes the opposite base to be flipped out of the duplex (Figure 6(b)), and the resultant duplex becomes synonymous in terms of the double helical conformation regardless of the opposite base component, which was verified from CD spectra, fluorescence measurements using the fluorescent base analog 2-aminopurine, and the mobility in polyacrylamide gel [41]. Although the stacking efficiency is largely influenced by the adjacent base pairs, the base flipping

conformation was suggested for any type of the closing base pairs.

We also tested the base flipping of an RNA strand. The DNA strand containing the base pair analog was hybridized with a complementary RNA sequence, thereby forming an RNA/DNA hybrid duplex. As predicted from the nearest-neighbor parameters determined for the hybrid duplexes [51], the thermal stability of RNA/DNA duplexes containing a mismatch site differed depending on the type of mismatch pair (the melting temperature T_m of the 11-mer natural duplexes forming A/A, A/G, A/C, and A/U pairs differed by 12.9°C). On the other hand, the duplexes containing A^{phe} or A^{naph} in place of A exhibited almost the same stability, within a 2.0°C difference in the T_m among the duplexes containing A^X/A, A^X/G, A^X/C, or A^X/U pair. This observation is consistent with the unstacked conformation of the

RNA base opposite A^X . The sugar-phosphate backbone of RNA perturbed due to the unstacking conformation can be preferentially hydrolyzed as a consequence of specific base catalysis at the site adopting the in-line attack arrangement [52, 53]. Indeed, we found highly site-selective cleavage at any ribonucleotide base opposite A^X in an RNA/DNA duplex [54]. The RNA-hydrolyzing activity agrees with the base flipping model in which A^X forces the opposite base to flip out in an unstacked position (Figure 6(c)). The rate of the RNA hydrolysis was relatively slow comparable to the nonspecific hydrolysis of a single-stranded RNA strand but much faster than those of the unmodified duplexes forming a mismatch pair [55]. Thus, it is likely that A^X induces the base flipping of a structurally unconstrained phosphodiester bond as much as ribonucleotides in a single-stranded state. A highly site-selective hydrolysis without base-pairing selectivity has a great advantage for biotechnology and therapeutic uses, thus the RNA cleavage using an oligonucleotide attaching artificial scissors of a metal ion-coordinating group, such as ion macrocycles, cationic amines, imidazole derivatives, and acridine derivatives, has been reported [56, 57]. However, it is usually difficult to restrict the site to be cleaved because of the difficulty in reducing the nonspecific hydrolysis due to conformational flexibility and distortion. The RNA hydrolysis by our base pair analogs is highly site-selective, which arises from minimized disruptions of the double helical structure and the thermal stability.

3.5. Dual Conformation Depending on the Interaction Energy.

Formation of the base flipping is somewhat surprising from the point view of interaction energy. The flipping conformation is adopted by moving the base from an intrahelical to an extrahelical position. Energetics of the equilibrium between the base stacking conformation and the flipped-out conformation are important to understand the mechanism of the flipping of bases out of a DNA duplex. However, the energy levels of these two conformations are usually largely different, and thus, the base stacking conformation cannot be formed without any cofactor. In nature, the base flipping is seen as an intermediate in the DNA base repair and DNA/RNA base modification pathways, mediated by uracil DNA glycosylase, DNA methyltransferase, and RNA adenosine deaminase, and so forth. These enzymes cause the target base to be flipped out of the duplex, where the energy cost for base flipping is compensated by intercalating the side chain of amino acids and/or forming direct interactions with the flipped-out base [58–60]. Likewise, the base flipping conformation can be prepared using the porphyrin and pyrene-modified nucleosides that compensate for the energy cost by the intercalation of the large stacking group [20, 21, 61]. Unlike these base analogs with a large aromatic group, our base pair analogs possessing a small aromatic group can provide enough interaction energy to stabilize the base flipping conformation with minimized structural disruptions of the double helical structure. Even when T is located opposite to A^X , the intercalation energy for the nonpolar group stacking is greater than the interaction energy for base pairing through interbase hydrogen bonds [41, 54].

It is an interesting finding that the pair-mimic nucleosides can recognize the complementary base under certain conditions, and the conformation changes depending on the interaction energies between the nonpolar group stacking and the base pairing through hydrogen bonds (Figure 7). When the deoxycytidine derivatives were investigated, the thermodynamic stability and the RNA hydrolysis data agreed with the flipping of A, C, T, or U opposite to C^{phe} . However, the base pair through hydrogen bonding is formed between C^{phe} and G by orienting the nonpolar aromatic group into the major groove of the duplex, rather than the guanine flipping conformation [55]. This observation is markedly different from the deoxyadenosine derivatives inability to form the base pair with T and U. The importance of the interaction energy in the conformation of C^{phe} was suggested from the studies using inosine (I) lacking the 2-amino group of G. The inosine base opposite C^{phe} was found to be flipped out. This observation clearly demonstrates that the phenyl group stacking overcomes the base pairing of C^{phe}/I through two hydrogen bonds but not that of C^{phe}/G through three hydrogen bonds (Figure 7(a)). Hence, our base pair analogs can discriminate the bases in accordance with the base pair interaction energy, such as G from I in which the base pair stability with cytosine differed by only about 1 kcal mol^{-1} . In a similar mechanism, recognition of a weakened base pair stability is suggested for the DNA damage searching by a DNA repair protein of human O6-alkylguanine alkyltransferase [60]. Additionally, we have also found that the triphosphate derivative of A^{phe} can be incorporated opposite T in a DNA template by DNA polymerases (manuscript in preparation), suggesting the conformational change in A^{phe} depending on the molecular environment (Figure 7(b)). It was also found that the equilibrium shift to the nonpolar group stacking conformation was more obvious when the phenyl group was replaced by the naphthyl group. We can conclude that the base pair-mimic nucleosides can potentially adopt a dual conformation in the nonpolar group stacking and the base pairing with a complementary base, depending on their interaction energies.

4. Concluding Remarks

This paper describes nucleic acid analogs with the modification and substitution of the purine or pyrimidine base used for the study of nucleic acid interactions. There are also many research studies on the modification at the sugar-phosphate backbone to enhance the hybridization efficiency and to provide greater nuclease resistance for oligonucleotides [62, 63]. For the interaction study, the backbone modification is important to reveal the role of the sugar-phosphate atoms on the cation binding and nucleotide hydration. In comparison to the case of base modifications, molecular design for the backbone modification is more difficult because interactions involved in the backbone atoms are not fully understood.

Particular base analogs with a nonpolar aromatic group can be applied for the interaction study between DNA and

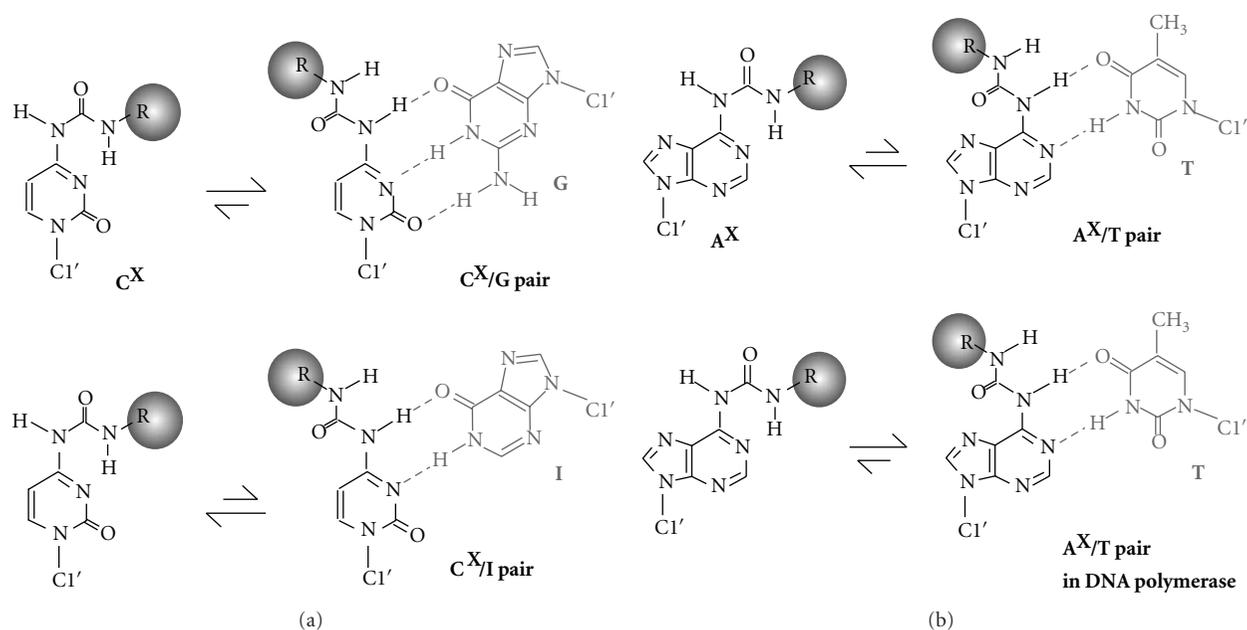


FIGURE 7: The equilibria between the conformations of the nonpolar group stacking and the base pairing of the deoxycytidine (a) and deoxyadenosine derivatives (b), where R indicates the nonpolar aromatic group of the phenyl or naphthyl group.

proteins. For example, nonpolar pyrimidine and purine analogs were used to investigate the base pair geometry in the selection of substrate nucleotides by DNA polymerases [64] and the deleterious effects of eliminating a particular base in a DNA duplex on the interaction with topoisomerases [65]. The pyrene-modified nucleotide in a DNA strand was found to be able to restore the catalytic activity of mutant uracil-DNA glycosylases by assisting the target uracil to be flipped out of the DNA duplex [25, 61]. The base pair analogs shown in Figure 3 are the distinguished Watson-Crick base pair analogs that are accommodated to the DNA duplex structure with minimum disruptions of the conformation and the thermal stability, and they can be used for the study of nucleic acid base interactions such as the base stacking, hydration, and DNA-protein interactions. The base flipping conformation prepared using the base pair-mimic nucleosides is useful to cleave a target RNA sequence and allows evaluation of the dynamics and energetics of the base flipping conformation found in the DNA repair and base-modification proteins and in RNA reactions of the mRNA splicing and ribozyme reactions [66, 67]. The base pair-mimic nucleosides also have an outstanding property to adopt the dual conformation responding to the condition, which is useful to investigate base flipping under the equilibrium with base pairing. Therefore, the molecular design using a flexible linker that tethers a modest stacking group to a purine or pyrimidine base is useful to explore base pair analogs useful for studying the biochemical properties of nucleic acid interactions. Modifications at the aromatic hydrocarbon group and the ureido linker may further expand the application of the base pair analogs such as in molecular biology and develop nucleic acid drugs.

Acknowledgments

The authors thank Yuuki Uotani, Hirohito Oka, Shoji Nakashima, Kazuya Uenishi, and Yosuke Anno for their contributions to the investigations of the base pair-mimic nucleosides. The work was supported in part by Grants-in-Aid for Scientific Research, the “Academic Frontier” Project (2004–2009), and the “Core Research” Project (2009–2014) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and the Hirao Taro Foundation of the Konan University Association for Academic Research.

References

- [1] S. Nakano and N. Sugimoto, “Energy of nucleic acid self-assemblies: from sequence to function through structure,” in *Bottom-Up Nanofabrication: Supramolecules, Self-Assemblies, and Organized Films*, K. Ariga, H. S. Nalwa et al., Eds., chapter 8, pp. 191–215, American Scientific Publishers, 2009.
- [2] D. H. Turner, N. Sugimoto, and S. M. Freier, “RNA structure prediction,” *Annual Review of Biophysics and Biophysical Chemistry*, vol. 17, pp. 167–192, 1988.
- [3] E. Protozanova, P. Yakovchuk, and M. D. Frank-Kamenetskii, “Stacked-unstacked equilibrium at the nick site of DNA,” *Journal of Molecular Biology*, vol. 342, no. 3, pp. 775–785, 2004.
- [4] J. Isaksson and J. Chattopadhyaya, “A uniform mechanism correlating dangling-end stabilization and stacking geometry,” *Biochemistry*, vol. 44, no. 14, pp. 5390–5401, 2005.
- [5] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii, “Base-stacking and base-pairing contributions into thermal stability of the DNA double helix,” *Nucleic Acids Research*, vol. 34, no. 2, pp. 564–574, 2006.
- [6] B. G. Moreira, Y. You, M. A. Behlke, and R. Owczarzy, “Effects of fluorescent dyes, quenchers, and dangling ends on

- DNA duplex stability," *Biochemical and Biophysical Research Communications*, vol. 327, no. 2, pp. 473–484, 2005.
- [7] M. T. Record, W. Zhang, and C. F. Anderson, "Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts," *Advances in Protein Chemistry*, vol. 51, pp. 281–353, 1998.
 - [8] G. G. Privé, K. Yanagi, and R. E. Dickerson, "Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers C-C-A-A-G-A-T-T-G-G and C-C-A-G-G-C-C-T-G-G," *Journal of Molecular Biology*, vol. 217, no. 1, pp. 177–199, 1991.
 - [9] B. Schneider and H. M. Berman, "Hydration of the DNA bases is local," *Biophysical Journal*, vol. 69, no. 6, pp. 2661–2669, 1995.
 - [10] R. I. Kraeva, D. B. Krastev, A. Roguev, A. Ivanova, M. N. Nedelcheva-Velleva, and S. S. Stoyanov, "Stability of mRNA/DNA and DNA/DNA duplexes affects mRNA transcription," *PLoS ONE*, vol. 2, no. 3, Article ID e290, 2007.
 - [11] D. H. Turner, N. Sugimoto, R. Kierzek, and S. D. Dreiker, "Free energy increments for hydrogen bonds in nucleic acid base pairs," *Journal of the American Chemical Society*, vol. 109, no. 12, pp. 3783–3785, 1987.
 - [12] J. SantaLucia, R. Kierzek, and D. H. Turner, "Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin," *Science*, vol. 256, no. 5054, pp. 217–219, 1992.
 - [13] E. M. Moody and P. C. Bevilacqua, "Thermodynamic coupling of the loop and stem in unusually stable DNA hairpins closed by CG base pairs," *Journal of the American Chemical Society*, vol. 125, no. 8, pp. 2032–2033, 2003.
 - [14] E. T. Kool, "Preorganization of DNA: design principles for improving nucleic acid recognition by synthetic oligonucleotides," *Chemical Reviews*, vol. 97, no. 5, pp. 1473–1487, 1997.
 - [15] D. Loakes, "The applications of universal DNA base analogues," *Nucleic Acids Research*, vol. 29, no. 12, pp. 2437–2447, 2001.
 - [16] K. Gao, "Mimicking Watson-Crick base pairs," *Recent Research Developments in Nucleosides & Nucleotides*, vol. 1, pp. 97–110, 2003.
 - [17] Z. Sun and L. W. McLaughlin, "Effects of the minor groove pyrimidine nucleobase functional groups on the stability of duplex DNA: the impact of uncompensated minor groove amino groups," *Biopolymers*, vol. 87, no. 2-3, pp. 183–195, 2007.
 - [18] A. N. Koller, J. Božilović, J. W. Engels, and H. Gohlke, "Aromatic N versus aromatic F: bioisosterism discovered in RNA base pairing interactions leads to a novel class of universal base analogs," *Nucleic Acids Research*, vol. 38, no. 9, pp. 3133–3146, 2010.
 - [19] K. M. Guckian, B. A. Schweitzer, R. X. F. Ren et al., "Experimental measurement of aromatic stacking affinities in the context of duplex DNA," *Journal of the American Chemical Society*, vol. 118, no. 34, pp. 8182–8183, 1996.
 - [20] L. A. Lipscomb, F. X. Zhou, S. R. Presnell et al., "Structure of a DNA-porphyrin complex," *Biochemistry*, vol. 35, no. 9, pp. 2818–2823, 1996.
 - [21] A. David, N. Bleimling, C. Beuck, J. M. Lehn, E. Weinhold, and M. P. Teulade-Fichou, "DNA mismatch-specific base flipping by a bisacridine macrocycle," *ChemBioChem*, vol. 4, no. 12, pp. 1326–1331, 2003.
 - [22] I. Singh, W. Hecker, A. K. Prasad, V. S. Parmar, and O. Seitz, "Local disruption of DNA-base stacking by bulky base surrogates," *Chemical Communications*, no. 5, pp. 500–501, 2002.
 - [23] C. Beuck, I. Singh, A. Bhattacharya et al., "Polycyclic aromatic DNA-base surrogates: high-affinity binding to an adenine-specific base-flipping DNA methyltransferase," *Angewandte Chemie—International Edition*, vol. 42, no. 33, pp. 3958–3960, 2003.
 - [24] J. Tuma, W. H. Connors, D. H. Stitelman, and C. Richert, "On the effect of covalently appended quinolones on termini of DNA duplexes," *Journal of the American Chemical Society*, vol. 124, no. 16, pp. 4236–4246, 2002.
 - [25] Y. L. Jiang, K. Kwon, and J. T. Stivers, "Turning on uracil-DNA glycosylase using a pyrene nucleotide switch," *Journal of Biological Chemistry*, vol. 276, no. 45, pp. 42347–42354, 2001.
 - [26] H. Morales-Rojas and E. T. Kool, "A porphyrin C-nucleoside incorporated into DNA," *Organic Letters*, vol. 4, no. 25, pp. 4377–4380, 2002.
 - [27] B. Bhat and N. J. Leonard, "Dimensional analogue of a dA·dT base pair devoid of propeller twist," *Journal of the American Chemical Society*, vol. 114, no. 19, pp. 7407–7410, 1992.
 - [28] K. Gao and L. E. Orgel, "Nucleic acid duplexes incorporating a dissociable covalent base pair," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 26, pp. 14837–14842, 1999.
 - [29] D. M. Noll, A. M. Noronha, and P. S. Miller, "Synthesis and characterization of DNA duplexes containing an N⁴C-ethyl-N⁴C interstrand cross-link," *Journal of the American Chemical Society*, vol. 123, no. 15, pp. 3405–3411, 2001.
 - [30] C. Dohno, A. Okamoto, and I. Saito, "Stable, specific, and reversible base pairing via schiff base," *Journal of the American Chemical Society*, vol. 127, no. 47, pp. 16681–16684, 2005.
 - [31] S. Atwell, E. Meggers, G. Spraggon, and P. G. Schultz, "Structure of a copper-mediated base pair in DNA," *Journal of the American Chemical Society*, vol. 123, no. 49, pp. 12364–12367, 2001.
 - [32] T. Ono, K. Yoshida, Y. Saotome, R. Sakabe, I. Okamoto, and A. Ono, "Synthesis of covalently linked parallel and antiparallel DNA duplexes containing the metal-mediated base pairs T-Hg(II)-T and C-Ag(I)-C," *Chemical Communications*, vol. 47, no. 5, pp. 1542–1544, 2011.
 - [33] M. Kimoto, T. Mitsui, Y. Harada, A. Sato, S. Yokoyama, and I. Hirao, "Fluorescent probing for RNA molecules by an unnatural base-pair system," *Nucleic Acids Research*, vol. 35, no. 16, pp. 5360–5369, 2007.
 - [34] A. T. Krueger and E. T. Kool, "Redesigning the architecture of the base pair: toward biochemical and biological function of new genetic sets," *Chemistry and Biology*, vol. 16, no. 3, pp. 242–248, 2009.
 - [35] T. Robak, E. Lech-Maranda, A. Korycka, and E. Robak, "Purine nucleoside analogs as immunosuppressive and antineoplastic agents: mechanism of action and clinical activity," *Current Medicinal Chemistry*, vol. 13, no. 26, pp. 3165–3189, 2006.
 - [36] Y. S. Lee, S. M. Park, H. M. Kim et al., "C5-modified nucleosides exhibiting anticancer activity," *Bioorganic and Medicinal Chemistry Letters*, vol. 19, no. 16, pp. 4688–4691, 2009.
 - [37] J. A. McDowell, L. He, X. Chen, and D. H. Turner, "Investigation of the structural basis for thermodynamic stabilities of tandem GU wobble pairs: NMR structures of (rGGAGUCC)₂ and (rGGAUGCC)₂," *Biochemistry*, vol. 36, no. 26, pp. 8030–8038, 1997.
 - [38] K. M. Guckian, B. A. Schweitzer, R. X. F. Ren, C. J. Sheils, D. C. Tahmassebi, and E. T. Kool, "Factors contributing to aromatic

- stacking in water: evaluation in the context of DNA,” *Journal of the American Chemical Society*, vol. 122, no. 10, pp. 2213–2222, 2000.
- [39] J. S. Lai, J. Qu, and E. T. Kool, “Fluorinated DNA bases as probes of electrostatic effects in DNA base stacking,” *Angewandte Chemie—International Edition*, vol. 42, no. 48, pp. 5973–5977, 2003.
- [40] W. K. Tae and E. T. Kool, “A series of nonpolar thymidine analogues of increasing size: DNA base pairing and stacking properties,” *Journal of Organic Chemistry*, vol. 70, no. 6, pp. 2048–2053, 2005.
- [41] S. Nakano, Y. Uotani, K. Uenishi, M. Fujii, and N. Sugimoto, “DNA base flipping by a base pair-mimic nucleoside,” *Nucleic Acids Research*, vol. 33, no. 22, pp. 7111–7119, 2005.
- [42] S. Bommarito, N. Peyret, and J. SantaLucia Jr., “Thermodynamic parameters for DNA sequences with dangling ends,” *Nucleic Acids Research*, vol. 28, no. 9, pp. 1929–1934, 2000.
- [43] T. Ohmichi, S. Nakano, D. Miyoshi, and N. Sugimoto, “Long RNA dangling end has large energetic contribution to duplex stability,” *Journal of the American Chemical Society*, vol. 124, no. 35, pp. 10367–10372, 2002.
- [44] S. Nakano, Y. Uotani, S. Nakashima, Y. Anno, M. Fujii, and N. Sugimoto, “Large stabilization of a DNA duplex by the deoxyadenosine derivatives tethering an aromatic hydrocarbon group,” *Journal of the American Chemical Society*, vol. 125, no. 27, pp. 8086–8087, 2003.
- [45] S. Nakano, H. Oka, Y. Uotani, K. Uenishi, M. Fujii, and N. Sugimoto, “Stacking interaction in the middle and at the end of a DNA helix studied with non-natural nucleotides,” *Molecular BioSystems*, vol. 6, no. 10, pp. 2023–2029, 2010.
- [46] S. Arai, T. Chatake, T. Ohhara et al., “Complicated water orientations in the minor groove of the B-DNA decamer d(CCATTAATGG)₂ observed by neutron diffraction measurements,” *Nucleic Acids Research*, vol. 33, no. 9, pp. 3017–3024, 2005.
- [47] P. Auffinger and E. Westhof, “Water and ion binding around r(UpA) and d(TpA) oligomers—comparison with RNA and DNA (CpG)₁₂ duplexes,” *Journal of Molecular Biology*, vol. 305, no. 5, pp. 1057–1072, 2001.
- [48] R. Brandes, R. R. Void, R. L. Void, and D. R. Kearns, “Effects of hydration on purine motion in solid DNA,” *Biochemistry*, vol. 25, no. 23, pp. 7744–7751, 1986.
- [49] T. V. Chalikian, A. P. Sarvazyan, G. E. Plum, and K. J. Breslauer, “Influence of base composition, base sequence, and duplex structure on DNA hydration: apparent molar volumes and apparent molar adiabatic compressibilities of synthetic and natural DNA duplexes at 25 °C,” *Biochemistry*, vol. 33, no. 9, pp. 2394–2401, 1994.
- [50] M. Feig and B. Montgomery Pettitt, “A molecular simulation picture of DNA hydration around A- And B-DNA,” *Biopolymers*, vol. 48, no. 4, pp. 199–209, 1998.
- [51] N. Sugimoto, M. Nakano, and S. Nakano, “Thermodynamics—structure relationship of single mismatches in RNA/DNA duplexes,” *Biochemistry*, vol. 39, no. 37, pp. 11270–11281, 2000.
- [52] G. A. Soukup and R. R. Breaker, “Relationship between internucleotide linkage geometry and the stability of RNA,” *RNA*, vol. 5, no. 10, pp. 1308–1325, 1999.
- [53] V. Tereshko, S. T. Wallace, N. Usman, F. E. Wincott, and M. Egli, “X-ray crystallographic observation of ”in-line” and ”adjacent” conformations in a bulged self-cleaving RNA/DNA hybrid,” *RNA*, vol. 7, no. 3, pp. 405–420, 2001.
- [54] S. Nakano, Y. Uotani, K. Uenishi, M. Fujii, and N. Sugimoto, “Site-selective RNA cleavage by DNA bearing a base pair-mimic nucleoside,” *Journal of the American Chemical Society*, vol. 127, no. 2, pp. 518–519, 2005.
- [55] S. Nakano, H. Oka, Y. Uotani, K. Uenishi, M. Fujii, and N. Sugimoto, “Dynamics and energetics of the base flipping conformation studied with base pair-mimic nucleosides,” *Biochemistry*, vol. 48, no. 47, pp. 11304–11311, 2009.
- [56] T. Niittymäki and H. Lönnberg, “Artificial ribonucleases,” *Organic and Biomolecular Chemistry*, vol. 4, no. 1, pp. 15–25, 2006.
- [57] A. Kuzuya and M. Komiyama, “Site-selective artificial ribonucleases and their applications,” *Current Organic Chemistry*, vol. 11, no. 16, pp. 1450–1459, 2007.
- [58] R. J. Roberts and X. Cheng, “Base flipping,” *Annual Review of Biochemistry*, vol. 67, pp. 181–198, 1998.
- [59] J. Bischerour and R. Chalmers, “Base flipping in Tn10 transposition: an active flip and capture mechanism,” *PLoS ONE*, vol. 4, no. 7, Article ID e6201, 2009.
- [60] C. G. Yang, K. Garcia, and C. He, “Damage detection and base flipping in direct DNA alkylation repair,” *ChemBioChem*, vol. 10, no. 3, pp. 417–423, 2009.
- [61] K. Kwon, Y. L. Jiang, and J. T. Stivers, “Rational engineering of a DNA glycosylase specific for an unnatural cytosine: pyrene base pair,” *Chemistry and Biology*, vol. 10, no. 4, pp. 351–359, 2003.
- [62] J. Micklefield, “Backbone modification of nucleic acids: synthesis, structure and therapeutic applications,” *Current Medicinal Chemistry*, vol. 8, no. 10, pp. 1157–1179, 2001.
- [63] M. Egli, G. Minasov, V. Tereshko et al., “Probing the influence of stereoelectronic effects on the biophysical properties of oligonucleotides: comprehensive analysis of the RNA affinity, nuclease resistance, and crystal structure of ten 2'-O-ribonucleic acid modifications,” *Biochemistry*, vol. 44, no. 25, pp. 9045–9057, 2005.
- [64] H. Hwang and J. S. Taylor, “Evidence for Watson-Crick and not Hoogsteen or wobble base pairing in the selection of nucleotides for insertion opposite pyrimidines and a thymine dimer by yeast DNA pol η ,” *Biochemistry*, vol. 44, no. 12, pp. 4850–4860, 2005.
- [65] L. Yakovleva, J. Lai, E. T. Kool, and S. Shuman, “Nonpolar nucleobase analogs illuminate requirements for site-specific DNA cleavage by vaccinia topoisomerase,” *Journal of Biological Chemistry*, vol. 281, no. 47, pp. 35914–35921, 2006.
- [66] A. A. Andersen and R. A. Collins, “Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis,” *Molecular Cell*, vol. 5, no. 3, pp. 469–478, 2000.
- [67] N. J. Reiter, H. Blad, F. Abildgaard, and S. E. Butcher, “Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change,” *Biochemistry*, vol. 43, no. 43, pp. 13739–13747, 2004.

Review Article

Radiolytic Reduction Characteristics of Artificial Oligodeoxynucleotides Possessing 2-Oxoalkyl Group or Disulfide Bonds

Kazuhito Tanabe, Takeo Ito, and Sei-ichi Nishimoto

Department of Energy and Hydrocarbon Chemistry, Graduate School of Engineering, Kyoto University, Katsura Campus, Kyoto 615-8510, Japan

Correspondence should be addressed to Kazuhito Tanabe, tanabeka@scl.kyoto-u.ac.jp and Sei-ichi Nishimoto, nishimot@scl.kyoto-u.ac.jp

Received 8 April 2011; Accepted 20 May 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Kazuhito Tanabe et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A number of advances have been made in the development of modified oligodeoxynucleotides (ODNs), and chemical or physical properties of which are controlled by external stimuli. These intelligent ODNs are promising for the next generation of gene diagnostics and therapy. This paper focuses on the molecular design of artificial ODNs that are activated by X-irradiation and their applications to regulation of hybridization properties, conformation change, radiation-activated DNAzyme, and decoy molecules.

1. Introduction

Regulation of chemical or physical properties of oligodeoxynucleotides (ODNs) is important for the development of future gene diagnostics and therapy [1, 2]. Because the function of ODNs is based on their conformation and hybridization properties with their complementary DNA or RNA, various attempts have been made to manipulate these basic characteristics using chemical modification and external stimuli [3–11]. These include the binding of a metal ion to mismatched [3, 4] or modified nucleobases [5, 6], interaction of boron compounds with modified riboses [7, 8], and photochemical methods that use nitrobenzene [9, 10] or azobenzene [11] functionalities on ODNs.

High-energy ionizing radiation is an attractive stimulus for controlling the activity of biomaterials, because the radiation reaction can be controlled spatially and temporally without any additives [12, 13]. In particular, X-ray has potential because it has high live-body permeability, and thus, has been extensively used for medical treatment and diagnosis. In this paper, we describe the current state of research on controlling the function of ODNs by X-irradiation. This paper includes our recent research on the

development of artificial ODNs possessing a 2-oxoalkyl group [14] or disulfide bonds [15, 16], whose properties and conformation can be regulated by X-irradiation. We applied their characteristics to regulation of hybridization, radiation-activated DNAzyme, regulation of the polymerase reaction, and conformation change of ODNs and decoy molecules for inhibition of protein-DNA interactions.

When diluted aqueous solutions are irradiated, practically all of the absorbed energy is deposited in water molecules, and the observed chemical changes are brought about indirectly by the molecular and, in particular, the radical products of water radiolysis. It is well known that ionization and excitation of water molecules by ionizing radiation occur and generate electronically excited states (H_2O^*), radical cations ($\text{H}_2\text{O}^{+\bullet}$), and dry electrons (e_{dry}^-). The excited water molecules H_2O^* dissociate in a homolytic manner to hydrogen atoms (H^\bullet) and hydroxyl radicals (OH^\bullet), while $\text{H}_2\text{O}^{+\bullet}$ deprotonates to OH^\bullet , and e_{dry}^- is solvated to e_{aq}^- . Eventually, e_{aq}^- , OH^\bullet , and H^\bullet are generated as the major active species with G values of 280, 280, and 60 nmol/J, respectively, in the radiolysis of water (Figure 1) [17, 18]. Among the active species, we draw attention to the high reactivity of hydrated electrons and hydrogen atoms, which

reduce a wide range of molecules under hypoxic conditions. We designed artificial ODNs that have high affinity for these reducing species and evaluated their reaction characteristics upon X-irradiation.

2. Design and Reaction Characteristics of 2-Oxoalkyl Caged Oligodeoxynucleotides

Previously, we designed and developed several types of radiation-activated prodrugs, which consisted of a cytotoxic agent and a substituent with high electron affinity [13, 19–22]. The nontoxic or less-toxic prodrugs underwent one-electron reduction by e_{aq}^- to release the active agent to exhibit its inherent cytotoxicity. During the process of developing radiation-activated prodrugs, we identified a series of 2-oxoalkyl groups [19–21] that act as efficiently removable substituents by X-ray treatment under hypoxic conditions in aqueous solution (Figure 2). An activation mechanism has been proposed that the 2-oxoalkyl group undergoes one-electron reduction by e_{aq}^- to form the corresponding π^* anion radical, followed by thermal activation into the σ^* anion radical, which is readily hydrolyzed to release the 2-oxoalkyl group. We have applied these characteristics of the 2-oxoalkyl group to develop several prodrugs of antitumor agents such as 5-fluorouracil and 2'-deoxy-5-fluorouridine (5-FdUrd). These contexts of previous research prompted us to apply 2-oxoalkyl substituents to radiolytic regulation of DNA duplex formation [14].

We synthesized ODNs caged by a 2-oxopropyl group at a given thymine N(3) position ($d^{oxo}T$), which has a characteristic structure similar to the previous radiolytic reduction-activated 5-FdUrd prodrug (Figure 3). One-electron reduction of ODNs possessing $d^{oxo}T$, initiated by hypoxic X-irradiation, occurred to remove exclusively the 2-oxopropyl group and produce the corresponding unmodified ODNs. The unmodified 5–9 mer ODNs were formed from caged ODNs with G values of 100–140 nmol/J for consumption of caged ODNs and 50–110 nmol/J for formation of unmodified ODNs. On the other hand, the radiolytic removal of the 2-oxopropyl group from caged ODNs under aerobic conditions was markedly inefficient, in contrast to the prompt formation of unmodified ODNs under hypoxic conditions. Compared with the noncaged ODN 1a ($T_m = 48.9^\circ C$), the caged ODN 1 ($T_m = 37^\circ C$) showed a dramatic decrease in melting temperature (T_m), indicating that the 2-oxopropyl group was effective in the destabilizing the duplex. To characterize the hybridization properties of caged ODNs, we also subjected a hypoxically irradiated ODN 1 containing $d^{oxo}T$ at the center of the *Swa* I recognition site to enzymatic digestion. After the hypoxic irradiation of ODN 1 and subsequent addition of an equimolar amount of ODN 2, which is complementary to ODN 1, treatment with *Swa* I was conducted. We observed strand cleavage of the resulting duplex at the corresponding restriction site, whereas no strand cleavage occurred when nonirradiated duplex was treated with enzyme. Thus, the ability for ordinary duplex formation was restored by hypoxic X-irradiation of caged ODNs.

Artificial DNA derivatives that can regulate their recognition properties and functions with the aid of external



FIGURE 1: The radical products of water radiolysis.

triggers are useful for a variety of applications. As an example, we tried to make DNAzyme, the function of which can be regulated by hypoxic X-irradiation. We conducted a modification of DNAzyme 17E which cleaves both RNA and DNA/RNA chimeric substrates in the presence of transition metal ions such as Zn^{2+} [23]. We designed DNAzyme 17E possessing $d^{oxo}T$ (ODN 3) at the crotch of the loop region and characterized its cleavage ability using DNA/RNA chimeric substrate (ODN 4), in which the fluorophore and quencher were incorporated into the 5'- and 3'-ends, respectively. The fluorescence emission of the fluorophore was suppressed intramolecularly by the quencher, while the cleavage of ODN 4 led to recovery of emission. Thus, we could monitor the strand cleavage of ODN 4 by fluorescence emission. Caged 33 mer DNAzyme, ODN 3, which was synthesized by automated DNA synthesis, was preirradiated and then incubated in the presence of DNA/RNA for 30 min at $25^\circ C$. As shown in Figure 4, we observed weak fluorescence emission, when ODN 3 without irradiation was incubated with ODN 4, indicating that the cleavage activity of DNAzyme 17E was efficiently reduced by incorporation of $d^{oxo}T$ into the strand. It is striking that the fluorescence intensity from ODN 4 increased with increasing radiation dose up to 40 Gy for caged DNAzyme. These results lead to the conclusion that hypoxic X-irradiation removed the 2-oxopropyl group on the caged DNAzyme, thereby resulting in the formation of active DNAzyme 17E which can cleave the DNA/RNA chimeric substrate.

We also applied a caged ODN to regulation of the DNA polymerase reaction [24]. DNA elongation of primer ODN 5 by the Klenow fragment of DNA polymerase I was regulated by caged ODN 7. Figure 5 shows the protocol and a representative gel picture. When 10 mer caged ODN 7 without X-irradiation was added to the mixture of primer ODN 5, template ODN 6 and enzyme, DNA polymerization occurred to form a 37 mer ODN, which was complementary to ODN 6 with a yield of 96%. This result indicates that caged ODN 7 did not act as a suppressor of enzymatic reaction as ODN 7 could not hybridize with template ODN 6 because of steric hindrance of the 2-oxoalkyl group. In contrast, X-irradiation resulted in removal of the 2-oxoalkyl group from ODN 7, and thereby addition of X-irradiated ODN 7 prevented DNA elongation by duplex formation with template ODN 6. The yield for elongation decreased to 29%. Thus, DNA polymerization has been successfully regulated by X-irradiation.

3. Radiolytic Reduction of Oligodeoxynucleotides Possessing Disulfide Bonds

Disulfides, thiols, and their corresponding radicals have been shown to induce unique reactions, including formation or

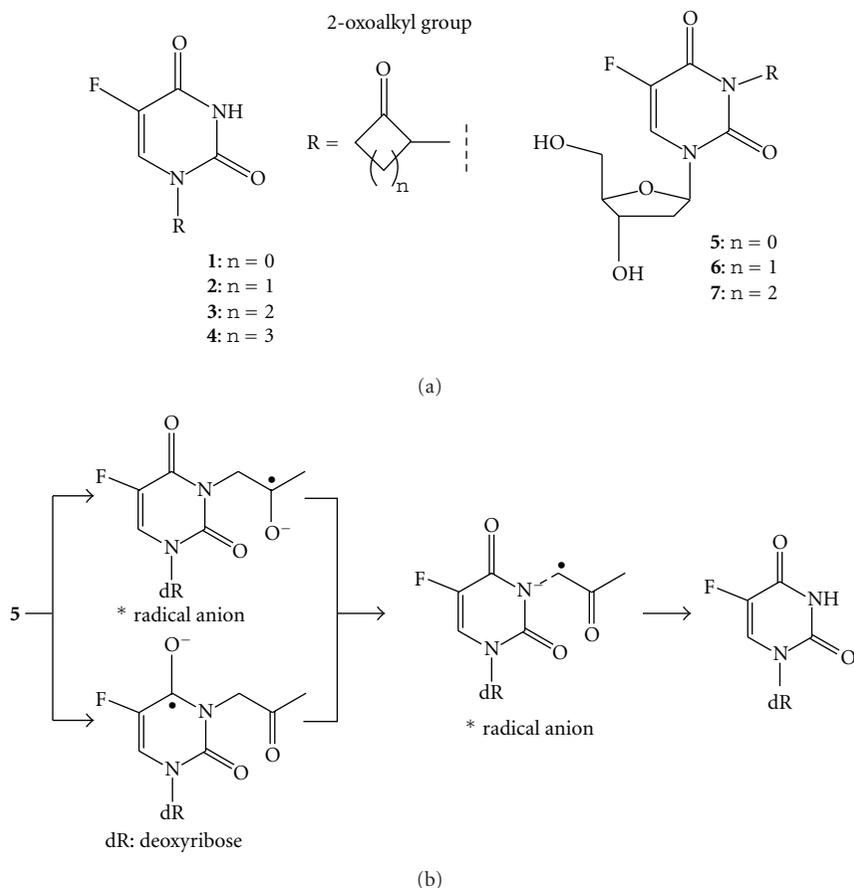


FIGURE 2: (a) Radiation-activated prodrugs of 5-fluorouracil (1-4) and 2'-deoxy-5-fluorouridine (5-7) possessing 2-oxoalkyl groups. (b) Plausible mechanism for reductive release of 2-oxoalkyl group.

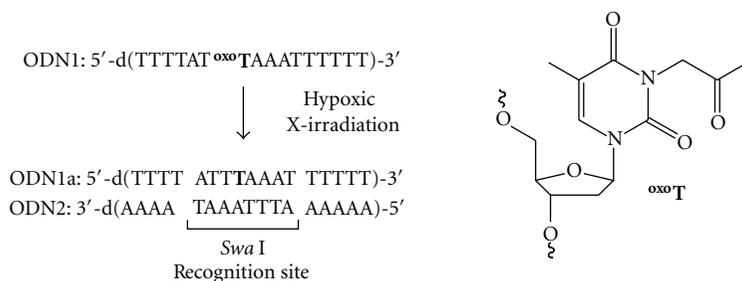


FIGURE 3: Radiolytic activation of 2-oxoalkyl caged oligodeoxynucleotides. Emergence of ordinary duplex formation by hypoxic X-irradiation.

rupture of covalent bonding under reduction conditions. For example, the disulfide radical anion (RSSR^{•-}) generated by one-electron reduction of a disulfide is converted into a sulfide anion (RS⁻) and a thiyl radical (RS[•]) [25, 26]. Sulfide anions tend to undergo disulfide exchange reactions in the presence of disulfide compounds, while thiyl radicals abstract a hydrogen atom to form thiols [27-29]. In addition, hydrogen atoms are assumed to reduce disulfide to form a radical anion [16]. These reaction characteristics motivated us to investigate radiolytic reduction of ODNs possessing disulfide bonds [15, 16].

We initially conducted radiolytic reduction of dinucleotides possessing a disulfide bond (ODN 8) [15]. After the hypoxic irradiation, two ODNs (ODN 9 and ODN 10) were formed as reaction products via an intramolecular strand-crossing reaction concomitantly with the regeneration of ODN 8. Mechanistic studies revealed that this reaction proceeded with multiple turnover process. We subsequently applied these unique reaction characteristics to template-directed strand crossing. In the presence of a complementary ODN (ODN 13), two ODNs possessing a disulfide bond (ODN 11 and ODN 12) produced a specified

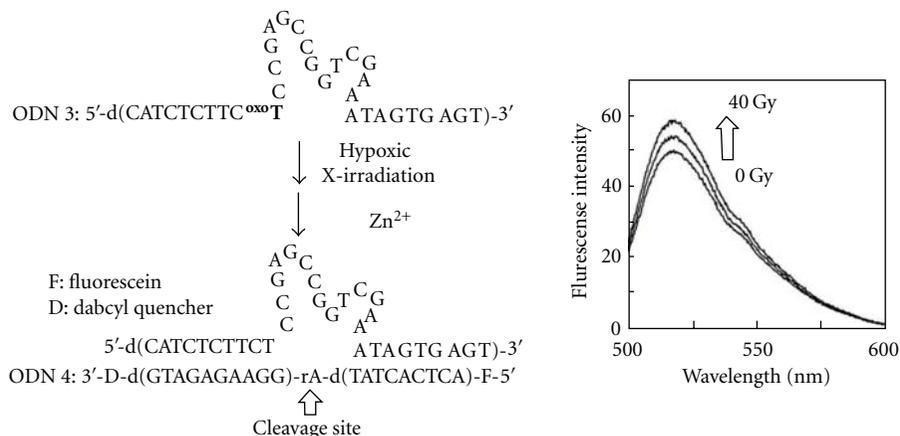


FIGURE 4: Design of caged DNAzyme activated by X-irradiation. (a) Schematic illustration of radiolytic activation of DNAzyme possessing 2-oxoalkyl group. (b) Fluorescence spectra of ODN 4 (500 nM) after addition of Zn²⁺ ion (100 μM) and X-irradiated (0, 20 and 40 Gy) ODN 3 (500 nM) under hypoxic conditions.

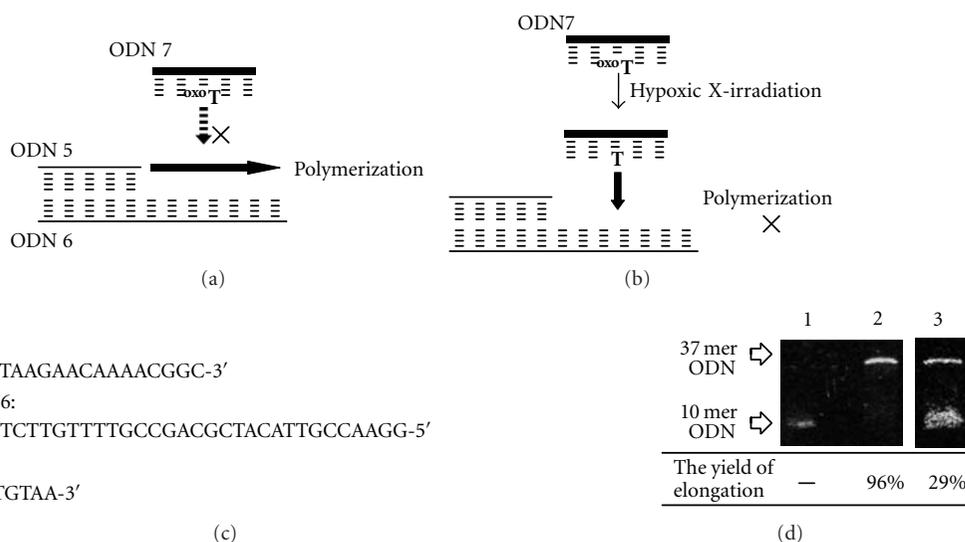


FIGURE 5: (a,b) Schematic illustration of regulation of DNA polymerase reaction by X-irradiation. The reaction in the presence of caged ODN 7 before X-irradiation (a) or after X-irradiation (b). (c) Sequences of oligodeoxynucleotides used for DNA polymerase reaction. (d) Representative gel electrophoresis of DNA elongation by the Klenow fragment of DNA polymerase I. ODN 7 (100 μM) was X-irradiated (1200 Gy) under hypoxic conditions, and then ODN 6 (5 μM), fluorescein-labeled ODN 5 (5 μM), and enzyme (0.2 unit) were added. After the incubation for 15 min at 32°C, the elongation was monitored by gel electrophoresis. Lane 1: marker (20 mer), Lane 2: the ODN 7 was added before X-irradiation. Lane 3: the ODN 7 was added after X-irradiation (1200 Gy).

ODN (ODN 14) via interstrand crossing upon hypoxic irradiation. Radiolytic reduction of ODNs possessing a disulfide bond resulted in the easy preparation of a specified ODN (Figure 6).

As well as these unique reaction characteristics, the disulfide bond plays an important role in the construction of the secondary or tertiary structure of DNA architecture [30–32]. Disulfide bond formation and dissociation are significantly affected by the redox state, and therefore the higher-order structure can be controlled by redox reactions. We next conducted the radiolytic reduction of hairpin-type ODNs possessing two disulfide bonds in order to establish guides for the design of radiation-associated systems to

regulate the DNA higher-order structure. We carried out the radiolytic reduction of 20 mer hairpin-type ODN 15 by the X-radiolysis of an argon-purged aqueous solution in the usual manner. As shown in Figure 7, the corresponding cyclized ODNs, which were identified by HPLC analysis and ESI-TOF mass measurements, were formed with G values of 100 nmol/J for the formation of cyclized ODNs. We also confirmed that the reaction efficiency decreased substantially upon increasing the ODN chain length, whereas it was improved by elevating the reaction temperature.

In a separate experiment, we characterized the effect of reactive species generated by radiolysis of water on the cyclization of ODNs possessing disulfide bonds. When

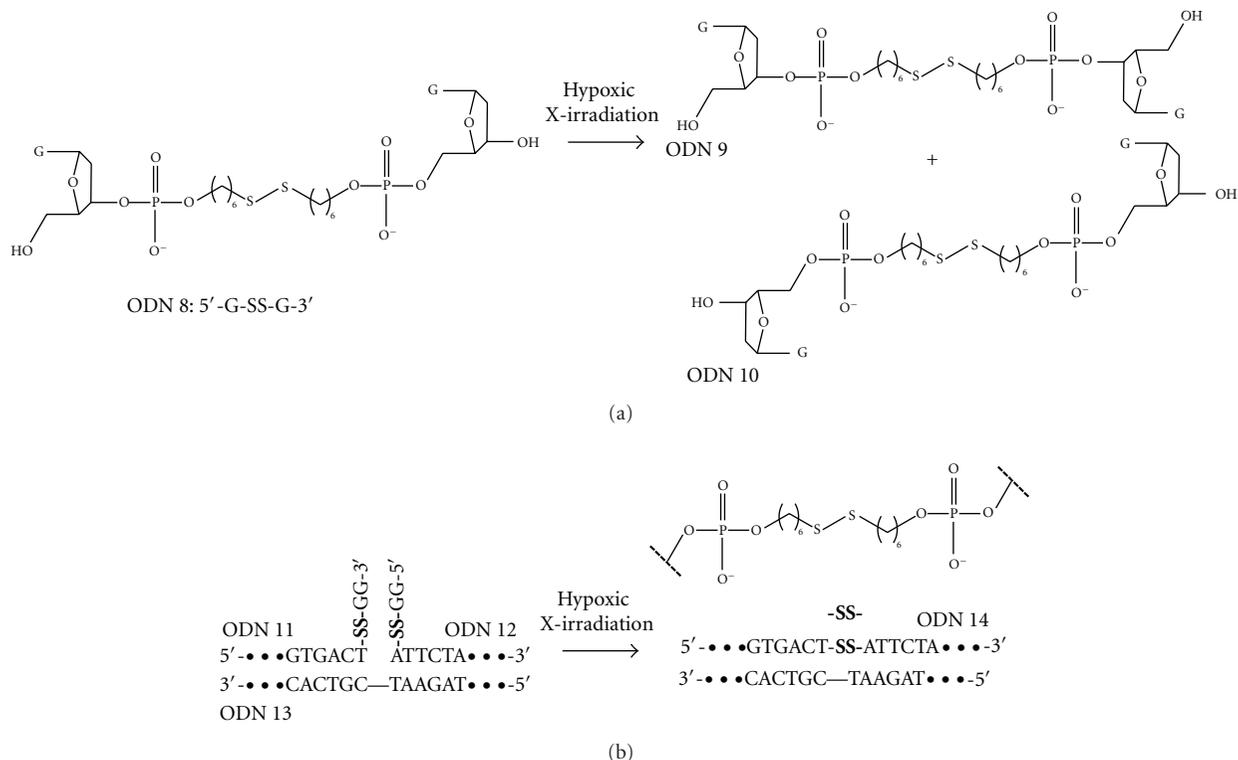


FIGURE 6: (a) Radiolytic strand exchange reaction of dinucleotide unit possessing disulfide bond. (b) Radiolytic template-directed ligation of oligodeoxynucleotides possessing disulfide bond.

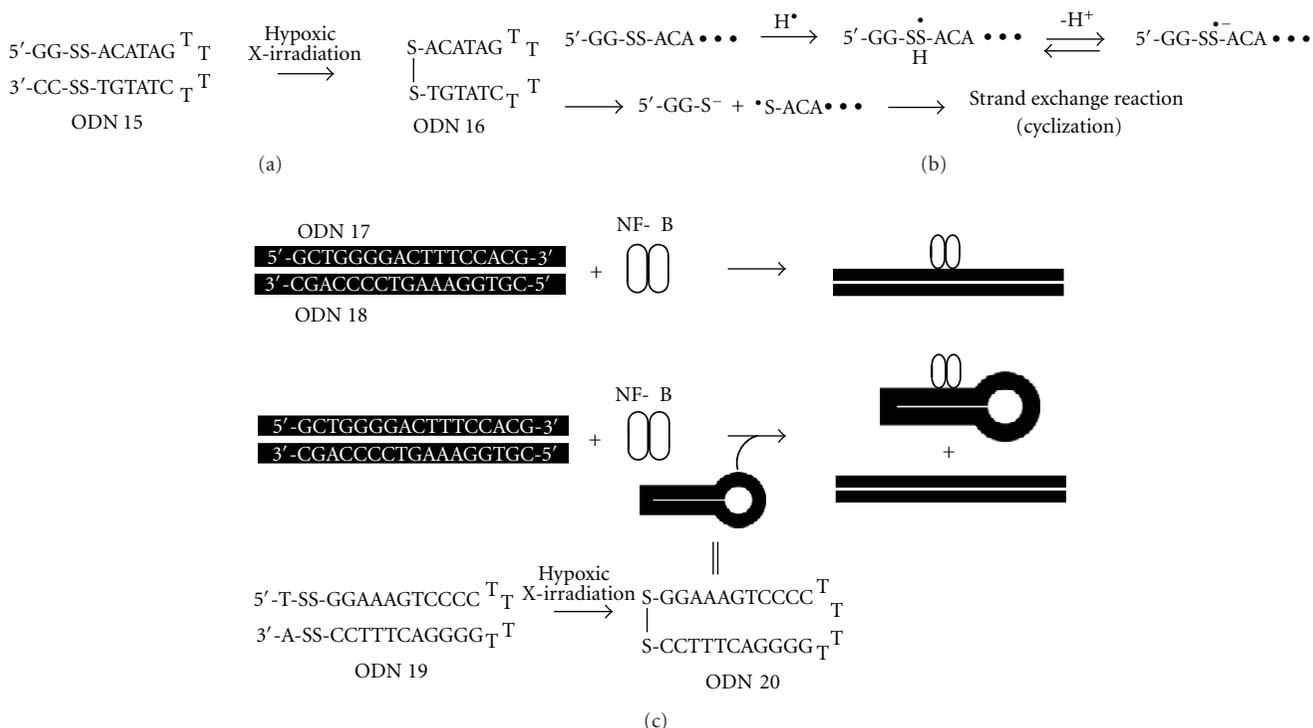


FIGURE 7: (a) Radiolytic cyclization of hairpin-type oligodeoxynucleotides possessing disulfide bonds. (b) The possible reaction mechanism for the cyclization of ODN 15 upon X-ray irradiation. (c) Cyclized oligodeoxynucleotides used for the transcriptional decoy strategy.

hairpin-type ODN 15 was X-irradiated in an aqueous solution containing OH[•]-scavenging 2-methyl-2-propanol purged with nitrous oxide (N₂O) gas [17], which efficiently captures e_{aq}⁻, cyclization was moderately suppressed. On the other hand, the formation of cyclized ODNs was dramatically suppressed in the presence of 2-propanol, which captures both OH[•] and H[•]. These results strongly underline the importance of H[•] as well as e_{aq}⁻ for the activation and strand-crossing reaction of disulfide bonds.

End-capped DNA has recently attracted much attention in the field of gene therapy because of its high stability under biological conditions [33–37]. Further attempts have been made to apply radiolytically cyclized ODN to a decoy strategy directly targeting transcription factors [16]. We prepared a 28 mer hairpin-type ODN 19 possessing two disulfide bonds and characterized its radiolytic reaction, stability, and binding properties against NF-κB. Similar to the results described above, ODN 19 was efficiently cyclized by hypoxic X-irradiation. The cyclized ODN 20 was almost-totally resistant to enzymatic digestion by snake venom phosphodiesterase, while prompt digestion was observed for hairpin-type ODN 19. We next demonstrated that radiolytically cyclized ODN 20 interacts with NF-κB as a decoy by in vitro competition assay. An electrophoretic mobility shift assay revealed that addition of cyclized ODN 20 effectively prevented the binding of NF-κB with a control duplex (ODN 17/ODN 18) in a sequence-selective manner. Thus, the radiolytic formation of cyclized ODNs is promising for a decoy strategy that controls gene expression by inhibition of specific transcription-regulation proteins.

4. Conclusion

Artificial oligonucleotides whose hybridization properties, conformation, and functions can be regulated by external stimuli are applicable to a wide range of gene research areas. From this perspective, we have summarized our recent advances in the development of caged oligonucleotides activated by X-irradiation. Radiolytic reduction by reactive species such as hydrated electrons or hydrogen atoms caused removal of the 2-oxoalkyl group on base units and exchange reaction of the disulfide bond. The hybridization properties and conformation of ODNs possessing these substituents are modulated by X-irradiation, and thereby they show promise as radiation-activated DNAzyme as modulators of the DNA polymerase reaction and as decoy molecules. However, these artificial ODNs are premature for practical use because high radiation doses are required to activate the caged ODNs. Because high doses of X-irradiation may cause damage to cellular DNA, it is difficult to apply the present system to in vivo experiments as it stands. One of the key strategies to overcome this problem is further chemical modification of radiation-activated substituents that have a high affinity for hydrated electrons and hydrogen atoms. Our current interest focuses on the construction of improved caged DNA systems that are highly sensitive to small doses of X-irradiation and their application to in vivo gene regulation.

References

- [1] I. Willner, "Photoswitchable biomaterials: en route to opto-bioelectronic systems," *Accounts of Chemical Research*, vol. 30, no. 9, pp. 347–355, 1997.
- [2] J. F. Milligan, M. D. Matteucci, and J. C. Martin, "Current concepts in antisense drug design," *Journal of Medicinal Chemistry*, vol. 36, no. 14, pp. 1923–1937, 1993.
- [3] A. Ono and H. Togashi, "Highly selective oligonucleotide-based sensor for mercury(II) in aqueous solutions," *Angewandte Chemie-International Edition*, vol. 43, no. 33, pp. 4300–4302, 2004.
- [4] Y. Miyake, H. Togashi, M. Tashiro et al., "Mercury^{II}-mediated formation of thymine-Hg^{II}-thymine base pairs in DNA duplexes," *Journal of the American Chemical Society*, vol. 128, no. 7, pp. 2172–2173, 2006.
- [5] E. Meggers, P. L. Holland, W. B. Tolman, F. E. Romesberg, and P. G. Schultz, "A novel copper-mediated DNA base pair," *Journal of the American Chemical Society*, vol. 122, no. 43, pp. 10714–10715, 2000.
- [6] K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, and M. Shionoya, "A discrete self-assembled metal array in artificial DNA," *Science*, vol. 299, no. 5610, pp. 1212–1213, 2003.
- [7] T. Wada, N. Minamimoto, Y. Inaki, and Y. Inoue, "Peptide ribonucleic acids (PRNA). 2. A novel strategy for active control of DNA recognition through borate ester formation," *Journal of the American Chemical Society*, vol. 122, no. 29, pp. 6900–6910, 2000.
- [8] H. Sato, Y. Hashimoto, T. Wada, and Y. Inoue, "Solid-phase synthesis of peptide ribonucleic acids (PRNA)," *Tetrahedron*, vol. 59, no. 40, pp. 7871–7878, 2003.
- [9] L. Kröck and A. Hechel, "Photoinduced transcription by using temporarily mismatched caged oligonucleotides," *Angewandte Chemie-International Edition*, vol. 44, no. 3, pp. 471–473, 2005.
- [10] G. Mayer and A. Hechel, "Biologically active molecules with a "light switch"," *Angewandte Chemie-International Edition*, vol. 45, no. 30, pp. 4900–4921, 2006.
- [11] H. Asanuma, T. Ito, T. Yoshida, X. Liang, and M. Komiyama, "Photoregulation of the formation and dissociation of a DNA duplex by using the *cis-trans* isomerization of azobenzene," *Angewandte Chemie-International Edition*, vol. 38, no. 16, pp. 2393–2395, 1999.
- [12] K. Tanabe, Z. Zhang, T. Ito, H. Hatta, and S. I. Nishimoto, "Current molecular design of intelligent drugs and imaging probes targeting tumor-specific microenvironments," *Organic and Biomolecular Chemistry*, vol. 5, no. 23, pp. 3745–3757, 2007.
- [13] T. Ito, K. Tanabe, H. Yamada, H. Hatta, and S. I. Nishimoto, "Radiation- and photo-induced activation of 5-fluorouracil prodrugs as a strategy for the selective treatment of solid tumors," *Molecules*, vol. 13, no. 10, pp. 2370–2384, 2008.
- [14] K. Tanabe, H. Kanezaki, H. Ishii, and S. I. Nishimoto, "2-Oxoalkyl caged oligonucleotides: one-electron reductive activation into emergence of ordinary hybridization property by hypoxic X-irradiation," *Organic and Biomolecular Chemistry*, vol. 5, no. 8, pp. 1242–1246, 2007.
- [15] K. Tanabe, E. Kuraseko, Y. Yamamoto, and S. I. Nishimoto, "One-electron reductive template-directed ligation of oligodeoxynucleotides possessing a disulfide bond," *Journal of the American Chemical Society*, vol. 130, no. 20, pp. 6302–6303, 2008.

- [16] K. Tanabe, E. Matsumoto, T. Ito, and S. I. Nishimoto, "Radiolytic cyclization of stem-and-loop structured oligodeoxynucleotide with neighboring arrangement of α,ω -bis-disulfides," *Organic and Biomolecular Chemistry*, vol. 8, no. 21, pp. 4837–4842, 2010.
- [17] G. V. Buxton, C. L. Greenstock, W. P. Helman, and A. B. Ross, "Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals ($\cdot\text{OH}/\cdot\text{O}^-$ in aqueous solution)," *Journal of Physical and Chemical Reference Data*, vol. 17, no. 2, pp. 513–886, 1988.
- [18] J. W. T. Spinks and R. J. Woods, *An Introduction to Radiation Chemistry*, Wiley-Interscience, New York, NY, USA, 1990.
- [19] M. Mori, H. Hatta, and S. I. Nishimoto, "Stereo-electronic effect on one-electron reductive release of 5-Fluorouracil from 5-fluoro-1-(2'-oxocycloalkyl)uracils as a new class of radiation-activated antitumor prodrugs," *Journal of Organic Chemistry*, vol. 65, no. 15, pp. 4641–4647, 2000.
- [20] Y. Shibamoto, L. Zhou, H. Hatta, M. Mori, and S. I. Nishimoto, "In vivo evaluation of a novel antitumor prodrug, 1-(2'-oxopropyl)-5-fluorouracil (OFU001), which releases 5-fluorouracil upon hypoxic irradiation," *International Journal of Radiation Oncology Biology Physics*, vol. 49, no. 2, pp. 407–413, 2001.
- [21] K. Tanabe, Y. Mimasu, A. Eto et al., "One-electron reduction characteristics of N(3)-substituted 5-fluorodeoxyuridines synthesized as radiation-activated prodrugs," *Bioorganic and Medicinal Chemistry*, vol. 11, no. 21, pp. 4551–4556, 2003.
- [22] K. Tanabe, Y. Makimura, Y. Tachi, A. Imagawa-Sato, and S. I. Nishimoto, "Hypoxia-selective activation of 5-fluorodeoxyuridine prodrug possessing indolequinone structure: radiolytic reduction and cytotoxicity characteristics," *Bioorganic and Medicinal Chemistry Letters*, vol. 15, no. 9, pp. 2321–2324, 2005.
- [23] J. Li, W. Zheng, A. H. Kwon, and Y. Lu, "In vitro selection and characterization of a highly efficient Zn(II)-dependent RNA-cleaving deoxyribozyme," *Nucleic Acids Research*, vol. 28, no. 2, pp. 481–488, 2000.
- [24] A. Yamazawa, X. Liang, H. Asanuma, and M. Komiyama, "Photoregulation of the DNA polymerase reaction by oligonucleotides bearing an azobenzene," *Angewandte Chemie-International Edition*, vol. 39, no. 13, pp. 2356–2357, 2000.
- [25] S. Antonello, R. Benassi, G. Gavioli, F. Taddei, and F. Maran, "Theoretical and electrochemical analysis of dissociative electron transfers proceeding through formation of loose radical anion species: reduction of symmetrical and unsymmetrical disulfides," *Journal of the American Chemical Society*, vol. 124, no. 25, pp. 7529–7538, 2002.
- [26] M. Z. Hoffman and E. Hayon, "One-electron reduction of the disulfide linkage in aqueous solution. Formation, protonation, and decay kinetics of the RSSR- radical," *Journal of the American Chemical Society*, vol. 94, no. 23, pp. 7950–7957, 1972.
- [27] N. K. Mathur, C. K. Narang, and R. E. Williams, *Polymers as Aids in Organic Chemistry*, Academic Press, New York, NY, USA, 1980.
- [28] M. Ishiguro, *SHki no Kagaku Shuushoku*, Japan Scientific Societies Press, Tokyo, Japan, 1978.
- [29] A. Wójcik, S. Naumov, B. Marciniak, and O. Brede, "Repair reactions of pyrimidine-derived radicals by aliphatic thiols," *Journal of Physical Chemistry B*, vol. 110, no. 25, pp. 12738–12748, 2006.
- [30] S. E. Osborne, J. Volker, S. Y. Stevens, K. J. Breslauer, and G. D. Glick, "Design, synthesis, and analysis of disulfide cross-linked DNA duplexes," *Journal of the American Chemical Society*, vol. 118, no. 48, pp. 11993–12003, 1996.
- [31] A. A. Gorodetsky and J. K. Barton, "DNA-mediated electrochemistry of disulfides on graphite," *Journal of the American Chemical Society*, vol. 129, no. 19, pp. 6074–6075, 2007.
- [32] M. Endo and T. Majima, "Structural arrangement of two DNA double helices using cross-linked oligonucleotide connectors," *Chemical Communications*, no. 11, pp. 1308–1309, 2004.
- [33] G. N. Parkinson, M. P. H. Lee, and S. Neidle, "Crystal structure of parallel quadruplexes from human telomeric DNA," *Nature*, vol. 417, no. 6891, pp. 876–880, 2002.
- [34] S. M. Gowan, J. R. Harrison, L. Patterson et al., "A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity," *Molecular Pharmacology*, vol. 61, no. 5, pp. 1154–1162, 2002.
- [35] M. Read, R. J. Harrison, B. Romagnoli et al., "Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 9, pp. 4844–4849, 2001.
- [36] M. Nakane, S. Ichikawa, and A. Matsuda, "Triazole-linked dumbbell oligodeoxynucleotides with NF- κ B binding ability as potential decoy molecules," *Journal of Organic Chemistry*, vol. 73, no. 5, pp. 1842–1851, 2008.
- [37] T. Hosoya, H. Takeuchi, Y. Kanesaka et al., "Sequence-specific inhibition of a transcription factor by circular dumbbell DNA oligonucleotides," *FEBS Letters*, vol. 461, no. 3, pp. 136–140, 1999.

Research Article

Coupling Strategies for the Synthesis of Peptide-Oligonucleotide Conjugates for Patterned Synthetic Biomineralization

Joshua D. Carter¹ and Thomas H. LaBean^{1,2}

¹Departments of Computer Science and Chemistry, Duke University, Durham, NC 27708, USA

²Department of Biomedical Engineering, Duke University Pratt School of Engineering, Durham, NC 27708, USA

Correspondence should be addressed to Thomas H. LaBean, thomas.labean@duke.edu

Received 17 May 2011; Accepted 25 June 2011

Academic Editor: F. C. Simmel

Copyright © 2011 J. D. Carter and T. H. LaBean. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work describes preparation strategies for peptide-oligonucleotide conjugates that combine the self-assembling behavior of DNA oligonucleotides with the molecular recognition capabilities of peptides. The syntheses include a solution-phase fragment coupling reaction and a solid-phase fragment coupling strategy where the oligonucleotide has been immobilized on DEAE Sepharose. The yield of four coupling reagents is evaluated, two reagents in water, EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) and DMTMM (4-(4,6-dimethoxy[1,3,5]triazin-2-yl)-4-methyl-morpholinium chloride), and two in dimethylformamide (DMF), PyBOP ((Benzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate) and HBTU (*O*-benzotriazole-*N,N,N'*-tetramethyluronium hexafluorophosphate), while the oligonucleotide fragment is either in solution or immobilized on DEAE. These coupling strategies rely on an unprotected 5' amino linker on the oligonucleotide reacting with the peptide C-terminus. The peptide, selected from a combinatorial library for its gold-binding behavior, was 12 amino acids long with an N-terminus acetyl cap. Formation of the conjugates was confirmed by gel electrophoresis and mass spectrometry while molecular recognition functionality of the peptide portion was verified using atomic force microscopy. Solution-phase yields were superior to their solid-phase counterparts. EDC resulted in the highest yield for both solution-phase (95%) and solid-phase strategies (24%), while the DMF-based reagents, PyBOP and HBTU, resulted in low yields with reduced recovery. All recoverable conjugates demonstrated gold nanoparticle templating capability.

1. Introduction

Development of oligonucleotide-based therapies began thirty years ago with the demonstration of gene regulation using synthetic oligonucleotides [1, 2]. Since that discovery, an increasing number of functional groups have been conjugated to oligonucleotides in efforts to enhance specific behaviors. Peptides, in particular, are an increasingly popular choice for conjugation with oligonucleotides because of their ability to aid in targeting specific biological functions while improving permeability through the cell membrane and stability against intracellular degradation. Peptide-oligonucleotide conjugates (POCs) have been used to target mRNA, double-stranded DNA, and proteins, in antisense, triple-helix, and aptamer-based therapy, respectively [3–5]. Our group has recently explored another use of POCs by

incorporating peptides used in artificial biomineralization [6–8] into self-assembling DNA nanostructures [9–11] for the demonstration of controlled biomineralization [12]. While the range of proposed applications for POCs continues to widen, scientists have relied on two principal strategies for joining the increasingly diverse peptide and oligonucleotide sequences together: in-line synthesis and fragment conjugation.

In-line synthesis typically involves the step-by-step synthesis of an oligomer covalently linked to a solid support and culminates with a linker through which the monomer-by-monomer assembly of the second oligomer can take place. The use of a solid-support during in-line synthesis of POCs greatly reduces the complexity of purification and can also be useful for the production of POC libraries. Furthermore, immobilizing the oligomers on solid-phase

eliminates the potential for differential solubility issues when attempting to synthesize a water-soluble oligonucleotide with a hydrophobic peptide component. However, difficulties in the production of POCs can arise due to the need for complex protecting group strategies to avoid damaging the first oligomer during synthesis of the second. This may lead to restrictions on sequence choices [13]. Moreover, each successive monomer addition results in a reduction of the overall yield making the production of long POCs difficult or impossible [14].

Fragment conjugation, in contrast to in-line synthesis, employs a strategy where the oligonucleotide and peptide are synthesized, deprotected, and purified separately, followed by a final step in which covalent linkage of the two pure fragments occurs. Although this strategy requires additional purification steps, it allows for the use of standard protecting group strategies in the production of the peptide and oligonucleotide fragments. The peptide does not need to be stable during oligonucleotide synthesis, and *vice versa*. Fragment conjugation also allows for the storage of the respective oligomer fragments for later use in multiple conjugate syntheses. Although fragment-coupling is rather straightforward, the relative solubilities of the oligomer fragments can cause problems. Oligonucleotides are water-soluble, while some peptides may be hydrophobic and require the use of organic solvents for coupling. One possible solution to this problem is the use of solid-phase fragment conjugation. In this strategy, the fragments are covalently linked while one oligomer is immobilized on solid-phase media.

Here, we describe both a solution-phase fragment coupling reaction and a nearly identical solid-phase fragment coupling where the oligonucleotide fragment has been temporarily immobilized on ion exchange resin via electrostatic interactions. Immobilizing the oligonucleotide fragment to a solid support allows us to consider only the solubility requirements of the peptide fragment when choosing a reaction solvent. The yields of both the solution-phase and solid-supported fragment coupling reactions were measured with two water-soluble coupling reagents, EDC and DMTMM, and two coupling reagents in dimethylformamide, PyBOP and HBTU. The coupling reaction relies on the formation of a peptide bond between an unprotected oligonucleotide with a 5' amino linker and the C-terminus of a 12 amino acid gold-binding peptide with an acetyl-capped N-terminus [15]. Characterization of the synthetic strategies using polyacrylamide gel electrophoresis and mass spectrometry will be presented, the effect of reaction solvent and coupling reagent will be discussed, and the choice of solution-phase or solid-phase media will be evaluated. Finally, incorporation of the purified POC into a self-assembling DNA lattice allows for atomic force microscope evaluation of the gold-binding capability of the free-peptide end.

2. Methods and Materials

2.1. Purification of Oligonucleotide. 15AmMC6/TTG TGA AGT TTT TCG ATC CTA GCA CCT CTG GAG TTT TTC TTG CC. Synthetic oligonucleotide with a 5'-Amino Modifier C6 functional group was purchased from Integrated

DNA Technologies (Coralville, IA) and separated from truncation products by polyacrylamide gel electrophoresis. Denaturing polyacrylamide gel electrophoresis (PAGE) was carried out on a 160 × 180 × 1.5 mm gel containing 10% acrylamide and 8.3 M urea in TBE buffer (90 mM Tris, 90 mM boric acid, and 2 mM ethylenediaminetetraacetate, pH 8). The gel loading buffer contained 90% formamide and 0.1% bromophenol blue. After electrophoresis (300 V for 45 minutes), the portion of the gel containing the desired DNA was excised, diced, and shaken overnight at 4°C in 500 μ L 0.5 M ammonium acetate, 10 mM magnesium acetate, and 2 mM ethylenediaminetetraacetate. The supernatant was removed and added to a centrifuge tube with 1 mL of 100% (v/v) ethanol and stored at -20°C overnight. The mixture was centrifuged for 30 minutes at 4°C and 16,000 g and the supernatant was discarded. The purified pellet of DNA was dissolved in pure water and then desalted using a Pierce spin column (#69705) containing Sephadex G-25 Medium (Amersham Biosciences). The concentration was determined by ultraviolet absorption at 260 nm wavelength.

2.2. Peptide Synthesis and Purification. The gold-binding peptide [15] (Ac-WALRRSIRRQSY-OH) was synthesized on Wang resin preloaded with Fmoc-L-Tyrosine (Novabiochem) at 0.1 mmol scale using a Protein Technologies PS3 automated peptide synthesizer. The coupling of standard Fmoc (9-fluorenylmethoxy-carbonyl)-protected amino acids (Chem-Impex) was achieved with HBTU (O-benzotriazole-*N,N,N'*-tetramethyluronium hexafluorophosphate; Novabiochem) in the presence of *N*-methylmorpholine (NMM) in *N,N'*-dimethylformamide (DMF) for 20-minute cycles. Fmoc deprotection was achieved using 20% piperidine in DMF (2 × 5 minutes). The N-terminus of the peptide was acetylated with acetic anhydride and NMM. Side-chain deprotection and peptide cleavage from the resin were achieved by treating the resin-bound peptide with 5 mL of 100% trifluoroacetic acid (TFA) for 2 hours under N₂. After evaporation of TFA under N₂, the peptide was washed three times with cold diethyl ether, air-dried, and purified by semipreparative reverse-phase HPLC on a YMC C18 column with a linear 40-minute gradient from 3 to 70% acetonitrile in water with 0.1% TFA. The mass of the peptide was confirmed using an Agilent ESI-MS.

2.3. Reaction Solvents in Peptide-Oligonucleotide Synthesis. Fragment-coupling reactions using PyBOP ((Benzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate; Aldrich) and HBTU were run completely in DMF using DNA immobilized on DEAE Sepharose media (DFF-100, Sigma). Reactions with water-soluble coupling reagents EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) and DMTMM (4-(4,6-dimethoxy[1,3,5]triazin-2-yl)-4-methyl-morpholinium chloride) were completed using both Sepharose media and in solution-phase. DMTMM-based coupling reactions were run in 200 mM MOPS buffer pH 7.0 (200 mM (3[N-morpholino]propanesulfonic acid, 20 mM sodium acetate, 10 mM EDTA) while EDC-based reactions were run in water.

TABLE 1: Reaction yields for POC formation in the listed conditions.

Reaction	Coupling reagent	Solvent	Supporting media	Temperature	Reaction time	Yield (%)
1	HBTU	DMF	Sepharose	room temperature	16 h	0
2	PyBOP	DMF	Sepharose	room temperature	16 h	4
3	DMTMM	Water	Sepharose	room temperature	16 h	15
4	EDC	Water	Sepharose	room temperature	16 h	24
5	DMTMM	Water	Sepharose	50°C	16 h	15
6	DMTMM	Water	Sepharose	room temperature	2 h	15
7	PyBOP	DMF	Sepharose	room temperature	2 h	4
8	DMTMM	Water	Solution	room temperature	16 h	50
9	EDC	Water	Solution	room temperature	16 h	95
10	DMTMM	Water	Solution	50°C	16 h	55

2.4. General Procedure for Coupling Reactions. All reactions were run in a constant volume in the solvent preferred by the coupling reagent. For DMTMM-coupled reactions, 450 pmoles of DNA was combined with peptide (270 nmoles, 600 eq.) and DMTMM (4.5 μ moles, 10,000 eq.). EDC-based coupling reactions followed the same procedure as above using EDC (4.5 μ moles, 10,000 eq) and N-hydroxy succinimide (450 nmoles, 1000 eq) in place of DMTMM. PyBOP and HBTU-based coupling reactions were accomplished in DMF with PyBOP or HBTU (4.5 μ moles, 10,000 eq.) and diisopropylethylamine (9.0 μ moles, 20,000 eq.).

2.5. Workup for Solution-Phase Coupling Reactions. The peptide-oligonucleotide conjugates were isolated from the completed solution-phase coupling reactions via ethanol precipitation. In this case, 1 mL of 100% (v/v) ethanol and 50 μ L of 3 M sodium acetate (150 μ moles) were added and allowed to sit overnight at -20°C . The precipitated conjugate was centrifuged for 30 minutes at 4°C and 16,000 g and the supernatant was discarded. The pellet was dried in a vacuum centrifuge for 2 hours before resuspension in 30 μ L water and analysis by denaturing PAGE (300 V for 45 minutes).

2.6. General Procedure for Solid-Phase Coupling Reactions. The procedure for immobilization of DNA on Sepharose was adapted from Halpin et al. [16]. Briefly, 250 μ L of Sepharose slurry was added to a Pierce spin column, promptly washed with three column volumes of DNA-loading buffer (10 mM Acetic Acid and 0.005% Triton X-100), and briefly centrifuged (one minute at 3,000 g). To bind the DNA on the Sepharose, 450 pmoles DNA (250 μ L loading buffer) was incubated with the Sepharose resin for ten minutes and centrifuged for one minute at 3,000 g. The resin was then washed with one column volume of the appropriate reaction solvent (DMF, buffer, or water) before addition of the solution containing coupling reagent and peptide to the immobilized DNA. The resulting mixture was allowed to react for the duration listed in Table 1. After the reaction, the exhausted reagents were removed using a centrifuge (1 minute at 3,000 g) and the column was washed with one column volume of DNA loading buffer. The peptide-oligonucleotide conjugate was eluted with $2 \times 200 \mu\text{L}$ elution buffer (1.5 M NaCl and 50 mM Tris-HCl

pH 8.0) and the combined fractions were added to 1 mL of 100% (v/v) ethanol and allowed to sit overnight at -20°C . The precipitated DNA was centrifuged for 30 minutes at 4°C and 16,000 g and the supernatant was discarded. The pellet was dried in a vacuum centrifuge for 2 hours before resuspension in 30 μ L water and analysis using denaturing PAGE (300 V for 45 minutes).

2.7. Denaturing PAGE Analysis of the Reaction Products. To determine the yield of a coupling reaction, 15 μ L of the DNA recovered from the reaction was added to a 10% denaturing polyacrylamide gel and run at 300 V for 45 minutes. The gel containing the contents of the peptide-oligonucleotide conjugate reaction was stained with ethidium bromide and imaged with a UV transilluminator. Reaction yields were determined by comparing fluorescent signals from unreacted oligonucleotide (faster moving band) to coupled peptide-oligonucleotide conjugate (slower moving band) using the UV transilluminator software (AlphaImager HP from AlphaInnotech).

2.8. MALDI Analysis of Peptide-Oligonucleotide Conjugates. MALDI-TOF mass spectrometry analysis was used to characterize the oligonucleotide starting material and peptide-oligonucleotide reaction product. The analysis was performed using an Applied Biosystems DE-Pro MalDI-MS in the Mass Spectrometry Facility in the Chemistry Department at Duke University. To prepare the samples, the products were recovered from the polyacrylamide gels (vide supra) and dissolved in pure water to a concentration of 30 μM . 10 μ L volumes of the recovered gel products were then stripped of cations using Ziptips (SCX, Millipore) and added to a mixture of 9 μ L of 50 mg/mL 3-hydroxypicolinic acid and 1 μ L 50 mg/mL diammonium citrate. The mass spectrometer was run in negative-ion mode and spectra were collected through the summing of 50 laser pulses.

2.9. DNA Strands. The self-assembling DNA lattice used in these experiments is described elsewhere [12].

2.10. DNA Nanostructure Formation. For DNA lattice formation, 18 individual PAGE-purified DNA oligos and the

peptide-oligo conjugate (POC) were mixed together stoichiometrically at $1.0 \mu\text{M}$ in $1 \times \text{TAE/Mg}^{2+}$ buffer (40 mM Tris-HCl (pH 8.0), 20 mM acetic acid, 2 mM EDTA, and 12.5 mM magnesium acetate) and slowly cooled from 95 to 20°C over a period of 16 hours. For AFM imaging, $3 \mu\text{L}$ of sample was spotted on freshly cleaved mica for 3 minutes. A $30 \mu\text{L}$ portion of $1 \times \text{TAE/Mg}^{2+}$ buffer was then placed onto the mica and another $30 \mu\text{L}$ of $1 \times \text{TAE/Mg}^{2+}$ buffer was placed onto the AFM tip (for a total of $60 \mu\text{L}$). AFM images were obtained on a Digital Instruments Nanoscope IIIa with a multimode head by tapping mode under buffer using NP-S tips (Veeco Inc.).

2.11. Gold Nanoparticle Preparation. Gold nanoparticles (AuNP, 5 nm) were purchased from Ted Pella (product no. 15702-20). To prevent nanoparticle aggregation in the high salt environment preferred by the DNA lattice, nanoparticles were pretreated with Tween 20 (Sigma Aldrich product no. P1379). 1 mL of stock AuNP solution was mixed with the appropriate amount of Tween 20 stabilizing agent to achieve a final Tween concentration of 0.2% (w/v). After addition of the Tween 20, the AuNP solutions were mixed overnight at room temperature. After incubation, the AuNP particles were concentrated by 6 hours of centrifugation at 16000 g. The resulting AuNP pellet was then isolated using a pipet and resuspended in $10 \mu\text{L}$ of supernatant. The concentrations of the AuNP were determined using the AuNP absorbance at 520 nm (ϵ 520 nm for 5 nm AuNP = $1 \times 10^7 \text{ M}^{-1} \text{ cm}^{-1}$) [17].

2.12. AFM Imaging of AuNP-Labeled DNA Lattice. AuNP-labeled DNA lattice was prepared by adding $6 \mu\text{L}$ annealed lattice solution to concentrated AuNP in $1 \times \text{TAE/Mg}^{2+}$ buffer. four equivalents of AuNP to one equivalent of POC were used in all cases. All samples combining DNA with AuNP were allowed to mix in solution for 10 minutes prior to deposition of $3 \mu\text{L}$ of the mixture on mica, followed by a 3 minute wait before the addition of $60 \mu\text{L}$ of buffer for tapping mode AFM analysis. AFM images were obtained on a Digital Instruments Nanoscope IIIa with a multimode head by tapping mode under buffer using NP-S tips (Veeco Inc.).

3. Results and Discussion

3.1. Peptide Synthesis and Purification. A recently published peptide sequence (Ac-WALRRSIRRQSY-OH) was selected to serve as the test molecule for making peptide-oligonucleotide conjugates [15]. The peptide, known for its ability to bind gold surfaces, was synthesized according to the procedure detailed in the methods and materials. Briefly, Fmoc coupling protocols were used to add amino acid residues to Wang resin preloaded with Fmoc-L-Tyrosine. The N-terminus of the peptide was acetylated to prevent the primary amine from participating in the coupling reaction with DNA. The final product was purified by semipreparative HPLC and analysis showed a purity greater than 95%. ESI-MS measurements of the final product gave an m/z of 1634.2

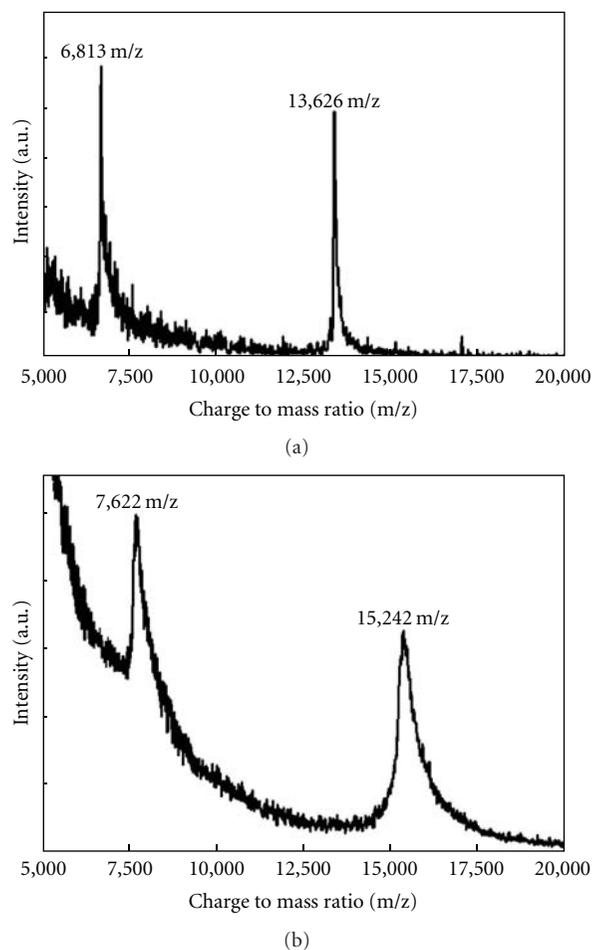


FIGURE 1: MALDI-TOF analysis of pure oligonucleotide starting material (a) and peptide-oligonucleotide conjugate (b).

($M + H^+$) which compared well with the calculated mass of 1633.9 g/mol.

3.2. Oligonucleotide Purification. The oligonucleotide sequence (TTG TGA AGT TTT TCG ATC CTA GCA CCT CTG GAG TTT TTC TTG CC) used in the test couplings was selected from a DNA nanostructure previously constructed in our lab [9]. Prior to coupling, the commercially synthesized oligonucleotide, containing a primary amine on its 5' end, was dissolved in pure water and purified using PAGE. After purification, the recovered DNA pellet was dissolved in pure water and desalted using size exclusion chromatography to produce the purified oligonucleotide in 30% yield. Incomplete removal of the salts remaining from DNA synthesis and purification was found to have an adverse effect on the DNA affinity for the Sepharose reaction media (data not shown).

3.3. MALDI Analysis of Peptide-Oligonucleotide Conjugates. Figure 1 shows the spectra obtained for pure oligonucleotide starting material and the peptide-oligonucleotide conjugate. As expected for the starting material, the measured mass

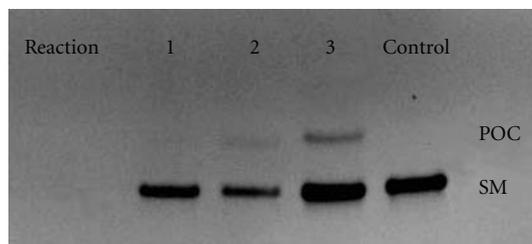


FIGURE 2: PAGE analysis of reactions 1–3 from Table 1. The coupling reagents used in lanes 1, 2, and 3 are HBTU, PyBOP, and DMTMM, respectively. As demonstrated by the control lane, starting material (SM) is below the slower-moving peptide-oligonucleotide conjugate (POC).

of 13,626 m/z corresponds well with the calculated theoretical mass of 13,630.9 g/mol . Similarly for the peptide oligonucleotide conjugate, the observed mass of 15,242 m/z corresponds well with the calculated theoretical mass of 15,246.8 g/mol .

3.4. Effects of Reaction Variables on Coupling Efficiency.

The coupling reagents and reaction variables tested for this report are detailed in Table 1. Conditions varied for this study include coupling reagent and solvent, reaction temperature, and reaction duration. Table 1 also delineates reactions that occurred on the Sepharose media and those that took place in solution. Reactions were evaluated based on the yield of peptide-oligonucleotide conjugate obtained for each coupling reagent. The reaction yield was determined by integrating the fluorescent signal from ethidium bromide-stained PAGE and comparing between unreacted oligonucleotide and newly formed (and slower moving) peptide-oligonucleotide conjugate. The reaction yield, or the amount of product obtained for each reaction, should be distinguished from the total recovery of DNA, the sum of product and unreacted oligonucleotide.

Reactions 1–4 listed in Table 1 compare the coupling efficiency of four coupling reagents at room temperature over 16 hours while the DNA is confined to the Sepharose surface. Under these conditions, EDC was the most effective coupling reagent and produced the highest reaction yield of 24%. At the other extreme, PyBOP and HBTU produced little or no product. PyBOP and HBTU also resulted in a reduced total recovery of the DNA from the Sepharose. This effect, demonstrated in Figure 2 (lanes 1 and 2) as a relatively weak fluorescent signal from the total DNA present in each lane, is in agreement with the findings of other researchers [16]. Aqueous washes of the Sepharose columns were used after DNA loading when coupling with either EDC or DMTMM and DMF washes were used after DNA loading in cases where PyBOP or HBTU was to be the coupling reagent. Although enhanced DNA solubility in the DMF wash solvent might have been suspected for the poor recovery of DNA from these latter reactions, PAGE analysis of DMF wash fractions collected from the Sepharose column demonstrated the absence of prematurely eluted DNA. This indicates irreversible binding between the DNA

and the Sepharose for PyBOP and HBTU-based reactions in DMF, but this could not be directly measured.

To test the effect of temperature on yield, reaction 3 (DMTMM, Sepharose) was repeated at a higher temperature (Reaction 5). As shown in Table 1, increasing the temperature to 50°C resulted in no increase in the yield for the DMTMM coupling reagent on Sepharose over 16 hours. Reaction 3 was repeated again on a shorter time scale (Reaction 6, Table 1) producing little effect on the yield indicating that the formation of peptide-oligonucleotide product is essentially complete within two-hour reaction time. Previous research into formation rates of peptide-oligoconjugates has demonstrated nearly complete reactions in as few as twenty minutes [18]. Although the quick completion of the reaction could be due to decay of the DMTMM through demethylation or hydrolysis, other groups have found the decay in this short time frame to be insignificant [19, 20]. The effect of decreasing reaction duration on coupling yield was repeated with PyBOP (Reaction 7), and as in the case of DMTMM, very little change in yield was observed. Reactions 8 and 9 are the solution-phase equivalent to the solid-phase reactions 3 and 4, respectively. As the data illustrate, dissolving the DNA in solution rather than binding it to Sepharose led to a dramatic increase in the reaction yield. One possible explanation for this observation is that the DNA conformation resulting from the electrostatic interactions between the Sepharose and the DNA could be adversely affecting the progress of the reaction. These interactions may prevent the amine group from reacting with the carboxylic acid of the peptide [21]. In the case of solution-phase coupling, DMTMM produced 50% product while EDC went nearly to completion (95%). The effect of temperature on DMTMM coupling was again tested in reaction 10, and in this case, a small improvement in the yield was observed.

3.5. Gold-Binding Functionality of the POCs after Incorporation into DNA Lattice.

POC preparation reactions that produced yields of at least 15% allowed for the purification and isolation of POC material (Table 1). To test the gold-binding capability of these POC materials, they were incorporated into the modified DNA lattice prior to addition of 5 nm AuNP and analysis with AFM. Figure 3 contains typical AFM images of both unmodified DNA lattice and DNA lattice incorporating POC after addition of 5 nm AuNP while Figure 4 shows the corresponding line profiles for both samples. AFM analysis of DNA lattices incorporating purified POC species isolated from the various preparation reactions listed in Table 1 showed no discernable difference in the areal density of AuNP reached with identical reaction times and AuNP concentrations. The steady gold-binding functionality of the purified POCs synthesized by the various means leads to two possible conclusions, either the peptide consistently formed a covalent linkage to the oligonucleotide as designed through the 5' amino linker on the oligonucleotide and the C-terminus of the peptide, or the gold-binding functionality of the peptide is not constrained by the location of the covalent linkage between oligonucleotide and peptide. Due to the uncertainty surrounding the gold-binding mechanism of the peptide either of these

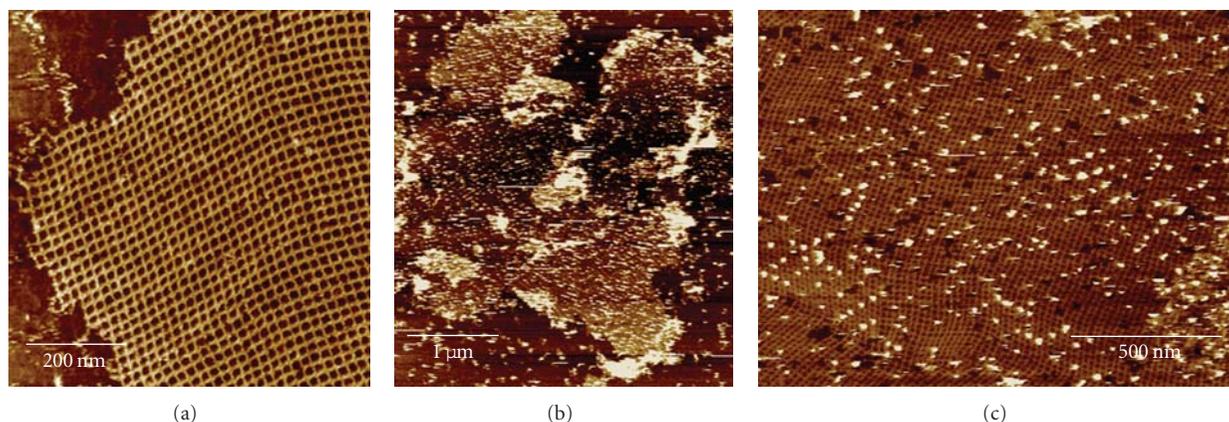


FIGURE 3: AFM images of a DNA lattice (a) and a DNA lattice incorporating POC isolated from reaction 3 after addition of 5 nm gold nanoparticles (b) and (c).

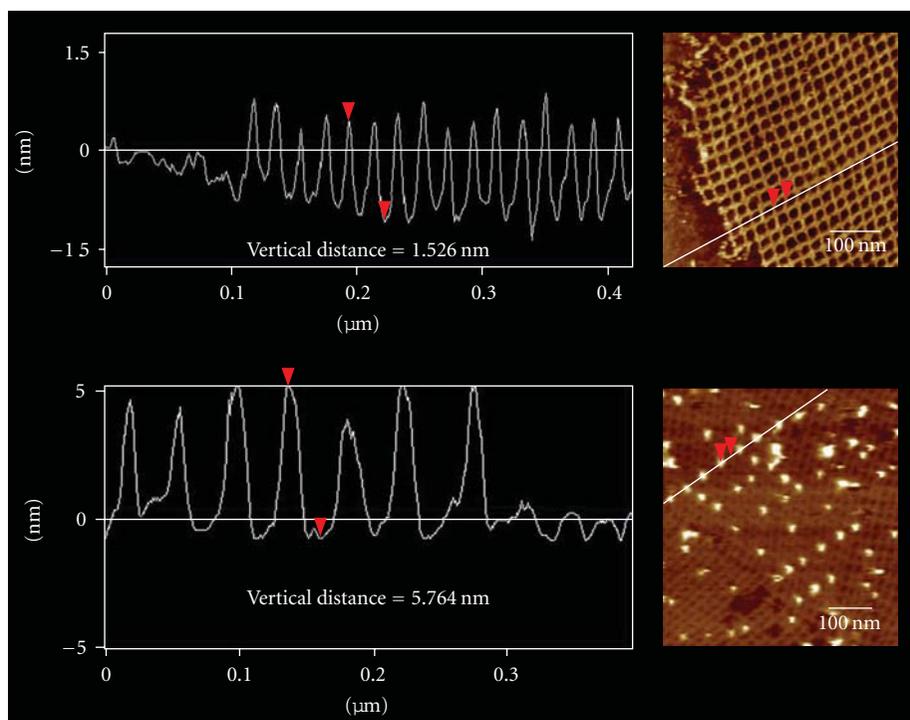


FIGURE 4: AFM profile scans of a DNA lattice (top) and a DNA lattice incorporating POC isolated from reaction 3 after addition of 5 nm gold nanoparticles (bottom).

scenarios remains plausible. Although higher areal densities have been achieved by using AuNP labeled with single-stranded DNA complementary to strands incorporated into DNA lattice [22], the simultaneous organization of multiple inorganic species favors an approach with greater selectivity, such as those exhibited by proteins in biological chemistry. Furthermore, experimental evidence indicated an inverse correlation between the number of AuNP equivalents added and the binding time required to reach near-saturating conditions. This suggests that low AuNP areal densities can be improved with increased reaction time and additional AuNP equivalents.

4. Conclusion

The work presented herein describes multiple strategies for the preparation of POCs that combine the capabilities of both self-assembling DNA oligonucleotides with the molecular-recognition capabilities of peptides. Four coupling reagents have been evaluated for their ability to form peptide oligonucleotide conjugates while the oligonucleotide is either in solution-phase or restricted to a solid-support. The molecular recognition capability of the peptide ends of the POC products was further evaluated using 5 nm AuNP after their incorporation into a self-assembling DNA

lattice. The various purified POCs, when incorporated into DNA lattices and allowed to mix with AuNP for equivalent durations, resulted in equivalent areal densities of AuNP when visualized using AFM. In the preparation of the POCs, EDC consistently produced the highest yield while water-soluble coupling reagents as a whole produced markedly higher yields compared to their hydrophobic counterparts. Solution-phase coupling yields were superior to their solid-phase alternative. DMF-based reactions resulted in reduced recovery of coupled product from the solid-phase resin. Furthermore, when choosing between DMF and water for the formation of peptide-oligonucleotide conjugates, the superior yields of EDC and DMTMM demonstrated here indicate water to be the superior solvent for coupling. However, researchers attempting to form peptide-oligonucleotide conjugates in the future may choose solvents other than water in order to balance starting material solubility with reaction yield. It is worth noting that although the results in this report were obtained using a water-soluble gold-binding peptide, the coupling yield of hydrophobic peptides targeting other species may be different under otherwise identical conditions.

In the characterization of peptide functionality after anchoring to a DNA oligonucleotide, all reactions from which viable yields of POC could be isolated demonstrated consistent nanoparticle templating capability. Although not tested for this report, there are studies concluding that the length of the amine linker may also affect the yield of the coupling reactions [18, 23]. This effect could provide researchers with additional means of optimizing the yield of peptide-oligoconjugate reactions. Another factor not tested for this report was the scalability of the coupling reactions. The manufacturer-published exchange capacity for DEAE Sepharose gel of 110–160 $\mu\text{eq/mL}$ means that the saturation point for the DNA used in this report is at least three orders of magnitude higher, allowing for the potential use of larger reaction scales.

Acknowledgments

The authors thank Katherine Franz for the use of the peptide synthesizer, Kurt Gothelf and Niels Vinther Voigt for experimental advice, and George Dubay for assistance with the MALDI measurements. The authors acknowledge funding for this project from NIH training Grant T32-EB001630 (JDC) and NSF Grants CCF-SGER-0650083, DMR-BMAT-0706397, and CCF-EMT-0829749.

References

- [1] M. L. Stephenson and P. C. Zamecnik, "Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 75, no. 1, pp. 285–288, 1978.
- [2] P. C. Zamecnik and M. L. Stephenson, "Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxyribonucleotide," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 75, no. 1, pp. 280–284, 1978.
- [3] J. Oehlke, G. Wallukat, Y. Wolf et al., "Enhancement of intracellular concentration and biological activity of PNA after conjugation with a cell-penetrating synthetic model peptide," *European Journal of Biochemistry*, vol. 271, no. 14, pp. 3043–3049, 2004.
- [4] D. S. Wilson and J. W. Szostak, "In vitro selection of functional nucleic acids," *Annual Review of Biochemistry*, vol. 68, pp. 611–647, 1999.
- [5] O. A. Sedelnikova, V. N. Karamychev, I. G. Panyutin, and R. D. Neumann, "Sequence-specific gene cleavage in intact mammalian cells by ^{125}I -labeled triplex-forming oligonucleotides conjugated with nuclear localization signal peptide," *Antisense and Nucleic Acid Drug Development*, vol. 12, no. 1, pp. 43–49, 2002.
- [6] M. Sarikaya, C. Tamerler, A. K. Y. Jen, K. Schulten, and F. Baneyx, "Molecular biomimetics: nanotechnology through biology," *Nature Materials*, vol. 2, no. 9, pp. 577–585, 2003.
- [7] J. N. Cha, G. D. Stucky, D. E. Morse, and T. J. Deming, "Biomimetic synthesis of ordered silica structures mediated by block copolypeptides," *Nature*, vol. 403, no. 6767, pp. 289–292, 2000.
- [8] E. C. Samano, M. Pilo-Pais, S. Goldberg, B. N. Vogen, G. Finkelstein, and T. H. LaBean, "Self-assembling DNA templates for programmed artificial biomineralization," *Soft Matter*, vol. 7, no. 7, pp. 3240–3245, 2011.
- [9] H. Yan, S. H. Park, G. Finkelstein, J. H. Reif, and T. H. LaBean, "DNA-templated self-assembly of protein arrays and highly conductive nanowires," *Science*, vol. 301, no. 5641, pp. 1882–1884, 2003.
- [10] H. Li, J. D. Carter, and T. H. LaBean, "Nanofabrication by DNA self-assembly," *Materials Today*, vol. 12, no. 5, pp. 20–28, 2009.
- [11] N. C. Seeman, "DNA in a material world," *Nature*, vol. 421, no. 6921, pp. 427–431, 2003.
- [12] J. D. Carter and T. H. LaBean, "Organization of inorganic nanomaterials via programmable DNA self-assembly and peptide molecular recognition," *ACS Nano*, vol. 5, no. 3, pp. 2200–2205, 2011.
- [13] C. H. Tung and S. Stein, "Preparation and applications of peptide-oligonucleotide conjugates," *Bioconjugate Chemistry*, vol. 11, no. 5, pp. 605–618, 2000.
- [14] N. Venkatesan and B. H. Kim, "Peptide conjugates of oligonucleotides: synthesis and applications," *Chemical Reviews*, vol. 106, no. 9, pp. 3712–3761, 2006.
- [15] M. Hnilova, E. E. Oren, U. O. S. Seker et al., "Effect of molecular conformations on the adsorption behavior of gold-binding peptides," *Langmuir*, vol. 24, no. 21, pp. 12440–12445, 2008.
- [16] D. R. Halpin, J. A. Lee, S. J. Wrenn, and P. B. Harbury, "DNA display III. Solid-phase organic synthesis on unprotected DNA," *PLoS Biology*, vol. 2, no. 7, p. E175, 2004.
- [17] T. A. Taton, "Preparation of gold nanoparticle-DNA conjugates," *Current Protocols in Nucleic Acid Chemistry*, vol. 12, no. 2, p. 11, 2002.
- [18] S. M. Viladkar, "Guanine rich oligonucleotide-amino acid/peptide conjugates: preparation and characterization," *Tetrahedron*, vol. 58, no. 3, pp. 495–502, 2002.
- [19] K. Hojo, M. Maeda, N. Tanakamaru, K. Mochida, and K. Kawasaki, "Solid phase peptide synthesis in water VI: evaluation of water-soluble coupling reagents for solid phase peptide synthesis in aqueous media," *Protein and Peptide Letters*, vol. 13, no. 2, pp. 189–192, 2006.

- [20] M. Kunishima, C. Kawachi, J. Morita, K. Terao, F. Iwasaki, and S. Tani, "4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride: an efficient condensing agent leading to the formation of amides and esters," *Tetrahedron*, vol. 55, no. 46, pp. 13159–13170, 1999.
- [21] D. C. Harris, X. Chu, and J. Jayawickramarajah, "DNA-small molecule chimera with responsive protein-binding ability," *Journal of the American Chemical Society*, vol. 130, no. 45, pp. 14950–14951, 2008.
- [22] J. Zhang, Y. Liu, Y. Ke, and H. Yan, "Periodic square-like gold nanoparticle arrays templated by self-assembled 2D DNA nanogrids on a surface," *Nano Letters*, vol. 6, no. 2, pp. 248–251, 2006.
- [23] S. Peyrottes, B. Mestre, F. Burlina, and M. J. Gait, "The synthesis of peptide-oligonucleotide conjugates by a fragment coupling approach," *Tetrahedron*, vol. 54, no. 41, pp. 12513–12522, 1998.

Research Article

The Effects of Molecular Crowding on the Structure and Stability of G-Quadruplexes with an Abasic Site

Takeshi Fujimoto,¹ Shu-ichi Nakano,^{1,2} Daisuke Miyoshi,^{1,2} and Naoki Sugimoto^{1,2}

¹ Faculty of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

² Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

Correspondence should be addressed to Daisuke Miyoshi, miyoshi@center.konan-u.ac.jp and Naoki Sugimoto, sugimoto@konan-u.ac.jp

Received 20 May 2011; Revised 23 June 2011; Accepted 24 June 2011

Academic Editor: Luis A. Marky

Copyright © 2011 Takeshi Fujimoto et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Both cellular environmental factors and chemical modifications critically affect the properties of nucleic acids. However, the structure and stability of DNA containing abasic sites under cell-mimicking molecular crowding conditions remain unclear. Here, we investigated the molecular crowding effects on the structure and stability of the G-quadruplexes including a single abasic site. Structural analysis by circular dichroism showed that molecular crowding by PEG200 did not affect the topology of the G-quadruplex structure with or without an abasic site. Thermodynamic analysis further demonstrated that the degree of stabilization of the G-quadruplex by molecular crowding decreased with substitution of an abasic site for a single guanine. Notably, we found that the molecular crowding effects on the enthalpy change for G-quadruplex formation had a linear relationship with the abasic site effects depending on its position. These results are useful for predicting the structure and stability of G-quadruplexes with abasic sites in the cell-mimicking conditions.

1. Introduction

Biomolecules have evolved to function within living cells, which contain a variety of macromolecules including nucleic acids, proteins, polysaccharides, and metabolites. These molecules make the intracellular environment extremely crowded; 20–40% of the total volume is physically occupied by biomolecules [1–5]. It has been reported that molecular crowding is a critical factor determining the structure, stability and function of proteins and nucleic acids [6–18]. Because of this importance, molecular crowding effects on the structure and stability of DNA duplexes, triplexes, G-quadruplexes, and other structures have been studied [9–18]. These studies have demonstrated that the molecular crowding effects depend on the patterns of base-base hydrogen bonding in the nucleic acid structure. The stability of DNA duplexes and triplexes comprised of Watson-Crick base pairs decreases by molecular crowding [11–13]. In contrast,

molecular crowding stabilizes DNA triplexes and quadruplexes formed by Hoogsteen base pairs [13, 14, 17]. These results suggest that the formation of noncanonical DNA structures such as triplexes and quadruplexes is induced under the molecular crowding conditions. In addition, a recent study has shown that the RNA cleavage activity of a ribozyme composed of noncanonical base-pairs and tertiary interactions was enhanced by molecular crowding [18]. These results indicate that noncanonical structures of nucleic acids can be stabilized by molecular crowding, leading to a polymorphic nature of function and structure in living cells. In fact, it is now generally accepted that noncanonical structures of nucleic acids, especially G-quadruplexes, play important roles in various biological systems [19–21].

G-quadruplexes are formed by intermolecular or intramolecular associations of guanine-rich sequences which formed four Hoogsteen base paired coplanar guanines, a structure called a G-quartet plane [22]. Bioinformatic studies

have demonstrated that there are more than 370,000 guanine-rich sequences in the human genome [19, 23]. Importantly, these guanine-rich sequences are enriched in important genomic regions involving oncogene promoter regions, short and long minisatellite repeats, ribosomal DNAs, immunoglobulin heavy chain switch regions, and telomere regions [23–29]. As seen in telomere DNA, various biological events related to these guanine-rich sequences should be regulated by the formation of G-quadruplexes, resulting in the DNA G-quadruplex as an important target for cancer therapies [30–33]. Notably, it has been reported that the function of telomere DNA is also affected by DNA lesions [34]. The most common lesions of genomic DNA are the abasic site (apurinic/aprimidinic (AP) site), mainly formed by hydrolysis of *N*-glycosidic bond of nucleotides [35–37]. Recently, the effects of abasic sites on the structure and stability of G-quadruplexes were reported independently by Esposito et al. and Školáková et al. [38, 39]. They each found that solution component, as well as the position of the abasic site, was important for the structure and stability of G-quadruplexes. However, the effects of abasic site on the structure and stability of G-quadruplexes under molecular crowding conditions remain unclear.

In the present study, we examined the molecular crowding effects on the structure and stability of G-quadruplexes including a single abasic site. We found that the degree of stabilization of the G-quadruplex by molecular crowding decreased by the substituting of an abasic site for a single guanine. Moreover, the molecular crowding effects on the thermodynamics of G-quadruplexes with abasic sites depended on the position of the abasic site. By comparing the molecular crowding effects and the abasic site effects on the enthalpy change during the formation of G-quadruplexes, we found that the molecular crowding effects on the enthalpy change decreased linearly as the abasic site effects increased. These results demonstrated that the effect of molecular crowding on the thermodynamics of G-quadruplexes is determined by the enthalpy change during formation, which depends on the position of the abasic site. These results are useful to predict the structure and stability of G-quadruplexes with abasic sites in cell-mimicking conditions.

2. Materials and Methods

2.1. Materials. DNA oligonucleotides with an abasic site were synthesized by a normal solid-phase synthetic procedure. For the abasic site, the dSpacer CE Phosphoramidite (5'-O-Dimethoxytrityl-1',2'-Dideoxyribose-3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite) (GLEN, Sterling, VA) was used. After purification using a C18 column and a polyacrylamide gel (acrylamide:bis (acrylamide) = 19:1) and electrophoresis in 7 M urea, the purity of the DNA oligonucleotides was confirmed to be ~95%. DNA oligonucleotide without abasic sites was high-performance liquid chromatography (HPLC) grade and was purchased from Hokkaido System Science (Sapporo, Japan). Single-strand concentrations of DNA oligonucleotides were determined by measuring the absorbance at 260 nm at a high

temperature using a Shimadzu 1700 spectrophotometer (Kyoto, Japan) connected to a thermo programmer. Single-strand extinction coefficients were calculated from mononucleotide and dinucleotide data using the nearest-neighbour approximation [40, 41].

2.2. Structural Analysis. Circular dichroism (CD) spectra utilizing a JASCO J-820 spectropolarimeter (Hachioji, Japan) were measured at 4°C in a 0.1-cm-path-length cuvette for 20 μM total strand concentration of DNA in a buffer of 100 mM NaCl, 50 mM Tris-HCl (pH 7.0), and 1 mM Na₂EDTA at 0 wt% (0 mol L⁻¹) and 40 wt% (2 mol L⁻¹) PEG200 (poly(ethylene glycol) with an average molecular weight of 200). The CD spectra were obtained by taking the average of at least three scans made from 200 to 350 nm. The temperature of the cell holder was regulated by a JASCO PTC-348 temperature controller, and the cuvette-holding chamber was flushed with a constant stream of dry N₂ gas to avoid condensation of water on the cuvette exterior. Before the measurement, the sample was heated to 80°C, gently cooled at a rate of 2°C min⁻¹, and incubated 1 h at 4°C.

2.3. Thermodynamic Analysis. The UV absorbance of the DNA oligonucleotides with were measured by a Shimadzu 1700 spectrophotometer (Kyoto, Japan) equipped with a temperature controller. The UV melting curves of G-quadruplexes were measured at 295 nm where G-quadruplexes showed hypochromic transition [42]. All experiments were carried out in a buffer of 100 mM NaCl, 50 mM Tris-HCl (pH 7.0), and 1 mM Na₂EDTA at 0 wt% (0 mol L⁻¹) and 40 wt% (2 mol L⁻¹) PEG200. The heating rate was 0.5°C min⁻¹. The thermodynamic parameters (ΔH° , ΔS° , ΔG_{37}°) were calculated from the fit of the melting curves (with at least four different concentrations of DNA) to a theoretical equation for a two-state model of intramolecular association as described previously [41]. The thermodynamic parameters are the average values obtained from the curve-fitting analysis. Before the measurement, the sample was heated to 80°C, gently cooled at a rate of 2°C min⁻¹, and incubated 1 h at 0°C.

2.4. Water Activity Measurement. The water activity was determined by the osmotic stressing method via vapor phase osmometry using a model 5520XR pressure osmometer (Wescor, Utah, USA.). In order to calculate the water activity with various concentrations of PEG200, we assumed that PEG200 did not interact directly with DNA [11, 12].

3. Results and Discussion

3.1. Sequence Design. We used the human telomere DNA sequence (dA(G₃T₂A)₃G₃) because previous studies have reported that this sequence folds into an intramolecular antiparallel G-quadruplex in the presence of Na⁺ (Figure 1(a)) [43]. In order to understand the effects of molecular crowding on the structure and stability of G-quadruplexes with abasic sites, we designed and synthesized human telomere DNA sequences with a single abasic site instead of a guanine base. The DNA sequences and their abbreviations used in

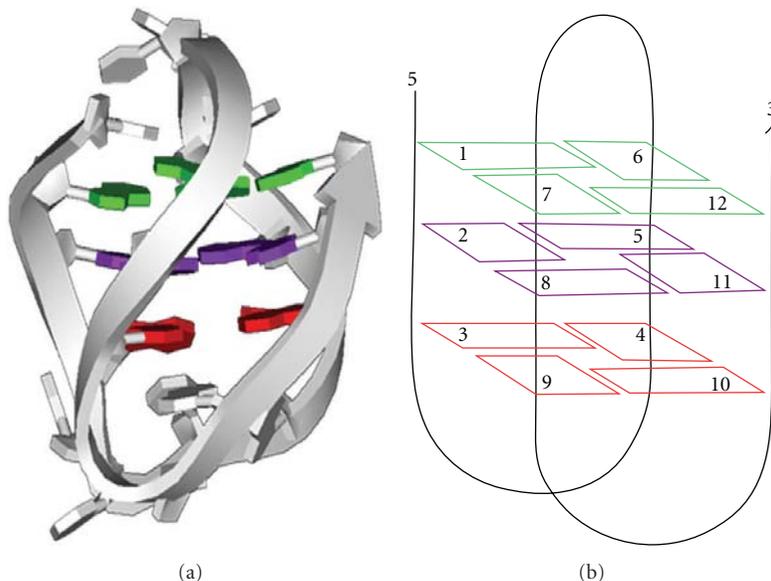


FIGURE 1: (a) Structure of human telomere (5'-A(G₃T₂A)₃G₅-3') G-quadruplex in the presence of Na⁺ (PDB ID: 143D). (b) Schematic illustration of G-quadruplex AP-0 with a top G-quartet plane (green), middle G-quartet plane (purple), and bottom G-quartet plane (red). The numbers indicate the position of abasic sites (AP-X (X indicates 1 to 11)).

TABLE 1: DNA sequences used in this study.

Abbreviation	Sequence
AP-0	5'-AGGGTTAGGGTTAGGGTTAGGG-3'
AP-1	5'-A \emptyset GGTTAGGGTTAGGGTTAGGG-3'
AP-2	5'-AG \emptyset GTTAGGGTTAGGGTTAGGG-3'
AP-3	5'-AGG \emptyset TTAGGGTTAGGGTTAGGG-3'
AP-4	5'-AGGGTTA \emptyset GGTTAGGGTTAGGG-3'
AP-5	5'-AGGGTTAG \emptyset GTTAGGGTTAGGG-3'
AP-6	5'-AGGGTTAGG \emptyset TTAGGGTTAGGG-3'
AP-7	5'-AGGGTTAGGGTTA \emptyset GGTTAGGG-3'
AP-8	5'-AGGGTTAGGGTTAG \emptyset GTTAGGG-3'
AP-9	5'-AGGGTTAGGGTTAGG \emptyset TTAGGG-3'
AP-10	5'-AGGGTTAGGGTTAGGGTTA \emptyset GG-3'
AP-11	5'-AGGGTTAGGGTTAGGGTTAG \emptyset G-3'

\emptyset indicates the abasic site.

this study are listed in Table 1. Figure 1(b) shows a schematic structure of the antiparallel G-quadruplex of AP-0 in the presence of Na⁺ [43]. When the first guanine is at the top of the G-quartet plane, AP-1, AP-6, and AP-7 contain the abasic site in the top G-quartet plane. AP-2, AP-5, AP-8, and AP-11 contain the abasic site in the middle G-quartet plane. AP-3, AP-4, AP-9, and AP-10 contain the abasic site in the bottom G-quartet plane. By comparing these sequences, it is possible to systematically investigate the molecular crowding effects on the G-quadruplex containing an abasic site at different positions.

3.2. Molecular Crowding Effects on the Structures of DNA Sequences with or without an Abasic Site. We initially studied

the folding topologies of DNA sequences by CD spectroscopy. Figure 2(a) shows CD spectra of AP-0, AP-1, AP-2, and AP-3 in 100 mM NaCl buffer at 0 wt% PEG200 as a dilute condition. The CD spectrum of AP-0 showed positive and negative peaks around 295 nm and 260 nm, respectively. This CD spectrum indicates the formation of an antiparallel G-quadruplex [43], which is consistent with a previous NMR study showing that AP-0 forms an intramolecular antiparallel G-quadruplex in the presence of Na⁺ (Figure 1(a)) [39]. The CD spectra of AP-1, AP-2, and AP-3 also showed positive and negative peaks around 295 nm and 260 nm, respectively, also showing the formation of an antiparallel G-quadruplex. Typical CD spectra for antiparallel G-quadruplexes were further observed for AP-0, AP-1, AP-2, and AP-3 under molecular crowding conditions with 40 wt% PEG200 (Figure 2(b)). We chose 40 wt% PEG200 as the molecular crowding condition because, inside living cells, 20–40 wt% total volume is physically occupied by biomolecules [1–5]. In addition, higher concentrations of PEG200 than 40 wt% sometime induce condensation of DNA. In addition, CD spectra of all DNA sequences at 0 wt% and 40 wt% PEG200 were almost identical with that of AP-0 (see Figure S1 in the Supplementary Material (available online at doi:10.4061/2011/857149) for CD spectra of DNA sequences from AP-4 to AP-11), showing that all DNA sequences form antiparallel G-quadruplexes under both dilute and molecular crowding conditions. These results demonstrated that molecular crowding did not alter the folding topology of G-quadruplexes. On the other hand, the CD intensities of the DNA sequences were different from each other, suggesting that loss of the stacking interaction in the antiparallel G-quadruplexes depends on the position of the abasic site.

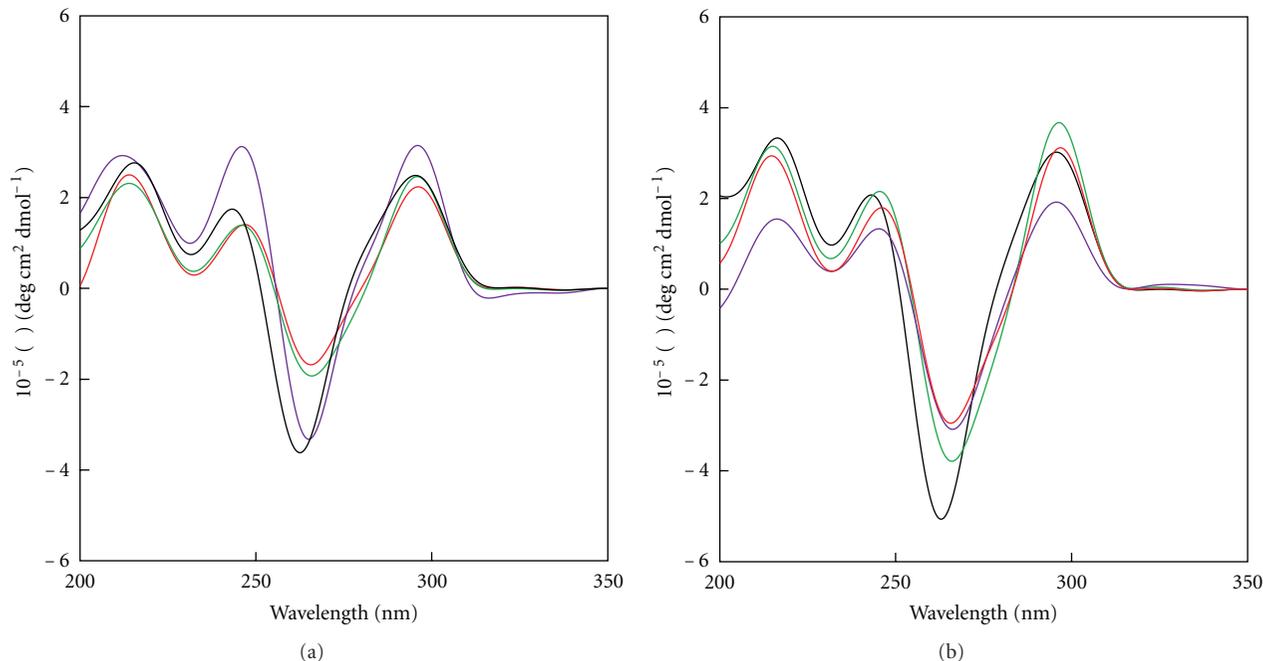


FIGURE 2: CD spectra for 20 μM AP-0 (black), AP-1 (green), AP-2 (purple), and AP-3 (red) in a buffer of 100 mM NaCl, 50 mM Tris-HCl (pH 7.0), and 1 mM Na₂EDTA at 0 wt% (0 mol L⁻¹) (a) and 40 wt% (2 mol L⁻¹) PEG200 (b) at 4°C.

3.3. Molecular Crowding Effects on the Thermodynamics of G-Quadruplexes with or without an Abasic Site. Since all sequences folded into antiparallel G-quadruplexes under dilute and molecular crowding conditions, we further attempted to evaluate the effects of molecular crowding on the thermodynamics of G-quadruplexes with or without an abasic site. Figure 3(a) shows normalized UV melting curves at 295 nm of AP-0 at 0 wt% and 40 wt% PEG200. Hypochromic transitions were observed, indicating that these melting behaviours corresponded to the dissociation of G-quadruplexes [42]. In addition, hysteresis was not observed between the melting and annealing curves for any sequence (data not shown), indicating a two-state transition of the G-quadruplexes. The values of melting temperature (T_m) of AP-0 at 0 wt% and 40 wt% PEG200 were estimated to be 58.1°C and 66.9°C, respectively (Table 2). The T_m values of AP-1, AP-2, and AP-3 at 0 wt% were 52.2°C, 44.2°C, and 50.0°C, respectively. At 40 wt% PEG200, the T_m values of the DNA sequences were 61.9°C, 53.2°C, and 57.2°C, respectively. In addition, the T_m values of DNA sequences from AP-4 to AP-11 at 0 wt% and 40 wt% PEG200 are listed in Table 2. All DNA sequences showed higher T_m at 40 wt% PEG200 than at 0 wt% PEG200. These results indicate that molecular crowding stabilizes G-quadruplexes with or without an abasic site.

The values of free energy changes at 37°C (ΔG_{37}°) for G-quadruplex formations at 0 wt% and 40 wt% PEG200 were further evaluated and are listed in Table 2. Moreover, we evaluated thermodynamic stability of G-quadruplexes of the natural AP-0 and modified AP-1 at 10, 20, and 30 wt% PEG200 as well as 0 and 40 wt% PEG200 to investigate a relationship between the stability of the G-quadruplex ($\ln K_{\text{obs}}$) and water activity ($\ln a_w$) at 37°C (Figure 4).

The plot showed that the stability of the G-quadruplex ($\ln K_{\text{obs}}$) decreased linearly with the increase in $\ln a_w$. Thus, the slope is approximately equal to the numbers of water molecules released upon the formation of the structure [12]. The numbers of water molecules released upon the G-quadruplex formation of AP-0 and AP-1 were estimated to be 74.7 and 56.8, respectively. These results are consistent with the previous reports for the numbers of water molecules through the formation of G-quadruplexes [12, 15]. Noteworthy, this linear relationship between $\ln K_{\text{obs}}$ and $\ln a_w$ indicates that cosolute and cation bindings to G-quadruplex are insignificant. We previously used various neutral cosolutes to induce molecular crowding conditions and to study direct interactions between cosolutes and DNA strand [11, 12]. We found that a cosolute with fewer hydroxyl groups in the vicinal position causes more water molecules to be released during formation of the G-quadruplex. From these results, it is possible to consider that the cosolutes affect the thermodynamics of G-quadruplex by regulating the hydration of the DNA molecule. On the other hand, it was reported that a direct interaction between PEG and DNA strand is thermodynamically unfavorable [45]. Thus, it appears that indirect interactions with cosolutes containing fewer hydroxyl groups, such as PEG, affect the hydration of the DNA structures. These results suggest that the stabilization effect of molecular crowding on G-quadruplex structures with or without an abasic site is due to the hydration change induced by the addition of PEG200.

Based on ΔG_{37}° , we further estimated the effects of molecular crowding as follows: $\Delta \Delta G_{37}^\circ(\text{MC}) = \{\Delta G_{37}^\circ(40 \text{ wt\% PEG200})\} - \{\Delta G_{37}^\circ(0 \text{ wt\% PEG200})\}$. The value of $\Delta \Delta G_{37}^\circ(\text{MC})$ for AP-0 was $-2.3 \text{ kcal mol}^{-1}$. This negative value of $\Delta \Delta G_{37}^\circ(\text{MC})$

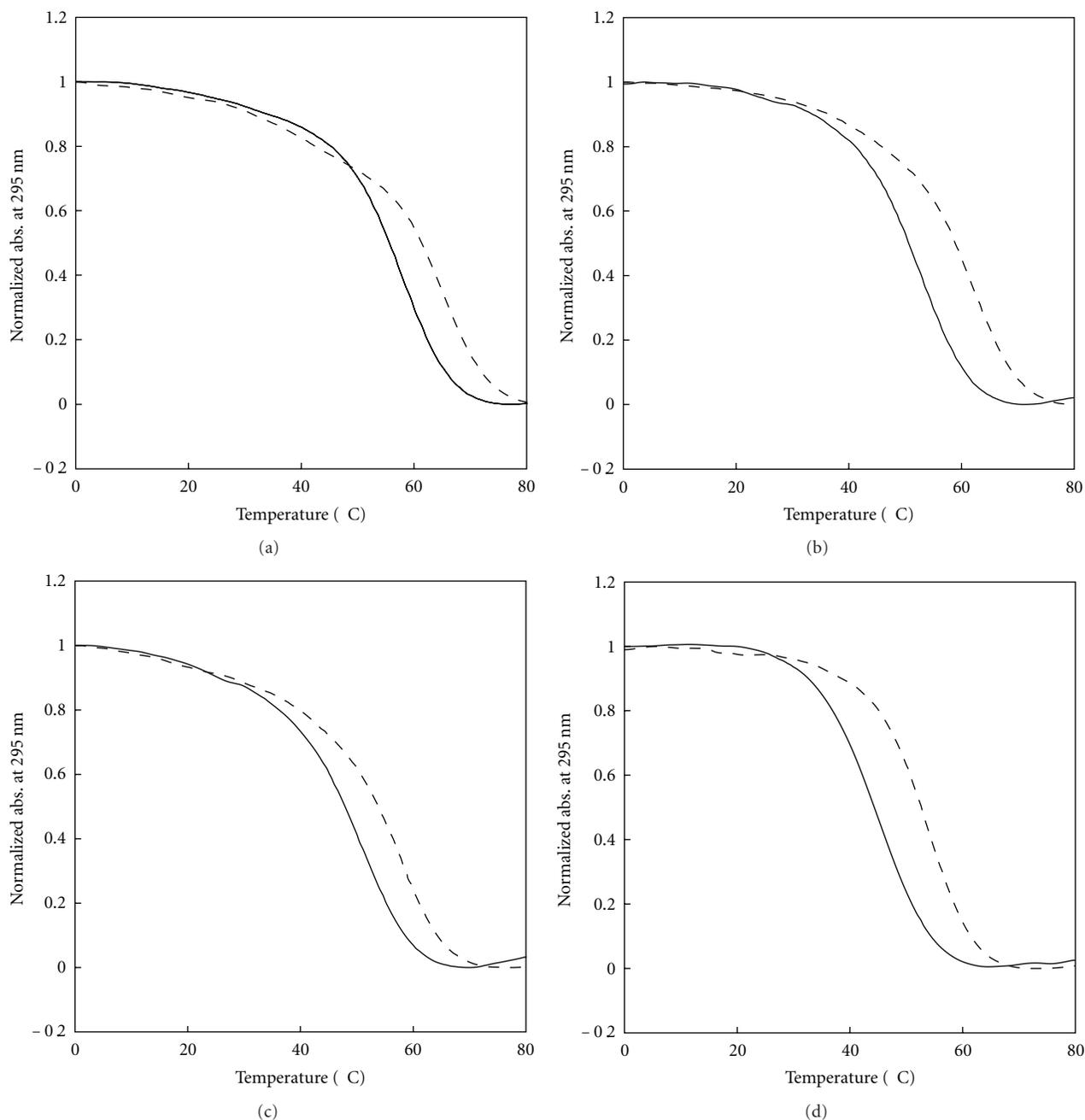


FIGURE 3: Normalized UV melting curves for 20 μM AP-0 (a), AP-1 (b), AP-2 (c), and AP-3 (d) in a buffer of 100 mM NaCl, 50 mM Tris-HCl (pH 7.0), and 1 mM Na_2EDTA at 0 wt% (0 mol L^{-1}) (solid line) and 40 wt% (2 mol L^{-1}) (dashed line) PEG200.

quantitatively shows that molecular crowding stabilizes natural G-quadruplexes, which is consistent with previous studies for molecular crowding effects on the antiparallel G-quadruplex stability [46]. By comparing $\Delta\Delta G_{37(\text{MC})}^\circ$ for AP-0 ($-2.3 \text{ kcal mol}^{-1}$) with those of AP-1 \sim AP-11 ($-1.1 \text{ kcal mol}^{-1} \sim -1.9 \text{ kcal mol}^{-1}$), we found that the degree of stabilization by molecular crowding was reduced by the introduction of an abasic site. Moreover, we found that the stabilizing effect of molecular crowding was the largest for G-quadruplexes with an abasic site in

the middle G-quartet plane (see Figure 1(b)). The average value of $\Delta\Delta G_{37(\text{MC})}^\circ$ for AP-2 ($-1.5 \text{ kcal mol}^{-1}$), AP-5 ($-1.7 \text{ kcal mol}^{-1}$), AP-8 ($-1.8 \text{ kcal mol}^{-1}$), and AP-11 ($-1.9 \text{ kcal mol}^{-1}$) was estimated to be $-1.7 \text{ kcal mol}^{-1}$, which is greater in magnitude than that of AP-1, AP-6, or AP-7 possessing abasic sites in the top G-quartet plane ($-1.5 \text{ kcal mol}^{-1}$), and larger in magnitude than that of AP-3, AP-4, AP-9, and AP-10 possessing an abasic site in the bottom G-quartet plane ($-1.3 \text{ kcal mol}^{-1}$). These results indicated that the abasic site position critically affected

TABLE 2: Thermodynamic parameters of AP series for the formation of G-quadruplex^a.

	PEG200 (wt%)	T_m^b (°C)	ΔH° (kcal mol ⁻¹)	$T\Delta S^\circ$ (kcal mol ⁻¹)	ΔG_{37}° (kcal mol ⁻¹)	$\Delta\Delta G_{37}^\circ(\text{MC})^c$ (kcal mol ⁻¹)	$\Delta\Delta G_{37}^\circ(\text{AP-X})^d$ (kcal mol ⁻¹)
AP-0	0	58.1	-47.9 ± 0.3	-44.8 ± 0.3	-3.1 ± 0.4	-2.3 ± 1.0	—
	40	66.9	-62.1 ± 1.0	-56.6 ± 0.8	-5.5 ± 1.0	—	—
AP-1	0	52.2	-41.3 ± 0.4	-39.4 ± 0.5	-1.9 ± 0.6	-1.6 ± 1.0	1.2 ± 0.7
	40	61.9	-46.5 ± 0.6	-43.0 ± 0.6	-3.5 ± 0.8	—	1.9 ± 1.2
AP-2	0	44.2	-40.8 ± 0.5	-39.9 ± 0.5	-0.9 ± 0.3	-1.5 ± 1.0	2.2 ± 0.5
	40	53.2	-48.2 ± 0.8	-45.8 ± 0.7	-2.4 ± 0.8	—	3.0 ± 1.3
AP-3	0	50.0	-42.1 ± 0.6	-40.4 ± 0.6	-1.7 ± 0.7	-1.2 ± 1.5	1.4 ± 0.8
	40	57.2	-47.9 ± 1.0	-45.0 ± 0.9	-2.9 ± 1.4	—	2.5 ± 1.7
AP-4	0	39.1	-38.6 ± 0.3	-38.3 ± 0.3	-0.3 ± 0.3	-1.1 ± 0.3	2.8 ± 0.5
	40	46.2	-47.9 ± 0.4	-46.3 ± 0.3	-1.4 ± 0.3	—	4.0 ± 1.0
AP-5	0	30.9	-32.6 ± 0.3	-33.3 ± 0.3	0.7 ± 0.3	-1.7 ± 0.6	3.8 ± 0.5
	40	43.8	-46.1 ± 0.4	-45.1 ± 0.4	-1.0 ± 0.5	—	4.4 ± 1.1
AP-6	0	47.0	-40.7 ± 0.4	-39.4 ± 0.4	-1.3 ± 0.6	-1.5 ± 0.9	1.8 ± 0.7
	40	56.1	-48.1 ± 0.6	-45.3 ± 0.6	-2.8 ± 0.7	—	2.6 ± 1.2
AP-7	0	43.5	-45.1 ± 0.4	-44.2 ± 0.4	-0.9 ± 0.5	-1.3 ± 0.9	2.2 ± 0.6
	40	51.3	-48.4 ± 0.4	-46.1 ± 0.4	-2.2 ± 0.5	—	3.2 ± 1.1
AP-8	0	33.5	-31.3 ± 0.3	-31.7 ± 0.3	0.4 ± 0.3	-1.8 ± 0.6	3.5 ± 0.5
	40	46.7	-46.4 ± 0.3	-45.0 ± 0.3	-1.4 ± 0.3	—	4.0 ± 1.0
AP-9	0	41.1	-35.2 ± 0.2	-33.3 ± 0.3	-0.5 ± 0.3	-1.4 ± 0.4	2.6 ± 0.5
	40	49.5	-47.5 ± 0.4	-45.6 ± 0.3	-1.9 ± 0.3	—	3.5 ± 1.0
AP-10	0	38.8	-37.5 ± 0.4	-37.3 ± 0.4	-0.2 ± 0.5	-1.3 ± 0.6	2.9 ± 0.6
	40	48.0	-44.8 ± 0.3	-43.3 ± 0.3	-1.5 ± 0.4	—	3.9 ± 1.1
AP-11	0	36.4	-33.3 ± 0.7	-33.5 ± 0.7	0.2 ± 0.7	-1.9 ± 1.1	3.3 ± 0.9
	40	48.1	-42.8 ± 0.7	-41.1 ± 0.6	-1.7 ± 0.7	—	3.7 ± 1.2

^aThe error values were calculate as shown in [44].

^b T_m is the melting temperature at 20 μM total strand concentration.

^c $\Delta\Delta G_{37}^\circ(\text{MC}) = \Delta G_{37}^\circ(40 \text{ wt}\% \text{ PEG200}) - \Delta G_{37}^\circ(0 \text{ wt}\% \text{ PEG200})$.

^d $\Delta\Delta G_{37}^\circ(\text{AP-X}) = \Delta G_{37}^\circ(\text{AP-X}) - \Delta G_{37}^\circ(\text{AP-0})$ (X indicates the numbers from 1 to 11).

the molecular crowding effect on the thermodynamics of G-quadruplex.

When the PEG200 concentration increased from 0 to 40 wt%, the enthalpy change (ΔH°) of AP-0 for G-quadruplex formation was more favorable, decreasing from -47.9 to -62.1 kcal mol⁻¹, and the entropy change at 37°C ($T\Delta S^\circ$) was more favorable, decreasing from -44.8 to -56.8 kcal mol⁻¹. Similarly, ΔH° and $T\Delta S^\circ$ values for AP-1, AP-2, and AP-3 for G-quadruplex formation decreased with increasing PEG concentration from 0 wt% to 40 wt%. These changes demonstrate that the promotion of G-quadruplex formation by molecular crowding is due to a favourable enthalpic contribution that exceeds the unfavourable entropic contribution. The enthalpic stabilization of G-quadruplexes by molecular crowding agrees with previous reports showing a stabilization effect of molecular crowding on a thrombin-binding aptamer forming an antiparallel G-quadruplex [12, 15]. Moreover, these results demonstrate that the enthalpic contribution of molecular crowding to the thermodynamics of G-quadruplexes depends on the position of the abasic site.

3.4. The Effects of the Abasic Site on the Thermodynamics of DNA G-Quadruplexes. In order to understand how the abasic site regulates the molecular crowding effects on the thermodynamics of G-quadruplexes, we compared the thermodynamic stability of nonmodified (AP-0) and modified G-quadruplexes (AP-1 ~ AP-11) under dilute conditions. The thermodynamic stability of AP-0 was higher than that of AP-1 ~ AP-11 (Table 2). These parameters are consistent with the destabilization of G-quadruplexes by abasic sites as was reported in a previous study [39]. The abasic site effects were quantitatively evaluated as follows: $\Delta\Delta G_{37}^\circ(\text{AP-X}) = \Delta G_{37}^\circ(\text{AP-X}) - \Delta G_{37}^\circ(\text{AP-0})$ (X indicates numbers from 1 to 11). For example, $\Delta\Delta G_{37}^\circ(\text{AP-1})$ at 0 wt% PEG200 was 1.2 kcal mol⁻¹. Moreover, we found that the value of $\Delta\Delta G_{37}^\circ(\text{AP-X})$ was the largest for G-quadruplexes with abasic sites in the middle G-quartet plane (AP-2, AP-5, AP-8, and AP-11). The average value was estimated to be 3.2 kcal mol⁻¹, which was larger than that of AP-1, AP-6, and AP-7 possessing abasic sites in the top G-quartet plane (1.7 kcal mol⁻¹), and larger than that of AP-3, AP-4,

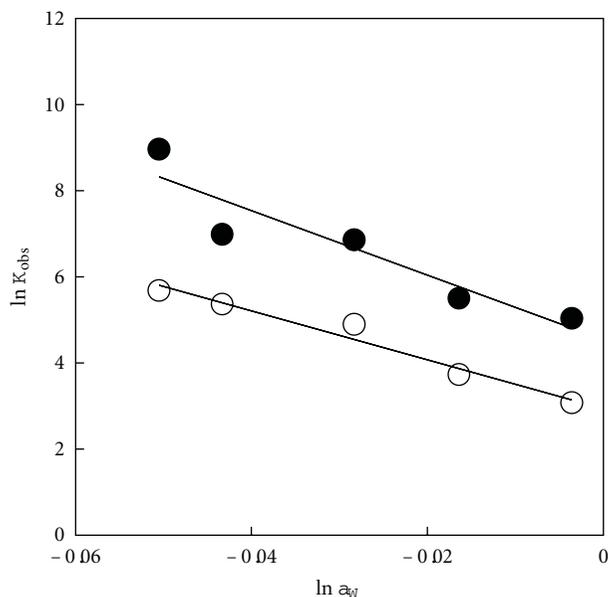


FIGURE 4: Plots of $\ln K_{\text{obs}}$ versus $\ln a_w$ for AP-0 (closed circle) and AP-1 (open circle) for the formation of the G-quadruplex in a buffer of 100 mM NaCl, 50 mM Tris-HCl (pH 7.0), and 1 mM Na_2EDTA containing 0 (0 mol L⁻¹), 10 (0.5 mol L⁻¹), 20 (1 mol L⁻¹), 30 (1.5 mol L⁻¹), or 40 (2 mol L⁻¹) wt% PEG200 at 37°C.

AP-9, and AP-10 possessing abasic sites in the bottom G-quartet plane (2.4 kcal mol⁻¹). These results show that the position of the abasic site significantly affects the abasic site effect on the thermodynamics of G-quadruplex. Interestingly, this order was the same as that observed for the values of $\Delta\Delta G_{37}^{\circ}(\text{MC})$, that is, the molecular crowding effect on the stabilizing of the G-quadruplexes with abasic sites at different positions. Moreover, a comparison of the values of ΔH° and $T\Delta S^{\circ}$ of AP-1 ~ AP-11 shows that the destabilization of G-quadruplexes by the abasic sites was due to unfavourable enthalpy changes. In addition, the values of ΔH° also depended on the position of the abasic site. Thus, it is possible that the enthalpy change of G-quadruplexes depending on the position of the abasic site is related to the molecular crowding effects on the thermodynamics of G-quadruplexes.

3.5. The Relationship between Molecular Crowding and Abasic Site Effects on the Thermodynamics of G-Quadruplexes. We found that stabilization of G-quadruplexes by molecular crowding and destabilization of G-quadruplexes by abasic sites were due to favourable and unfavourable enthalpy changes, respectively. In order to understand a possible mechanism for molecular crowding effects that depend on the position of the abasic site, we compared molecular crowding and abasic site effects on enthalpy changes for G-quadruplex formation. Figure 5 shows the molecular crowding effects on the enthalpy change for the formation of G-quadruplexes containing abasic sites $\{\Delta H_{(40 \text{ wt}\% \text{ PEG200})}^{\circ} - \Delta H_{(0 \text{ wt}\% \text{ PEG200})}^{\circ}\}$ versus the abasic site effects on enthalpy changes for G-quadruplex formation under dilute conditions

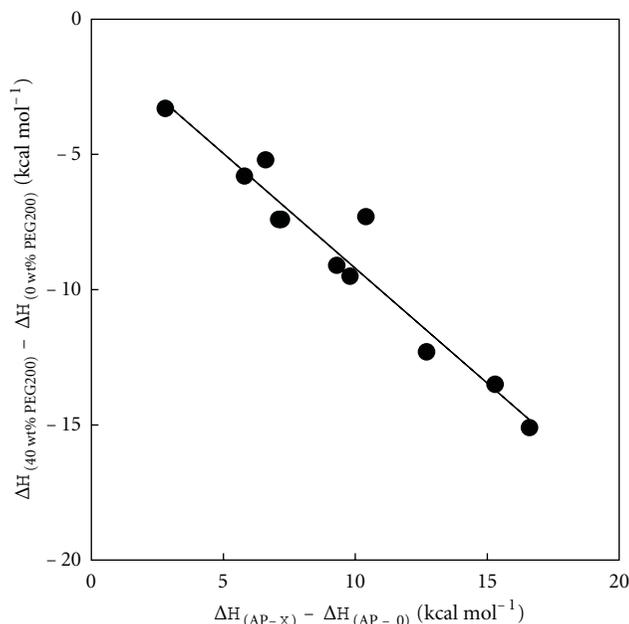


FIGURE 5: Plots of the molecular crowding effects on enthalpy change for G-quadruplex formation $\Delta H_{(40 \text{ wt}\% \text{ PEG200})}^{\circ} - \Delta H_{(0 \text{ wt}\% \text{ PEG200})}^{\circ}$ versus abasic site effects on enthalpy change for G-quadruplex formation $\Delta H_{(\text{AP-X})}^{\circ} - \Delta H_{(\text{AP-0})}^{\circ}$ at 0 wt% PEG200 (X indicates the numbers from 1 to 11) in a buffer of 100 mM NaCl, 50 mM Tris-HCl (pH 7.0), and 1 mM Na_2EDTA at 0 wt% (0 mol L⁻¹) and 40 wt% (2 mol L⁻¹) PEG200.

$\Delta H_{(\text{AP-X})}^{\circ} - \Delta H_{(\text{AP-0})}^{\circ}$, where X indicates numbers from 1 to 11. The plot reveals that the molecular crowding effects on the enthalpy change for the formation of G-quadruplexes containing abasic sites at different positions decreased linearly as the abasic site effects on enthalpy change increased. The negative slope shows that a G-quadruplex having a more unfavourable enthalpic change induced by an abasic site is more enthalpically stabilized by molecular crowding. Interestingly, the slope was around -1 . This compensation relationship between the enthalpy changes suggests that molecular crowding can restore an unfavourable enthalpic loss induced by an abasic site, which may be useful to maintain a G-quadruplex structure with an abasic site under molecular crowding conditions.

Olsen et al. reported that base substitution in the loops causes a decrease in the thermal stability of the G-quadruplex, which can be explained in terms of stacked loops on the G-quartets and base-base stacking within the loops, as well as their overall hydration contributions [15]. Moreover, in case of a DNA duplex with an abasic site, destabilization depends on neighbouring bases due to the different stacking conformations and energies between the substituted base and neighbouring bases [47–49]. Thus, it is possible that the molecular crowding effects on the thermodynamics of G-quadruplexes including a single abasic site depended on position of the abasic site due to local stacking energies altered. However, further studies are required of base substitution in the loop region of G-quadruplexes under

different topologies. In contrast to duplexes, the effect of an abasic site on the thermodynamics of G-quadruplexes largely depends on the stacking position. This dependency can create a polymorphic nature in the structure and thermal stability of G-quadruplexes, which should affect biological systems related to guanine-rich DNA sequences. However, quantitative results obtained here imply that molecular crowding can reduce such risks. This buffering effect to reduce responses of biomolecules against extended stimuli has been proposed as one of the important functions of molecular crowding to maintain homeostasis of cells [50].

4. Conclusion

We examined the effects of molecular crowding on the structure and stability of G-quadruplexes with abasic sites at different positions. The CD spectra showed that molecular crowding did not affect the topology of G-quadruplexes. On the other hand, UV melting analysis showed that the effects of molecular crowding on the thermodynamic stability of G-quadruplexes depended on the position of the abasic site. Moreover, a systematic comparison of the effects of molecular crowding and abasic sites on the thermodynamic parameters demonstrated that the molecular crowding effects on enthalpy changes during the formation of G-quadruplexes with abasic sites at different positions decreased linearly as the abasic site effects increased. This compensatory relationship suggests that molecular crowding plays a role in the maintenance of G-quadruplex structures with abasic sites in cells.

Acknowledgment

This work was supported in part by Grants-in-Aid for Scientific Research and the “Core Research” project (2009–2014), from the Ministry of Education, Culture, Sports, Science and Technology, Japan, the Hiraio Taro Foundation of the Konan University Association for Academic Research, and the Long-Range Research Initiative Project of Japan Chemical Industry Association.

References

- [1] S. B. Zimmerman and A. P. Minton, “Macromolecular crowding: biochemical, biophysical, and physiological consequences,” *Annual Review of Biophysics and Biomolecular Structure*, vol. 22, pp. 27–65, 1993.
- [2] A. P. Minton, “The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media,” *Journal of Biological Chemistry*, vol. 276, no. 14, pp. 10577–10580, 2001.
- [3] R. J. Ellis and A. P. Minton, “Cell biology: join the crowd,” *Nature*, vol. 425, no. 6953, pp. 27–28, 2003.
- [4] O. Medalia, I. Weber, A. S. Frangakis, D. Nicastro, G. Gerisch, and W. Baumeister, “Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography,” *Science*, vol. 298, no. 5596, pp. 1209–1213, 2002.
- [5] R. J. Ellis, “Macromolecular crowding: an important but neglected aspect of the intracellular environment,” *Current Opinion in Structural Biology*, vol. 11, no. 1, pp. 114–119, 2001.
- [6] P. R. Davis-Searles, A. J. Saunders, D. A. Erie, D. J. Winzor, and G. J. Pielak, “Interpreting the effects of small uncharged solutes on protein-folding equilibria,” *Annual Review of Biophysics and Biomolecular Structure*, vol. 30, pp. 271–306, 2001.
- [7] J. R. Wenner and V. A. Bloomfield, “Crowding effects on EcoRV kinetics and binding,” *Biophysical Journal*, vol. 77, no. 6, pp. 3234–3241, 1999.
- [8] Y. Sasaki, D. Miyoshi, and N. Sugimoto, “Regulation of DNA nucleases by molecular crowding,” *Nucleic Acids Research*, vol. 35, no. 12, pp. 4086–4093, 2007.
- [9] C. H. Spink and J. B. Chaires, “Selective stabilization of triplex DNA by poly(ethylene glycols),” *Journal of the American Chemical Society*, vol. 117, no. 51, pp. 12887–12888, 1995.
- [10] R. Goobes and A. Minsky, “Thermodynamic aspects of triplex DNA formation in crowded environments,” *Journal of the American Chemical Society*, vol. 123, no. 50, pp. 12692–12693, 2001.
- [11] S. Nakano, H. Karimata, T. Ohmichi, J. Kawakami, and N. Sugimoto, “The effect of molecular crowding with nucleotide length and cosolute structure on DNA duplex stability,” *Journal of the American Chemical Society*, vol. 126, no. 44, pp. 14330–14331, 2004.
- [12] D. Miyoshi, H. Karimata, and N. Sugimoto, “Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions,” *Journal of the American Chemical Society*, vol. 128, no. 24, pp. 7957–7963, 2006.
- [13] D. Miyoshi, K. Nakamura, H. Tateishi-Karimata, T. Ohmichi, and N. Sugimoto, “Hydration of Watson-Crick base pairs and dehydration of Hoogsteen base pairs inducing structural polymorphism under molecular crowding conditions,” *Journal of the American Chemical Society*, vol. 131, no. 10, pp. 3522–3531, 2009.
- [14] M. C. Miller, R. Buscaglia, J. B. Chaires, A. N. Lane, and J. O. Trent, “Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3′ sequence of d(AG₃(TTAG₃)₃),” *Journal of the American Chemical Society*, vol. 132, no. 48, pp. 17105–17107, 2010.
- [15] C. M. Olsen, H. T. Lee, and L. A. Marky, “Unfolding thermodynamics of intramolecular G-quadruplexes: base sequence contributions of the loops,” *Journal of Physical Chemistry B*, vol. 113, no. 9, pp. 2587–2595, 2009.
- [16] A. Arora and S. Maiti, “Stability and molecular recognition of quadruplexes with different loop length in the absence and presence of molecular crowding agents,” *Journal of Physical Chemistry B*, vol. 113, no. 25, pp. 8784–8792, 2009.
- [17] A. Rajendran, S. Nakano, and N. Sugimoto, “Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH,” *Chemical Communications*, vol. 46, no. 8, pp. 1299–1301, 2010.
- [18] S. Nakano, H. T. Karimata, Y. Kitagawa, and N. Sugimoto, “Facilitation of RNA enzyme activity in the molecular crowding media of cosolutes,” *Journal of the American Chemical Society*, vol. 131, no. 46, pp. 16881–16888, 2009.
- [19] A. K. Todd, M. Johnston, and S. Neidle, “Highly prevalent putative quadruplex sequence motifs in human DNA,” *Nucleic Acids Research*, vol. 33, no. 9, pp. 2901–2907, 2005.
- [20] R. C. Allsopp, H. Vaziri, C. Patterson et al., “Telomere length predicts replicative capacity of human fibroblasts,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 21, pp. 10114–10118, 1992.
- [21] L. A. Klobutcher, M. T. Swanton, P. Donini, and D. M. Prescott, “All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3′ terminus,” *Proceedings of the National Academy of Sciences of*

- the United States of America*, vol. 78, no. 5, pp. 3015–3019, 1981.
- [22] J. R. Williamson, “G-quartet structures in telomeric DNA,” *Annual Review of Biophysics and Biomolecular Structure*, vol. 23, pp. 703–730, 1994.
- [23] J. L. Huppert and S. Balasubramanian, “Prevalence of quadruplexes in the human genome,” *Nucleic Acids Research*, vol. 33, no. 9, pp. 2908–2916, 2005.
- [24] E. H. Blackburn and P. B. Challoner, “Identification of a telomeric DNA sequence in *Trypanosoma brucei*,” *Cell*, vol. 36, no. 2, pp. 447–457, 1984.
- [25] E. H. Blackburn, “Structure and function of telomeres,” *Nature*, vol. 350, no. 6319, pp. 569–573, 1991.
- [26] D. J. Patel, A. T. Phan, and V. Kuryavii, “Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics,” *Nucleic Acids Research*, vol. 35, no. 22, pp. 7429–7455, 2007.
- [27] A. C. Vallur and N. Maizels, “Activities of human exonuclease 1 that promote cleavage of transcribed immunoglobulin switch regions,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 43, pp. 16508–16512, 2008.
- [28] J. W. Szostak and E. H. Blackburn, “Cloning yeast telomeres on linear plasmid vectors,” *Cell*, vol. 29, no. 1, pp. 245–255, 1982.
- [29] C. W. Greider and E. H. Blackburn, “The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity,” *Cell*, vol. 51, no. 6, pp. 887–898, 1987.
- [30] E. S. Baker, J. T. Lee, J. L. Sessler, and M. T. Bowers, “Cyclo[n]pyrroles: size and site-specific binding to G-quadruplexes,” *Journal of the American Chemical Society*, vol. 128, no. 8, pp. 2641–2648, 2006.
- [31] J. Seenisamy, S. Bashyam, V. Gokhale et al., “Design and synthesis of an expanded porphyrin that has selectivity for the c-MYC G-quadruplex structure,” *Journal of the American Chemical Society*, vol. 127, no. 9, pp. 2944–2959, 2005.
- [32] P. S. Shirude, E. R. Gillies, S. Ladame et al., “Macrocyclic and helical oligoamides as a new class of G-quadruplex ligands,” *Journal of the American Chemical Society*, vol. 129, no. 39, pp. 11890–11891, 2007.
- [33] H. Yaku, T. Murashima, D. Miyoshi, and N. Sugimoto, “Anionic phthalocyanines targeting G-quadruplexes and inhibiting telomerase activity in the presence of excessive DNA duplexes,” *Chemical Communications*, vol. 46, no. 31, pp. 5740–5742, 2010.
- [34] B. van Steensel, A. Smogorzewska, and T. De Lange, “TRF2 protects human telomeres from end-to-end fusions,” *Cell*, vol. 92, no. 3, pp. 401–413, 1998.
- [35] T. Lindahl, “Instability and decay of the primary structure of DNA,” *Nature*, vol. 362, no. 6422, pp. 709–715, 1993.
- [36] T. Lindahl and B. Nyberg, “Rate of depurination of native deoxyribonucleic acid,” *Biochemistry*, vol. 11, no. 19, pp. 3610–3618, 1972.
- [37] J. Tchou, H. Kasai, S. Shibutani et al., “8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 11, pp. 4690–4694, 1991.
- [38] V. Esposito, L. Martino, G. Citarella et al., “Effects of abasic sites on structural, thermodynamic and kinetic properties of quadruplex structures,” *Nucleic Acids Research*, vol. 38, no. 6, pp. 2069–2080, 2009.
- [39] P. Školáková, K. Bednářová, M. Vorlíčková, and J. Sági, “Quadruplexes of human telomere dG₃(TTAG₃)₃ sequences containing guanine abasic sites,” *Biochemical and Biophysical Research Communications*, vol. 399, no. 2, pp. 203–208, 2010.
- [40] N. Sugimoto, S. Nakano, M. Katoh et al., “Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes,” *Biochemistry*, vol. 34, no. 35, pp. 11211–11216, 1995.
- [41] S. Nakano, T. Kanzaki, and N. Sugimoto, “Influences of Ribonucleotide on a duplex conformation and its thermal stability: study with the chimeric RNA-DNA strands,” *Journal of the American Chemical Society*, vol. 126, no. 4, pp. 1088–1095, 2004.
- [42] J. L. Mergny, J. Li, L. Lacroix, S. Amrane, and J. B. Chaires, “Thermal difference spectra: a specific signature for nucleic acid structures,” *Nucleic Acids Research*, vol. 33, no. 16, p. e138, 2005.
- [43] Y. Wang and D. J. Patel, “Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex,” *Structure*, vol. 1, no. 4, pp. 263–282, 1993.
- [44] N. Sugimoto, M. Nakano, and S. Nakano, “Thermodynamics—structure relationship of single mismatches in RNA/DNA duplexes,” *Biochemistry*, vol. 39, no. 37, pp. 11270–11281, 2000.
- [45] V. V. Vasilevskaya, A. R. Khokhlov, Y. Matsuzawa, and K. Yoshikawa, “Collapse of single DNA molecule in poly(ethylene glycol) solutions,” *The Journal of Chemical Physics*, vol. 102, no. 16, pp. 6595–6602, 1995.
- [46] Z. Chen, K. W. Zheng, Y. H. Hao, and Z. Tan, “Reduced or diminished stabilization of the telomere G-quadruplex and inhibition of telomerase by small chemical ligands under molecular crowding condition,” *Journal of the American Chemical Society*, vol. 131, no. 30, pp. 10430–10438, 2009.
- [47] J. Sági, B. Hang, and B. Singer, “Sequence-dependent repair of synthetic AP sites in 15-mer and 35-mer oligonucleotides: role of thermodynamic stability imposed by neighbor bases,” *Chemical Research in Toxicology*, vol. 12, no. 10, pp. 917–923, 1999.
- [48] J. Sági, A. B. Guliaev, and B. Singer, “15-Mer DNA duplexes containing an abasic site are thermodynamically more stable with adjacent purines than with pyrimidines,” *Biochemistry*, vol. 40, no. 13, pp. 3859–3868, 2001.
- [49] C. A. Gelfand, G. E. Plum, A. P. Grollman, F. Johnson, and K. J. Breslauer, “Thermodynamic consequences of an abasic lesion in duplex DNA are strongly dependent on base sequence,” *Biochemistry*, vol. 37, no. 20, pp. 7321–7327, 1998.
- [50] R. Goobes, N. Kahana, O. Cohen, and A. Minsky, “Metabolic buffering exerted by macromolecular crowding on DNA-DNA interactions: origin and physiological significance,” *Biochemistry*, vol. 42, no. 8, pp. 2431–2440, 2003.

Research Article

Direct Detection of Thrombin Binding to 8-Bromodeoxyguanosine-Modified Aptamer: Effects of Modification on Affinity and Kinetics

Shou Goji and Jun Matsui

Department of Nanobiochemistry, FIRST, Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

Correspondence should be addressed to Jun Matsui, matsui@konan-u.ac.jp

Received 14 May 2011; Revised 15 July 2011; Accepted 21 July 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 S. Goji and J. Matsui. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The affinity of an 8-bromodeoxyguanosine- (8-BrdG-) substituted thrombin-binding aptamer (TBA-Br), which has the 1st and 10th guanosine residues replaced with 8-BrdG, was estimated using reflectometric interference spectroscopy (RIfS). When comparing TBA-Br with unmodified TBA (TBA-H), it was demonstrated that the modification effectively improved the affinity of TBA; dissociation constants (K_D) of TBA-H and TBA-Br were 45.4 nM and 1.99 nM, respectively. These values, which were obtained by direct observation of thrombin binding using RIfS, have the same order of magnitude as those obtained in our previous study utilizing conformational changes in TBA to detect thrombin binding, thus confirming the validity of the obtained K_D values. RIfS measurements also revealed that the 8-BrdG modification resulted in a lower dissociation rate constant (k_d), which suggests that the enhancement of affinity can be attributed to the stabilization of the G-quadruplex structure on introduction of 8-BrdG.

1. Introduction

Chemical modification of nucleic acids is a useful approach to improve the stability of higher-order structures [1–5]. Thrombin-binding aptamer (TBA), d(GGTTGGTGTGGT-TGG), is a nucleic acid for which modification has attracted considerable attention, as the improved stability may lead to enhanced affinity for the target species, thrombin, which is potentially useful for the diagnosis and treatment of various conditions. Because the relationship between higher-order structure and affinity is well known, its modification has been widely studied. TBA is known to fold into a G-quadruplex structure consisting of two G-planes connected via a TGT loop and two TT loops, as shown in Figure 1(a) [6–11].

Strategies for modifying TBA can be classified into two categories: modification of the nucleotide backbone such as Locked Nucleic Acid (LNA) [12, 13] or 2'-fluoro-arabinonucleic acid (2'-F-ANA) [14] and modification of nucleobases, such as alkylation or phenylation [15]. Both strategies have been shown to be effective for stabilizing higher-order

structure by stabilizing specific glycosidic bond conformations in the G-plane.

We recently demonstrated that the higher-order structure of TBA can be stabilized by introducing 8-bromodeoxyguanosine (8-BrdG, Figure 1(b)), which stabilizes a *syn* conformation of a glycosidic bond by steric hindrance between the bromo group at the C8 position and the deoxyribose moiety [16–18]. When two 8-BrdG residues were introduced in place of the 1st and 10th guanosine residues, which have a *syn* conformation, the higher-order structure of TBA was significantly stabilized, with an increase in melting point of 15.8°C, and the affinity for thrombin showed a 12.5-fold greater binding constant [19]. In a previous study, however, the binding of thrombin to TBA was detected indirectly; conformational transition of TBA from a single strand to a G-quadruplex, which is induced by the binding of thrombin, was monitored using a fluorescent dye. Therefore, direct measurement of the thrombin-binding phenomenon has been required in order to verify the estimated binding constant values and the effects of 8-BrdG on enhancing the affinity of TBA for thrombin. Furthermore, the practical

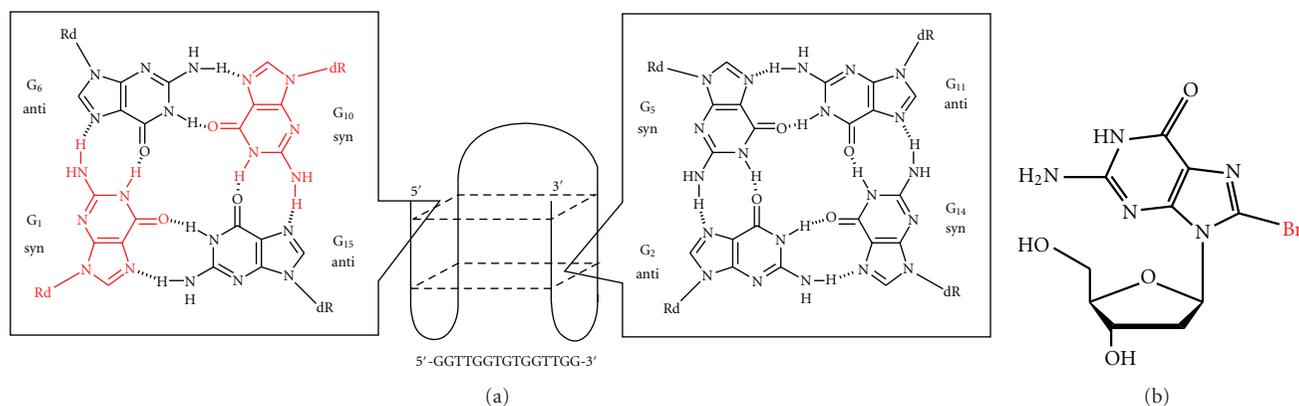


FIGURE 1: Structure of (a) G-quadruplex of thrombin binding aptamer and (b) 8-bromodeoxyguanosine. Modification sites, G1 and G10, are indicated in red.

usefulness of the 8-BrdG-modified aptamer has not been investigated.

In this study, we performed direct observation of thrombin binding to an 8-BrdG-modified TBA (TBA-Br) using reflectometric interference spectroscopy (RiFS), which can measure changes in optical thickness of a molecular layer on a sensor chip and is, therefore, expected to detect thrombin binding to the modified TBA immobilized on a chip [20–23] in order to assess the binding constants of TBA-Br and the effectiveness of TBA-Br for developing a sensitive TBA sensor.

2. Materials and Methods

2.1. Materials. Human α -thrombin was purchased from Haematologic Technologies Inc. (Essex, VT). *N*-(6-Maleimidocaproyloxy)sulfosuccinimide sodium salt (Sulfo-EMCS) and 6-Hydroxy-1-hexanethiol (HTT) were purchased from Dojindo (Kumamoto, Japan). 3-Aminopropyltriethoxysilane (APTS) was purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan). All DNA amidites and 5'-thiol modifier C6 were purchased from Glen Research Corp. (Sterling, VA). 5'-Thiolated unmodified thrombin-binding aptamer (TBA-H) was obtained from Hokkaido System Science Co., Ltd. (Sapporo, Japan). The RiFS sensor MI-affinity LCR-01 and its consumable supplies (SiN sensor chips and PDMS flow cells) were purchased from Konica Minolta Opto, Inc. (Tokyo, Japan). Other reagents were purchased from Wako Pure Chemical (Tokyo, Japan).

Base sequences of aptamers were as follows.

TBA-H:



TBA-Br:



2.2. DNA Synthesis. TBA-Br was synthesized using an AB 3400 DNA synthesizer (Applied Biosystems Inc., Tokyo,

Japan). Synthesized oligonucleotides were purified by 20% denaturing PAGE. Concentration of single-stranded oligonucleotides was determined by measuring the absorbance at 260 nm at a high temperature using a UV-1700 spectrometer connected with TMSPC-8 thermoprogrammer (Shimadzu Co., Ltd Kyoto, Japan). Oligonucleotides were treated by dithiothreitol prior to use.

2.3. Chip Preparation. The SiN Chip surface was irradiated with UV under vacuum for 1 h using UER20-172VB (Ushio Inc., Tokyo, Japan). UV-treated chips were soaked in acetone containing 1% APTS for 1 h, and after washing with acetone, chips were heated at 110°C for 30 min. Thrombin-binding aptamers were immobilized *in situ* using RiFS by injecting 1 mM Sulfo-EMCS, 50 μM aptamer, and 5 mM HTT in PBS buffer at 30°C. An Econoflo Syringe Pump (Harvard Apparatus, Holliston, Mass, USA) was used to move PBS buffer at a flow rate of 2.0 $\mu\text{L min}^{-1}$. Obtained sensor chips were stored in PBS buffer at 4°C.

2.4. RiFS Measurements. A RiFS sensor (MI-affinity LCR-01) was used for direct monitoring of the association/dissociation process. All measurements were conducted using binding buffer (20 mM Tris buffer, pH 7.4) containing 140 mM NaCl, 5 mM KCl, 1 mM MgCl_2 , 1 mM CaCl_2 , and 1% PEG at 25°C and at a flow rate of 20 $\mu\text{L min}^{-1}$. Kinetic analysis was carried out using software associated with the MI-affinity LCR-01. Triplicate measurements were conducted, and average data are shown in Table 1.

3. Results and Discussion

3.1. Chip Preparation. Sensor chips were prepared *in situ* using RiFS (Figure 2). A hetero bifunctional cross-linker, Sulfo-EMCS, which bears a maleimide group and succinimide group at each end, was used to immobilize a thiol-terminated aptamer onto an aminated SiN chip. HTT was used to cap the unreacted maleimide termini of the immobilized Sulfo-EMCS moieties. After injection of aptamer for immobilization, an approximately 1.2-nm shift in wavelength was

TABLE 1: Binding properties of each TBA.

	k_a $M^{-1} s^{-1}$	k_d s^{-1}	K_D nM
Nonmodified TBA	7.45 ± 0.20 $\times 10^4$	3.38 ± 0.33 $\times 10^{-3}$	45.4 ± 5.6
8-BrdG substituted TBA	1.58 ± 0.10 $\times 10^5$	3.15 ± 0.46 $\times 10^{-4}$	1.99 ± 0.29

observed with both the unmodified (TBA-H) and modified aptamers (TBA-Br), thus confirming the same degree of immobilization for both aptamers. TBA-Br possesses two 8-BrdG residues at the 1st and 10th guanosine sites, and based on our previous study, is optimized in terms of the number and position of the 8-BrdG groups [19].

3.2. Kinetic Measurement. The sensor chips immobilized with TBA-H and TBA-Br were assessed by injection of 100 nM thrombin. The response was reproducible with variations of less than 8.7%. When compared to the TBA-H-immobilized chip (Figure 3(a)), the TBA-Br-immobilized chip showed a larger response (Figure 3(b)). These results suggest that TBA-Br has a higher affinity for thrombin than the TBA-H and accordingly captured a larger number of thrombin molecules.

Using the sensorgrams obtained on injection of 100 nM thrombin, kinetic analysis was performed (Table 1). TBA-Br showed a larger association rate constant (k_a) ($1.58 \times 10^5 M^{-1} s^{-1}$) than TBA-H ($7.45 \times 10^4 M^{-1} s^{-1}$). However, the effect on dissociation rate constant was more significant; TBA-Br showed a dissociation rate constant an order of magnitude lower ($k_d = 3.15 \times 10^{-4} s^{-1}$) than TBA-H ($k_d = 3.38 \times 10^{-3} s^{-1}$). Because thrombin binding with TBA is accompanied by G-quadruplex formation of TBA, the decrease in dissociation rate constant may be attributed to the stabilization of G-quadruplex structure of TBA (TBA-Br) due to the introduction of 8-BrdG; the T_m value for TBA-H was 50.7°C while that for TBA-Br was 66.5°C [19].

Dissociation constants (K_D) for TBA-H and TBA-Br were 45.4 nM and 1.99 nM, respectively. These values are within the same order of magnitude as those obtained by our previous study using a fluorescent dye [19]. It is also notable that the dissociation constant for TBA-H reported elsewhere is between 20 and 102.6 nM, which was obtained by other detection methods, such as SPR [24, 25], QCM [26], ITC [27], and an electrochemical sensor [28], thus supporting the suitability of the applied conditions and estimation protocol. Direct detection of binding between TBA and thrombin appeared to be successful using RfS, and the enhancement of affinity by introducing 8-BrdG to TBA was confirmed. Interestingly, the dissociation constant (K_D) for TBA-Br estimated by RfS was slightly lower than that obtained using the fluorescent dye [19], whereas the dissociation constant (K_D) for TBA-H estimated by RfS was larger than that obtained using the fluorescent dye. Although the cause of these discrepancies is not currently understood, it is possible that even in the absence of thrombin, TBA-Br forms a G-quadruplex in the binding buffer at the temperature used

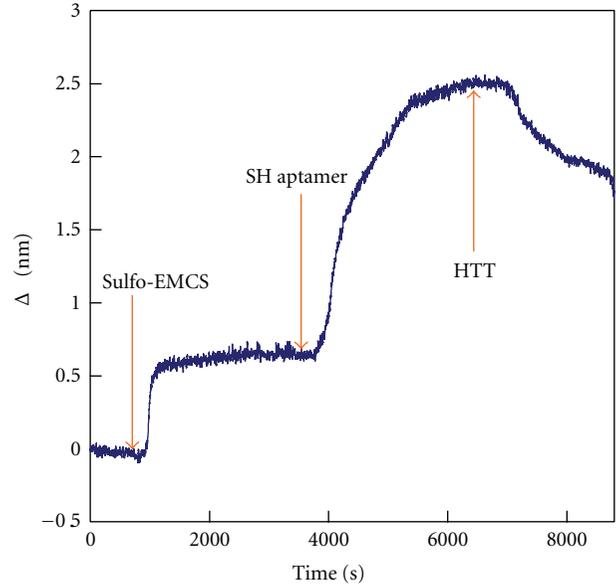


FIGURE 2: *In situ* chip preparation process monitored by reflectometric interference spectroscopy (RfS). Chips were prepared in PBS at 30°C. Flow rate was $2.0 \mu L \text{ min}^{-1}$.

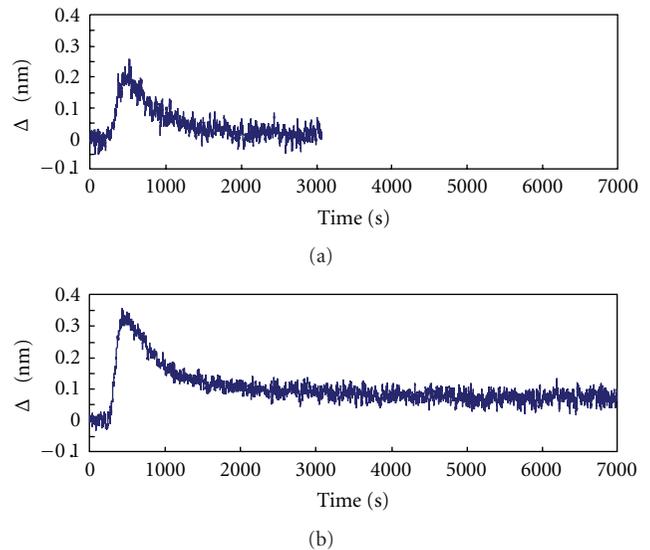


FIGURE 3: Sensorgrams on injection of 100 nM thrombin with (a) TBA-H-immobilized chip and (b) TBA-Br-immobilized chip. Measurements were conducted in binding buffer at 25°C. Flow rate was $20 \mu L \text{ min}^{-1}$.

for RfS measurements due to the improved stability, which would be favorable for binding with thrombin as a result of preorganization effects.

3.3. Effects of Aptamer Modification on Sensitivity. Sensors that sensitively detect thrombin are potentially useful for diagnosis of thrombosis. The TBA-Br-immobilized chip is expected to offer improved sensitivity due to its high affinity for thrombin. Thus, we investigated the sensitivity of the RfS

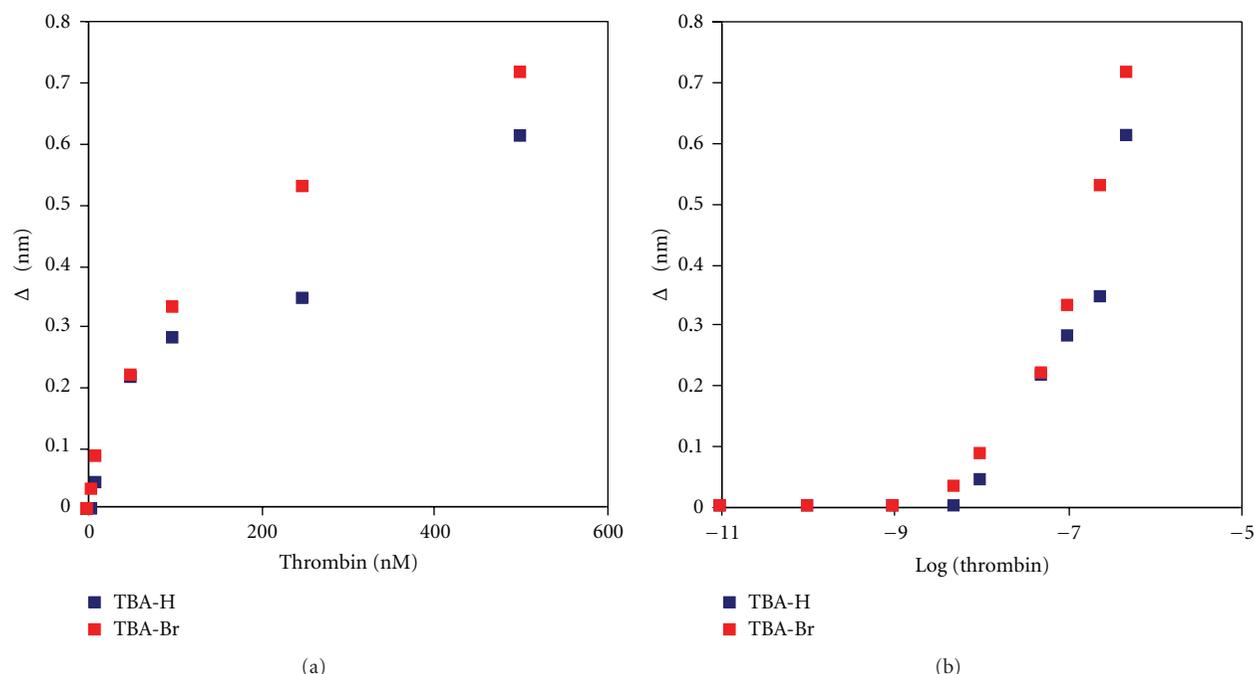


FIGURE 4: Calibration curves for thrombin on TBA-H-immobilized chip and TBA-Br-immobilized chip. (a) Change in wavelength versus thrombin concentration. (b) Change in wavelength versus log (thrombin concentration). Measurements were conducted in binding buffer at 25°C. Flow rate was 20 $\mu\text{L min}^{-1}$.

sensor with an TBA-Br-immobilized chip. Responses to various concentrations of thrombin (from 0.01 nM to 500 nM) are plotted in Figure 4. The TBA-Br-immobilized chip showed a significant response to 1 nM thrombin with variations of less than 5%, whereas the detection limit of the TBA-H-immobilized chip was higher than 5 nM. Although the degree was moderate, it confirms that modification of TBA to improve affinity can contribute to enhancing thrombin-sensor sensitivity.

4. Conclusions

In this study, we directly observed the association between TBA-Br and thrombin by RfS in order to estimate the kinetic rate and affinity constants. The estimated dissociation constants had the same order of magnitude as those obtained in a previous study using a fluorescent dye, thus suggesting that modification with 8-BrdG at appropriate residues in the G-plane increases the affinity of TBA. In addition, the decrease observed in dissociation rate constant by introduction of 8-BrdG suggests that affinity enhancement is due to the effect of 8-BrdG in stabilizing the G-quadruplex structure of TBA. A preliminary test on the sensitivity of TBA-immobilized sensor chips also demonstrated that modification to improve affinity resulted in a better detection limit although the effect is currently moderate.

Acknowledgment

The authors are grateful to Konica Minolta Opto, Inc., for technical assistance and for lending them a prototype of the MI-affinity LCR-01 RfS sensor.

References

- [1] S. E. Osborne, R. J. Cain, and G. D. Glick, "Structure and dynamics of disulfide cross-linked DNA triple helices," *Journal of the American Chemical Society*, vol. 119, no. 6, pp. 1171–1182, 1997.
- [2] K. S. Schmidt, S. Borkowski, J. Kurreck et al., "Application of locked nucleic acids to improve aptamer in vivo stability and targeting function," *Nucleic Acids Research*, vol. 32, no. 19, pp. 5757–5765, 2004.
- [3] B. Saccà, L. Lacroix, and J. L. Mergny, "The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides," *Nucleic Acids Research*, vol. 33, no. 4, pp. 1182–1192, 2005.
- [4] D. Shangguan, Z. Tang, P. Mallikaratchy, Z. Xiao, and W. Tan, "Optimization and modifications of aptamers selected from live cancer cell lines," *A European Journal of Chemical Biology*, vol. 8, no. 6, pp. 603–606, 2007.
- [5] E. B. Pedersen, J. T. Nielsen, C. Nielsen, and V. V. Filichev, "Enhanced anti-HIV-1 activity of G-quadruplexes comprising locked nucleic acids and intercalating nucleic acids," *Nucleic Acids Research*, vol. 39, no. 6, pp. 2470–2481, 2011.
- [6] L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas, and J. J. Toole, "Selection of single-stranded DNA molecules that bind and inhibit human thrombin," *Nature*, vol. 355, no. 6360, pp. 564–566, 1992.
- [7] R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, and J. Feigon, "Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 8, pp. 3745–3749, 1993.
- [8] P. Schultze, R. F. Macaya, and J. Feigon, "Three-dimensional solution structure of the thrombin-binding DNA aptamer

- d(GGTTGGTGTGGTTGG)," *Journal of Molecular Biology*, vol. 235, no. 5, pp. 1532–1547, 1994.
- [9] K. Y. Wang, S. McCurdy, R. G. Shea, S. Swaminathan, and P. H. Bolton, "A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA," *Biochemistry*, vol. 32, no. 8, pp. 1899–1904, 1993.
- [10] K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler, and A. Tulinsky, "The structure of α -thrombin inhibited by a 15-mer single-stranded DNA aptamer," *Journal of Biological Chemistry*, vol. 268, no. 24, pp. 17651–17654, 1993.
- [11] K. Padmanabhan and A. Tulinsky, "An ambiguous structure of a DNA 15-mer thrombin complex," *Acta Crystallographica Section D*, vol. 52, no. 2, pp. 272–282, 1996.
- [12] A. Virno, A. Randazzo, C. Giancola, M. Bucci, G. Cirino, and L. Mayol, "A novel thrombin binding aptamer containing a G-LNA residue," *Bioorganic and Medicinal Chemistry*, vol. 15, no. 17, pp. 5710–5718, 2007.
- [13] L. Bonifacio, F. C. Church, and M. B. Jarstfer, "Effect of locked-nucleic acid on a biologically active G-quadruplex. A structure-activity relationship of the thrombin aptamer," *International Journal of Molecular Sciences*, vol. 9, no. 3, pp. 422–433, 2008.
- [14] C. G. Peng and M. J. Damha, "G-quadruplex induced stabilization by 2'-deoxy-2'-fluoro-d-arabinonucleic acids (2'F-ANA)," *Nucleic Acids Research*, vol. 35, no. 15, pp. 4977–4988, 2007.
- [15] G. X. He, S. H. Krawczyk, S. Swaminathan et al., "N₂- and C₈-substituted oligodeoxynucleotides with enhanced thrombin inhibitory activity in vitro and in vivo," *Journal of Medicinal Chemistry*, vol. 41, no. 13, pp. 2234–2242, 1998.
- [16] V. Esposito, A. Randazzo, G. Piccialli, L. Petraccone, C. Giancola, and L. Mayol, "Effects of an 8-bromodeoxyguanosine incorporation on the parallel quadruplex structure [d(TGG-GT)]₄," *Organic and Biomolecular Chemistry*, vol. 2, no. 3, pp. 313–318, 2004.
- [17] L. Petraccone, I. Duro, A. Randazzo, A. Virno, L. Mayol, and C. Giancola, "Biophysical properties of quadruplexes containing two or three 8-bromodeoxyguanosine residues," *Nucleosides, Nucleotides and Nucleic Acids*, vol. 26, no. 6-7, pp. 669–674, 2007.
- [18] E. Dias, J. L. Battiste, and J. R. Williamson, "Chemical probe for glycosidic conformation in telomeric DNAs," *Journal of the American Chemical Society*, vol. 116, no. 10, pp. 4479–4480, 1997.
- [19] J. Matsui and S. Goji, to be submitted.
- [20] F. Pröll, B. Möhrle, M. Kumpf, and G. Gauglitz, "Label-free characterisation of oligonucleotide hybridisation using reflectometric interference spectroscopy," *Analytical and Bioanalytical Chemistry*, vol. 382, no. 8, pp. 1889–1894, 2005.
- [21] B. P. Möhrle, M. Kumpf, and G. Gauglitz, "Determination of affinity constants of locked nucleic acid (LNA) and DNA duplex formation using label free sensor technology," *Analyst*, vol. 130, no. 12, pp. 1634–1638, 2005.
- [22] J. Piehler, A. Brecht, G. Gauglitz, and M. Zerin, "Label-free monitoring of DNA-ligand interactions," *Analytical Biochemistry*, vol. 249, no. 1, pp. 94–102, 1997.
- [23] T. Hattori, M. Umetsu, T. Nakanishi et al., "High affinity anti-inorganic material antibody generation by integrating graft and evolution technologies: potential of antibodies as biointerface molecules," *Journal of Biological Chemistry*, vol. 285, no. 10, pp. 7784–7793, 2010.
- [24] H. Hasegawa, K. I. Taira, K. Sode, and K. Ikebukuro, "Improvement of aptamer affinity by dimerization," *Sensors*, vol. 8, no. 2, pp. 1090–1098, 2008.
- [25] A. Pastemak, F. J. Hernandez, L. M. Rasmussen, B. Vester, and J. Wengel, "Improved thrombin binding aptamer by incorporation of a single unlocked nucleic acid monomer," *Nucleic Acids Research*, vol. 39, pp. 1155–1164, 2011.
- [26] T. Hianik, V. Ostatná, M. Sonlajtnerova, and I. Grman, "Influence of ionic strength, pH and aptamer configuration for binding affinity to thrombin," *Bioelectrochemistry*, vol. 70, no. 1, pp. 127–133, 2007.
- [27] S. R. Nallagatla, B. Heuberger, A. Haque, and C. Switzer, "Combinatorial synthesis of thrombin-binding aptamers containing iso-guanine," *Journal of Combinatorial Chemistry*, vol. 11, no. 3, pp. 364–369, 2009.
- [28] X. Li, L. Shen, D. Zhang et al., "Electrochemical impedance spectroscopy for study of aptamer-thrombin interfacial interactions," *Biosensors and Bioelectronics*, vol. 23, no. 11, pp. 1624–1630, 2008.

Research Article

Use of a Designed Peptide Library to Screen for Binders to a Particular DNA G-Quadruplex Sequence

Keita Kobayashi,¹ Noriko Matsui,¹ and Kenji Usui^{1,2}

¹ Faculty of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan

² Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan

Correspondence should be addressed to Kenji Usui, kusui@center.konan-u.ac.jp

Received 14 May 2011; Revised 6 July 2011; Accepted 6 July 2011

Academic Editor: Souvik Maiti

Copyright © 2011 Keita Kobayashi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We demonstrated a method to screen for binders to a particular G-quadruplex sequence using easily designed short peptides consisting of naturally occurring amino acids and mining of binding data using statistical methods such as hierarchical clustering analysis (HCA). Despite the small size of the library used in this study, candidates of specific binders were identified. In addition, a selected peptide stabilized the G-quadruplex structure of a DNA oligonucleotide derived from the promoter region of the protooncogene *c-MYC*. This study illustrates how a peptide library can be designed and presents a screening guideline for construction of G-quadruplex binders. Such G-quadruplex peptide binders could be functionally modified to enable switching, cellular penetration, and organelle-targeting for cell and tissue engineering.

1. Introduction

Research over the last few decades has revealed that some DNA and RNA secondary structures modulate a variety of cellular events. One secondary structure in particular, the G-quadruplex [1] regulates cellular events such as transcription, translation, pre-RNA splicing, and telomerase elongation, all of which play roles in various serious diseases and cellular aging [2–6]. Systems capable of controlling DNA and RNA G-quadruplex structures would therefore be useful for the modulation of various cellular events for the purpose of producing biological effects. Because of their biological importance, many G-quadruplex-targeting ligands [7, 8] have been described, including phthalocyanine derivatives [9], porphyrin derivatives [10], and others [11–14]. However, the next generation of binders should have more G-quadruplex sequence specificity, higher inducing or collapsing ability of the structure, and a greater degree of functionality including binding on-off switching, cellular penetration, and the ability to target organelles.

De novo designed peptides (peptides not derived from domains of binding proteins) are promising next generation G-quadruplex binding candidates because of the following advantages they offer: (i) peptides are easier to design and synthesize than antibodies or recombinant proteins; (ii) they can mimic protein-G-quadruplex interactions; (iii) analyses based on peptide libraries can be used to elucidate binding properties of DNAs; (iv) in addition to naturally occurring amino acids, various functional moieties (e.g., artificial amino acids) can be employed as building blocks in designed peptides; (v) because certain peptide sequences may exhibit transmembrane or hormonal properties, combining peptides with these functional sequences with G-quadruplex-binding peptides can produce multifunctional molecular systems useful in cell and tissue engineering.

To increase the utility of the peptide library technology, we designed peptide microarrays composed of various secondary structures. The designed peptide arrays were initially applied to protein analysis [15–21]. Upon addition of various

proteins to these peptide arrays, library peptides containing fluorescent probes showed different binding responses depending on the peptide sequence. These response patterns served as “protein fingerprints” (PPFs), which can be used to establish the identity of a target protein and correlate the recognition properties of a target protein to a particular peptide [15, 16, 21]. In addition, by applying statistical analyses such as hierarchical clustering analysis (HCA) and principal component analysis (PCA) to PPFs, researchers can draw high-confidence correlations between target proteins and biological function, based primarily upon peptide charge and hydrophobicity data [19, 21]. We successfully applied our system to screen for peptide ligands that tightly bind to a target protein and simultaneously control the function of a protein related to the target [20]. This approach has several advantages, such as ease of peptide library design and robust selection of ligands with novel structures for the control of signaling pathways and/or cascades.

Here, we demonstrate a model screening of binders to a particular G-quadruplex sequence using easily designed short peptides consisting of naturally occurring amino acids. We also examined the stability of the DNA G-quadruplex structure upon addition of a G-quadruplex-binding peptide and checked whether the peptides could induce or collapse the G-quadruplex structure. This study illustrates how a peptide library can be designed and presents a screening guideline for the construction of next-generation ligands with increased specificity to particular G-quadruplexes and increased functionality, including on-off switching and the ability to penetrate cells and target organelles.

2. Experimental Methods

2.1. General Remarks. All chemicals and solvents were of reagent or HPLC grade and were used without further purification. Oligodeoxynucleotide samples purified by HPLC were purchased from Hokkaido System Science (Sapporo, Japan). HPLC was performed on a GL-7400 HPLC system (GL sciences, Tokyo, Japan) using an Inertsil ODS-3 (10 × 250 mm; GL Science) column for preparative purification with a linear acetonitrile/0.1% trifluoroacetic acid (TFA) gradient at a flow rate of 3.0 mL/min. Peptides were analyzed using MALDI-TOF MS on an Autoflex III (Bruker Daltonics, Billerica, MA, USA) mass spectrometer with 3,5-dimethoxy-4-hydroxycinnamic acid as the matrix.

2.2. Synthesis of a Designed Peptide Minilibrary. The designed peptide library was synthesized on Fmoc-NH-SAL-PEG Resin (Watanabe Chemical Industries, Hiroshima, Japan) using an automatic synthesizer (Advanced ChemTech Model 357 FBS) with Fmoc chemistry [22] using Fmoc-AA-OH (4 eq., Watanabe Chemical Industries) according to the *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU, Watanabe Chemical) method. Side chain protections were as follows: *t*-butyl (tBu) for Glu and Tyr, trityl (Trt) for His, and *t*-butyloxycarbonyl (Boc) for Lys. The peptides were cleaved from the resin

and side chain protections were removed by incubating the peptides for 2 h in TFA (Watanabe Chemical Industries)/H₂O/triisopropylsilane (Wako Pure Chemical Industries, Tokyo, Japan) (20:1:1, v/v). The peptides were precipitated by addition of cold diethylether and then collected by centrifugation. The recovered peptides were dissolved in water to about 1 mM and stored at 4°C. In order to determine the concentration of library peptide stock solutions, the absorbance at 280 nm (for Trp ($\epsilon = 5500$) and Tyr ($\epsilon = 1490$) [23]) of a diluted solution of each peptide was measured on a U-1900 UV spectrometer (Hitachi High-Technologies, Tokyo, Japan). After a screening assay in 3.2, selected peptides were purified by RP-HPLC and characterized by MALDI-TOF MS for further experiments in 3.3–3.5. Purified peptides were dissolved in water to about 1 mM and their concentrations were measured by their absorbance at 280 nm, after which they were stored at 4°C.

2.3. Screening of Library Peptides Using Gel Electrophoresis. Prior to analysis, each sample in 10 mM Tris-HCl (pH 7.0) buffer containing 0.1 mM EDTA was heated to 85°C for 5 min then gently cooled to room temperature at a rate of 1.0°C min⁻¹. Native gel electrophoresis was performed using nondenaturing gels containing 13% polyacrylamide. Loading buffer (2 μ L) was mixed with 2 μ L of 5 μ M DNA sample with/without 50 μ M library peptide in 10 mM Tris-HCl (pH 7.0) buffer containing 0.1 mM EDTA. A 4 μ L aliquot of each sample was loaded onto the gel and electrophoresed at 10.0 V cm⁻¹ for 2 h at room temperature. Gels were stained with SYBR Gold Nucleic Acid Gel Stain (Invitrogen, Carlsbad, CA, USA) and imaged using a FLA-7000 imager (Fuji Film, Tokyo, Japan). Band intensities were quantified using Malti Gauge software (V.3.2) for Windows. The bound DNA percentage was calculated according to the following equation: $\{(\text{intensity of DNA band without peptide}) - (\text{intensity of DNA band with peptide})\} / (\text{intensity of DNA band without peptide}) \times 100$ (%).

2.4. Preparation of the Color Scale Bar. Data were manipulated according to previously reported standard procedures to prepare the black-red-yellow color scale bar [16]. A portable-pixel-map (.ppm) file format was used. Each grid position was first assigned a three whole-number code based on its RGB color-code response (0, 0, 0 = full black, lowest bound DNA percentage; 255, 0, 0 = red, moderate percentage; 255, 255, 0 = yellow, highest percentage), which corresponds to all the bound DNA percentages divided into 511 levels. The numbers of the grid were saved as a comma-separated-value (.csv) file including the three (or four) lines of the .ppm setting at the top of the file. The file was then saved in the portable-pixel-map format by simply adding a “.ppm” extension to the filename. This file can be opened using graphic viewer software, resized, then saved in another file format, such as bitmap (.bmp).

2.5. Hierarchical Clustering Analysis (HCA). The Euclidean distance [24, 25], a common measure of the distance between two vectors, was used to determine the similarity

G series: H-Z₁X₁Z₂GKWK-NH₂
P series: H-Z₁X₁Z₂PKWK-NH₂
N-terminal C-terminal

X₁ = H, W, F, Y
Z₁ = E, K
Z₂ = E, K

001	E H E G K W K	017	E H E P K W K
002	K H E G K W K	018	K H E P K W K
003	E H K G K W K	019	E H K P K W K
004	K H K G K W K	020	K H K P K W K
005	E W E G K W K	021	E W E P K W K
006	K W E G K W K	022	K W E P K W K
007	E W K G K W K	023	E W K P K W K
008	K W K G K W K	024	K W K P K W K
009	E F E G K W K	025	E F E P K W K
010	K F E G K W K	026	K F E P K W K
011	E F K G K W K	027	E F K P K W K
012	K F K G K W K	028	K F K P K W K
013	E Y E G K W K	029	E Y E P K W K
014	K Y E G K W K	030	K Y E P K W K
015	E Y K G K W K	031	E Y K P K W K
016	K Y K G K W K	032	K Y K P K W K

(a)

		G series X ₁ residue			
		H	W	F	Y
Numbers of amines	3	001	005	009	013
	4	002	006	010	014
	4	003	007	011	015
	5	004	008	012	016

(b)

		P series X ₁ residue			
		H	W	F	Y
Numbers of amines	3	017	021	025	029
	4	018	022	026	030
	4	019	023	027	031
	5	020	024	028	032

(c)

FIGURE 1: (a) Design of the G and P series peptide libraries. (b) Peptide identification number and sequence of each library peptide. (c) Distribution of peptides into the G or P series.

between two binding-color images obtained from different target DNAs. After the similarities between the binding-color images were determined, HCA was conducted. Ward's clustering algorithm was used and the dendrogram was obtained from the Euclidean distances using the Excel Macro program [26]. The horizontal axis represents the distance between vectors (left for high similarity and right for low similarity).

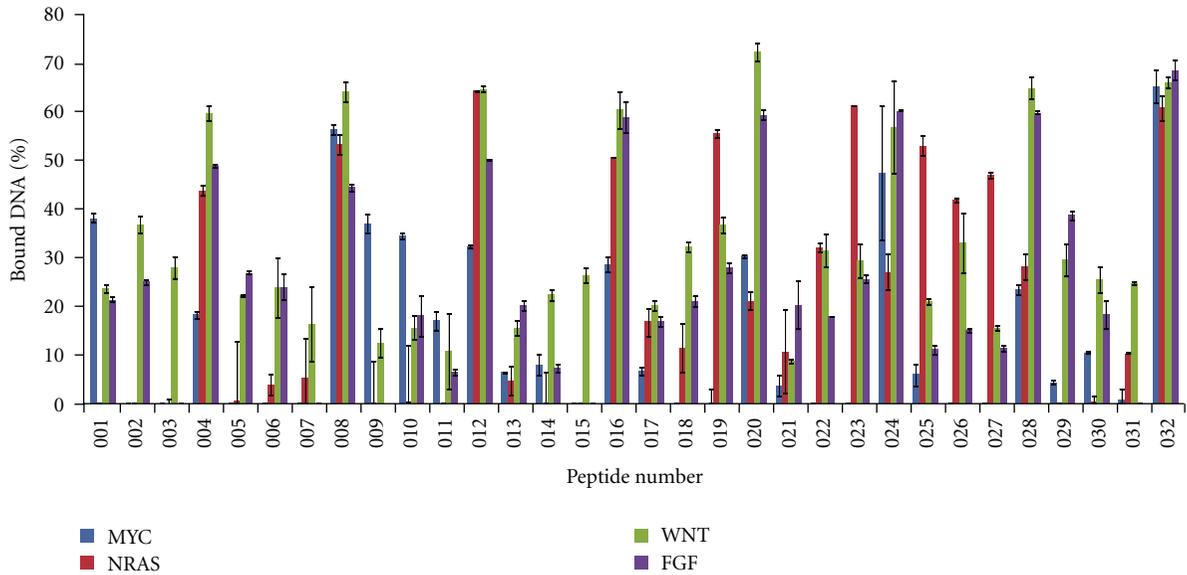
2.6. Circular Dichroism (CD) Spectroscopy. Circular dichroism (CD) spectroscopy was performed using DNA (1 μ M) and peptide no. 010 (0 or 100 μ M) in 20 mM Tris-HCl (pH 7.0) containing 0.1 mM KCl and 0.1 mM EDTA. A J-820 spectropolarimeter (JASCO, Hachioji, Japan) with a thermoregulator using a quartz cell with a 1 cm path length at 25°C was used for CD measurements. Prior to analysis, each sample was heated to 85°C for 5 min, then gently cooled to room temperature at a rate of 1.0°C min⁻¹.

2.7. UV Melting of G-Quadruplex Structures. The UV absorbance was measured using a Shimadzu 1800 spectrophotometer equipped with a temperature controller (Shimadzu, Kyoto, Japan). Melting curves for the G-quadruplex structures were obtained by measuring the UV absorbance at 295 nm in 10 mM Tris-HCl (pH 7.0) containing 0.1 mM KCl and 0.1 mM EDTA at a heating rate of 1.0°C min⁻¹. The melting temperature (T_m) values for 5 μ M DNAs with/without peptide no. 010 (100 μ M) were obtained from UV melting curves as described previously [27, 28].

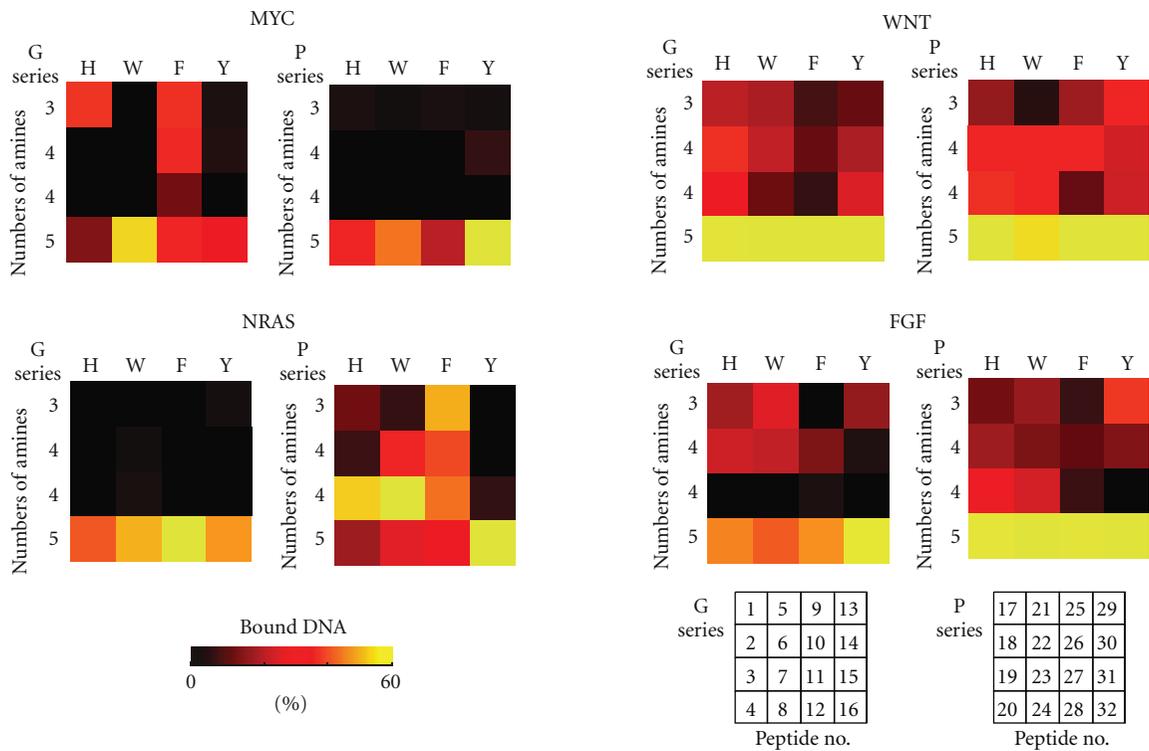
3. Results and Discussion

3.1. Design and Synthesis of the Peptide Minilibrary. We constructed a minilibrary consisting of 32 peptides of varying charge and/or hydrophobicity using the strategy shown in Figure 1(a). The KWK motif is known for its ability to bind to DNA [29]. We designed peptides in which four residues were added to the N-terminus of the KWK motif. Addition of a G or P at 4th residue allowed for varying the flexibility of the peptide main chain. Addition of an H, W, F, or Y at residue X₁ allowed for varying the aromatic character of the peptide, while addition of a K or E at residues Z₁ and Z₂ allowed for varying the peptide charge. Figures 1(b) and 1(c) show all the library peptide species. The column and row headings in Figure 1(c) denote the aromatic residues (H, W, F, or Y at X₁) and numbers of amines (E, or K at Z₁ and Z₂; numbers of amines = 3, 4, or 5), respectively, and each cell displays the number of the synthesized peptide. Hence, the library consists of two series (a G series and a P series) of designed peptides, each of which contains 16 systematically designed peptides.

3.2. Screening of the Designed Peptide Minilibrary and Data Mining. Library peptides were assayed for their ability to bind to four different parallel G-quadruplex sequences from the promoter regions or the 5' untranslated regions of *human* protooncogenes (MYC from *c-MYC*, NRAS from *NRAS*, WNT from *WNT 3*, and FGF from *FGF 3* [30–32] (Table 1)) using electrophoresis. The bound DNA percentages varied according to the peptide sequence (Figure 2),



(a)



(b)



(c)

FIGURE 2: (a) Bound DNA percentages to each library peptide. (b) Bound DNA percentages to each peptide shown as a color illustration. (c) Typical MYC bands in PAGE. MYC alone (A representative of 0% bound DNA percentage), MYC with no. 021 (a representative of 0–10% bound DNA percentage), MYC with no. 020 (a representative of 30–40% bound DNA percentage), and MYC with no. 024 (a representative of 50–70% bound DNA percentage).

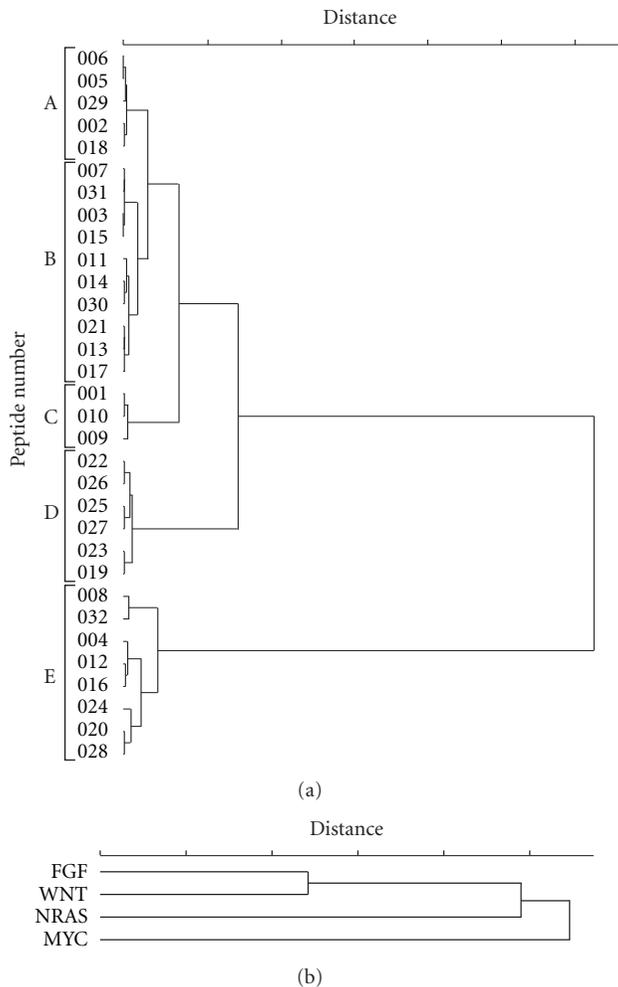


FIGURE 3: (a) Peptide-binding divergence clustering dendrogram generated by Euclidean distance analysis. (b) DNA-binding property clustering dendrogram generated by Euclidean distance analysis.

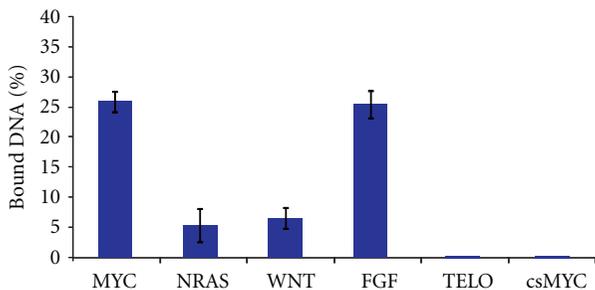


FIGURE 4: Binding percentages of various DNAs with purified peptide no. 010 (100 μM).

and the binding percentages ranged from high to low. To visualize the binding properties more clearly, these results were converted into black-red-yellow images as shown in Figure 2(b). The color images correspond to two series of peptides (Figure 1(c)), and the color of each cell indicates the binding response of each peptide against each DNA.

TABLE 1: Sequences of the G-quadruplex DNAs used in this study.

Oligo name	Gene	Sequence (5'-3')
MYC	<i>c-MYC</i>	AGGGTGGGGAGGGTGGGGA
NRAS	<i>NRAS</i>	GGGAGGGGCGGGTCTGGG
WNT	<i>WNT 3</i>	GGGCCGGGCGGGAGGGG
FGF	<i>FGF 3</i>	GGGGGAGGGGCGGGTAGGG

TABLE 2: Melting temperature (T_m) of G-quadruplex DNAs with/without peptide no. 010.

Oligo name	No. 10 peptide	T_m (°C)
MYC	0 μM	41.3 ± 0.2
	100 μM	49.3 ± 0.3
FGF	0 μM	25.5 ± 0.9
	100 μM	29.2 ± 0.3

The RGB color codes represent the degree of binding: black (lowest percentage), to red (moderate percentage), to yellow (highest percentage), divided into 511 levels. Some peptides, such as nos. 008 and 032, bound to all DNAs tested, while peptide no. 011, for example, little bound to the DNAs tested. The cells in the bottom portion of each series were colored in either yellow or red, indicating that DNAs preferentially bind cationic peptides. In addition, while the MYC G-quadruplex tended to bind to G-series peptides and the NRAS G-quadruplex tended to bind to P-series peptides, the FGF and WNT G-quadruplexes tended to bind to both peptide series. This result implies that imparting variation of flexibility to the peptides by adding a G and P to the middle of the sequence provides DNA selectivity.

Similarities between the peptides in their ability to bind DNAs were analyzed quantitatively using HCA with Euclidean distances. As shown in Figure 3(a), the library peptides could be sorted into five groups (Groups A–E). Group A peptides exhibited a relatively high binding percentage for the FGF and WNT G-quadruplexes, but had less binding percentages for the MYC and NRAS G-quadruplexes. Peptides classified into Group B showed relatively low binding percentages for all DNAs except the WNT G-quadruplex, to which they bound tightly. Peptides classified into Group C had high binding percentages for the MYC G-quadruplex, but lower binding percentages for the NRAS, FGF, and WNT G-quadruplexes. Peptides in Group D demonstrated an intermediate binding percentage for all quadruplexes except MYC, to which they did not bind. Finally, peptides classified into Group E, which are characterized by 5 amines, bound to all the DNAs tested with high binding percentages. Because peptides with 3 or 4 amines demonstrated diverse binding percentages, we concluded that a greater number of specific binders could be obtained by designing and synthesizing a wider array of peptides with 3 or 4 amines. Despite the small size of the library used in this study, we were able to obtain peptides that demonstrated sequence-specific binding. For example, our results show that peptides 009 and 010 are MYC-specific binders, and peptides 003, 007, and 015 are WNT-specific binders.

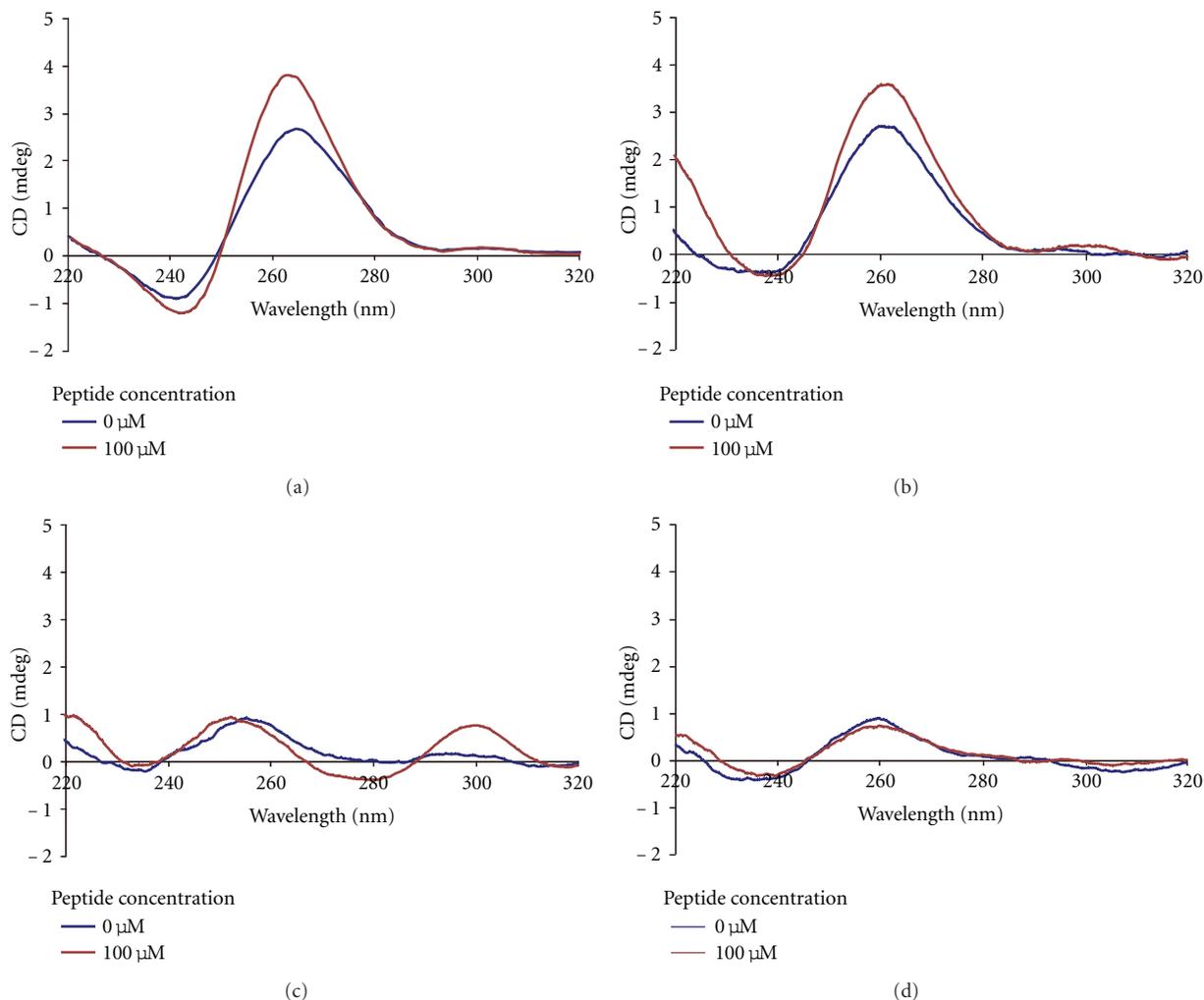


FIGURE 5: CD spectra of (a) MYC, (b) FGF, (c) NRAS, and (d) WNT with/without peptide no. 010.

Similarities between G-quadruplex binding properties were analyzed quantitatively using the HCA method. A clustering dendrogram (Figure 3(b)) was generated by analysis of Euclidean distances. The horizontal axis indicates the distance between the binding percentages of the G-quadruplexes for library peptides (left indicates high similarity and right indicates low similarity). The clustering dendrogram discriminated MYC from the other G-quadruplexes, which tended to cluster. We suspect that some of the library peptides recognized a particular sequence (GGGCGGG) present in all the G-quadruplexes tested except for MYC. Although much additional data regarding the binding of library peptides to various DNA types are needed, these results imply that the statistical approaches used in this study could be used to characterize the binding properties of a variety of other G-quadruplexes.

3.3. Confirmation of DNA Binding. We selected peptide no. 010 as a MYC-specific binder and after the peptide was purified, an electrophoresis assay was conducted to confirm no. 010 binding to MYC, (Though we also selected and

purified peptide no. 009, the purified no. 009 peptide did not strongly bind to MYC (data not shown)). Figure 4 shows the bound DNA percentages at which purified peptide no. 010 bound to various DNAs. Peptide no. 010 strongly bound to MYC, while the peptide weakly bound to NRAS and WNT or little bound to csMYC (complementary sequence of MYC, 5'-TCCCCACCCTCCCCACCCT-3') and TELO (representative antiparallel G-quadruplex sequence from human telomeric DNA, 5'-AGGGTTAGGGTTAGGGTTAGGG-3' [33]). Although G-quadruplex FGF as well as MYC highly bound to peptide no. 010, the data clearly suggest that peptide no. 010 bound only to the parallel G-quadruplex sequence, not to the antiparallel G-quadruplex sequence or to the other sequences, including the MYC complementary sequence. Although additional assays and/or detailed confirmation experiments are needed, these results indicate that the electrophoresis in 3.2 is one of the promising tools for screening DNA-binding peptides using peptide libraries.

3.4. Circular Dichroism (CD) Spectroscopy of G-Quadruplex DNAs with/without Peptide no. 010. We performed CD

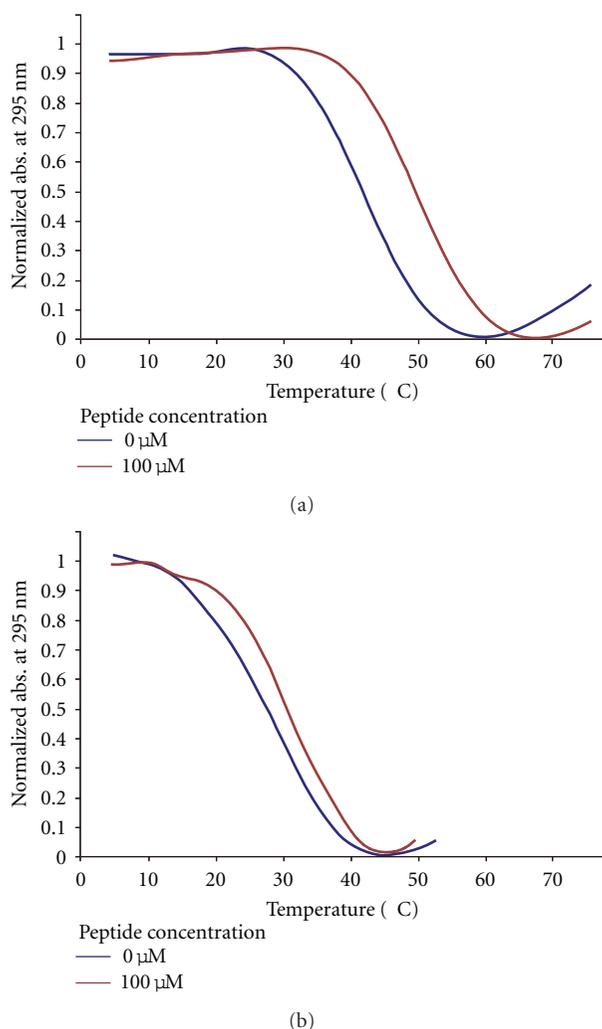


FIGURE 6: Normalized UV melting curves for (a) MYC and (b) FGF with/without 100 μM peptide no. 010.

experiments to investigate induction of structural and conformational changes in the MYC G-quadruplex upon interaction with peptide no. 010. The MYC G-quadruplex yielded spectra that were characteristic of parallel quadruplexes, with a maximum at 260 nm and a minimum at 240 nm (Figure 5(a)), a result that is consistent with a prior study [10]. Interestingly, addition of peptide no. 010 led to an increase in the parallel G-quadruplex signature. Furthermore we performed CD experiments with the other G-quadruplex DNAs upon addition of peptide no. 010. The FGF G-quadruplex showed similar results to those of MYC (Figure 5(b)), whereas the NRAS and the WNT did not show significant changes upon addition of peptide no. 010 (Figures 5(c) and 5(d)). These results were correlated to the DNA-binding results in 3.3.

3.5. Thermodynamic Stability of the DNA G-Quadruplexes Bound to Peptide no. 010. We investigated the effect of peptide binding on the thermodynamic stability of the DNA G-quadruplexes. Table 2 shows the T_m of the MYC

and FGF G-quadruplexes in the presence and absence of 100 μM peptide no. 010. (Melting curves of MYC and FGF with/without no. 010 are shown in Figure 6.) Surprisingly, the MYC G-quadruplex T_m changed appreciably in the presence of the peptide, increasing by about 8°C while the FGF G-quadruplex T_m changed in the presence of the peptide, increasing by only about 4°C.

These results imply that peptide no. 010 binds to the MYC G-quadruplex by a kind of specific association. Although we could not screen peptides with a high binding selectivity, despite the limited size of our library, we found that peptide no. 010 acts as a promising stabilizer of the G-quadruplex structure as well as a binder to MYC.

4. Conclusions

In this study, we demonstrated a novel designed peptide library method to screen for binders to a particular G-quadruplex and also demonstrated the mining of data generated from binding results using statistical methods such as HCA. Our results suggest that the use of a designed peptide library enables the discrimination of G-quadruplex sequences and could therefore provide useful information for the design of peptides for targeting specific G-quadruplexes. Despite the small size of the library used in this study, some candidates of specific binders were identified. By improving the design of the library peptides and the screening methods, our system could be used to screen for peptides that bind to a particular G-quadruplex and alter its thermodynamic properties. It would then be possible to find binders with strong specificity to which specific functional attributes can be added, such as the ability to penetrate cells in order to control DNA and/or RNA events for the purposes of cell and tissue engineering.

Acknowledgments

This study was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT). K. Usui is grateful to the Grants-in-Aid for Scientific Research, the “Core research” project (2009–2014) from MEXT.

References

- [1] M. Gellert, M. N. Lipsett, and D. R. Davies, “Helix formation by guanylic acid,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 48, no. 12, pp. 2013–2018, 1962.
- [2] E. Henderson, C. C. Hardin, S. K. Walk, I. Tinoco Jr., and E. H. Blackburn, “Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs,” *Cell*, vol. 51, no. 6, pp. 899–908, 1987.
- [3] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, and L. H. Hurley, “Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 18, pp. 11593–11598, 2002.
- [4] S. Rankin, A. P. Reszka, J. Huppert et al., “Putative DNA quadruplex formation within the human c-kit oncogene,” *Journal*

- of the *American Chemical Society*, vol. 127, no. 30, pp. 10584–10589, 2005.
- [5] D. Miyoshi, S. Matsumura, S. I. Nakano, and N. Sugimoto, “Duplex dDissociation of telomere DNAs induced by molecular crowding,” *Journal of the American Chemical Society*, vol. 126, no. 1, pp. 165–169, 2004.
 - [6] H. Q. Yu, D. Miyoshi, and N. Sugimoto, “Characterization of structure and stability of long telomeric DNA G-quadruplexes,” *Journal of the American Chemical Society*, vol. 128, no. 48, pp. 15461–15468, 2006.
 - [7] A. De Cian, L. Lacroix, C. Douarre et al., “Targeting telomeres and telomerase,” *Biochimie*, vol. 90, no. 1, pp. 131–155, 2008.
 - [8] L. H. Hurley, “DNA and its associated processes as targets for cancer therapy,” *Nature Reviews Cancer*, vol. 2, no. 3, pp. 188–200, 2002.
 - [9] H. Yaku, T. Murashima, D. Miyoshi, and N. Sugimoto, “Anionic phthalocyanines targeting G-quadruplexes and inhibiting telomerase activity in the presence of excessive DNA duplexes,” *Chemical Communications*, vol. 46, no. 31, pp. 5740–5742, 2010.
 - [10] J. Seenisamy, E. M. Rezler, T. J. Powell et al., “The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4,” *Journal of the American Chemical Society*, vol. 126, no. 28, pp. 8702–8709, 2004.
 - [11] S. Roy, F. A. Tanious, W. D. Wilson, D. H. Ly, and B. A. Armitage, “High-affinity homologous peptide nucleic acid probes for targeting a quadruplex-forming sequence from a MYC promoter element,” *Biochemistry*, vol. 46, no. 37, pp. 10433–10443, 2007.
 - [12] S. Müller, G. D. Pantoş, R. Rodriguez, and S. Balasubramanian, “Controlled-folding of a small molecule modulates DNA G-quadruplex recognition,” *Chemical Communications*, no. 1, pp. 80–82, 2009.
 - [13] T. K. Chakraborty, A. Arora, S. Roy, N. Kumar, and S. Maiti, “Furan based cyclic oligopeptides selectively target G-quadruplex,” *Journal of Medicinal Chemistry*, vol. 50, no. 23, pp. 5539–5542, 2007.
 - [14] J. A. Schouten, S. Ladame, S. J. Mason, M. A. Cooper, and S. Balasubramanian, “G-quadruplex-specific peptide-hemicyanine ligands by partial combinatorial selection,” *Journal of the American Chemical Society*, vol. 125, no. 19, pp. 5594–5595, 2003.
 - [15] K. Usui, T. Ojima, M. Takahashi, K. Nokihara, and H. Mihara, “Peptide arrays with designed secondary structures for protein characterization using fluorescent fingerprint patterns,” *Biopolymers*, vol. 76, no. 2, pp. 129–139, 2004.
 - [16] K. Usui, M. Takahashi, K. Nokihara, and H. Mihara, “Peptide arrays with designed α -helical structures for characterization of proteins from FRET fingerprint patterns,” *Molecular Diversity*, vol. 8, no. 3, pp. 209–218, 2004.
 - [17] K. Y. Tomizaki, K. Usui, and H. Mihara, “Protein-detecting microarrays: current accomplishments and requirements,” *ChemBioChem*, vol. 6, no. 5, pp. 782–799, 2005.
 - [18] K. Usui, K. Y. Tomizaki, T. Ohyama, K. Nokihara, and H. Mihara, “A novel peptide microarray for protein detection and analysis utilizing a dry peptide array system,” *Molecular BioSystems*, vol. 2, no. 2, pp. 113–121, 2006.
 - [19] K. Usui, K. Y. Tomizaki, and H. Mihara, “Protein-fingerprint data mining of a designed α -helical peptide array,” *Molecular BioSystems*, vol. 2, no. 9, pp. 417–420, 2006.
 - [20] K. Usui, K. Y. Tomizaki, and H. Mihara, “Screening of α -helical peptide ligands controlling a calcineurin-phosphatase activity,” *Bioorganic and Medicinal Chemistry Letters*, vol. 17, no. 1, pp. 167–171, 2007.
 - [21] K. Usui, K. Y. Tomizaki, and H. Mihara, “A designed peptide chip: protein fingerprinting technology with a dry peptide array and statistical data mining,” *Methods in Molecular Biology*, vol. 570, pp. 273–284, 2009.
 - [22] W. C. Chan and P. D. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, New York, NY, USA, 2000.
 - [23] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, and T. Gray, “How to measure and predict the molar absorption coefficient of a protein,” *Protein Science*, vol. 4, no. 11, pp. 2411–2423, 1995.
 - [24] J. Grognum and J. L. Reymond, “Classifying enzymes from selectivity fingerprints,” *ChemBioChem*, vol. 5, no. 6, pp. 826–831, 2004.
 - [25] J. P. Goddard and J. L. Reymond, “Enzyme activity fingerprinting with substrate cocktails,” *Journal of the American Chemical Society*, vol. 126, no. 36, pp. 11116–11117, 2004.
 - [26] <http://aoki2.si.gunma-u.ac.jp/lecture/stats-by-excel/vba/html/clustan.html>.
 - [27] N. Sugimoto, S. I. Nakano, M. Katoh et al., “Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes,” *Biochemistry*, vol. 34, no. 35, pp. 11211–11216, 1995.
 - [28] S. I. Nakano, T. Kanzaki, and N. Sugimoto, “Influences of ribonucleotide on a duplex conformation and its thermal stability: study with the chimeric RNA-DNA strands,” *Journal of the American Chemical Society*, vol. 126, no. 4, pp. 1088–1095, 2004.
 - [29] N. P. Johnson, H. Mazarguil, and A. Lopez, “Strandedness discrimination in peptide-polynucleotide complexes,” *Journal of Biological Chemistry*, vol. 271, no. 33, pp. 19675–19679, 1996.
 - [30] D. J. Patel, A. T. Phan, and V. Kuryavyi, “Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics,” *Nucleic Acids Research*, vol. 35, no. 22, pp. 7429–7455, 2007.
 - [31] S. Kumari, A. Bugaut, J. L. Huppert, and S. Balasubramanian, “An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation,” *Nature Chemical Biology*, vol. 3, no. 4, pp. 218–221, 2007.
 - [32] N. Kumar and S. Maiti, “A thermodynamic overview of naturally occurring intramolecular DNA quadruplexes,” *Nucleic Acids Research*, vol. 36, no. 17, pp. 5610–5622, 2008.
 - [33] Y. Wang and D. J. Patel, “Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex,” *Structure*, vol. 1, no. 4, pp. 263–282, 1993.

Research Article

The Stability of a Model Substrate for Topoisomerase 1-Mediated DNA Religation Depends on the Presence of Mismatched Base Pairs

William H. Gmeiner,¹ Freddie Salsbury Jr.,² Chris M. Olsen,³ and Luis A. Marky³

¹ Department of Cancer Biology, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA

² Department of Physics, Wake Forest University, Winston-Salem, NC 27109, USA

³ Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, NE 68198, USA

Correspondence should be addressed to William H. Gmeiner, bgmeiner@wakehealth.edu

Received 9 May 2011; Revised 7 June 2011; Accepted 11 June 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 William H. Gmeiner et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Topoisomerase 1 (Top1) enzymes regulate DNA superhelicity by forming covalent cleavage complexes that undergo controlled rotation. Substitution of nucleoside analogs at the +1 position of the DNA duplex relative to the Top1 cleavage site inhibits DNA religation. The reduced efficiency for Top1-mediated religation contributes to the anticancer activity of widely used anticancer drugs including fluoropyrimidines and gemcitabine. In the present study, we report that mismatched base pairs at the +1 position destabilize the duplex DNA components for a model Top1 cleavage complex formation even though one duplex component does not directly include a mismatched base pair. Molecular dynamics simulations reveal G-dU and G-FdU mismatched base pairs, but not a G-T mismatched base pair, increase flexibility at the Top1 cleavage site, and affect coupling between the regions required for the religation reaction to occur. These results demonstrate that substitution of dT analogs into the +1 position of the non-scissile strand alters the stability and flexibility of DNA contributing to the reduced efficiency for Top1-mediated DNA religation. These effects are inherent in the DNA duplex and do not require formation of the Top1:DNA complex. These results provide a biophysical rationale for the inhibition of Top1-mediated DNA religation by nucleotide analog substitution.

1. Introduction

DNA topoisomerases regulate the topological state of DNA as required to relieve superhelical density for important biological processes such as replication and transcription [1–3]. DNA topoisomerase 1 (Top1) is expressed at elevated levels during S-phase of the cell cycle and is the topoisomerase primarily responsible for relieving superhelical density generated in front of advancing replication forks in mammalian cells. Top1 preferentially binds superhelical DNA and forms a covalent complex as a result of nucleophilic attack by the hydroxyl of Tyr 723 on the phosphodiester backbone of the scissile strand of the DNA duplex. DNA superhelical density is reduced by controlled rotation of the scissile strand about the non-scissile strand in the cleavage complex [4, 5]. Following release of superhelical tension, the

cleavage complex is dissociated by nucleophilic attack of the free 5'-OH of the scissile strand to reform the phosphodiester backbone. DNA sequences that have several A-tracts flanking a conserved DNA duplex motif are also substrates for DNA Top1 and serve as a model system for understanding DNA recognition and catalysis by Top1 [6].

Top1 is the sole target for the camptothecin (CPT) class of anticancer drugs. CPT forms a stable ternary complex upon binding to the Top1:DNA covalent cleavage complex. Stabilization of cleavage complexes by CPT converts Top1 into a cellular poison since collision of advancing replication forks with trapped Top1 cleavage complexes results in DNA double-strand breaks. Thus, CPT not only inhibits Top1 activity, but also converts the enzymatic activity into DNA damage that is potentially lethal to the cell. Over the last decade, it has been shown that a variety of nonnative

nucleotide substitutions that may result from oxidative damage to DNA (e.g., 8-oxo-dG) or covalent modification of DNA nucleobases (e.g., benzpyrene adducts) also cause trapping of Top1 cleavage complexes and result in DNA DSB formation [7]. Work from our laboratory in collaboration with the Pommier lab has shown that misincorporation of deoxyribonucleotide analogs that have anticancer activity, such as FdU [8] and gemcitabine [9], into Top1 cleavage sites also causes trapping of Top1 cleavage complexes [10]. Poisoning of Top1 by FdU-substituted DNA contributes to the cytotoxicity and antitumor activities of fluoropyrimidines [8].

The structural basis for trapping of Top1 cleavage complexes by damaged nucleobases or misincorporation of nucleotide analogs into the non-scissile strand of DNA remains an area of investigation. Although the DNA sequence used in most model studies of Top1:DNA interactions contains several A-tracts [6], X-ray crystal structures do not reveal any bending of this DNA in either covalent or noncovalent complexes with DNA [5]. One question that remains unanswered is how introduction of nonnative nucleotides into the non-scissile strand of DNA inhibits the religation reaction. To investigate this issue, we have constructed a model Top1 cleavage site consisting of a 39 mer DNA hairpin consisting of 13 base pairs with a 10 mer single-stranded overhang (Figure 1). We have investigated the thermal stability of this DNA hairpin consisting of all native nucleotides and have compared the stability of the native sequence to sequences that contain a single C → dU, C → FdU (5-fluoro-2'-deoxyuridine), or C → T substitution. These substitutions result in a single mismatched base pair at the site corresponding to the +1 site relative to the site of Top1 cleavage. As expected, introduction of a mismatched base pair decreases the stability of the DNA hairpin by approximately 3°C. We also investigated the stability of DNA duplex formation between these DNA hairpins and a DNA 10 mer that is complementary to the 10-nucleotide overhang of the hairpin. The interaction of this DNA 10 mer with the hairpin provides a model system for the Top1 religation reaction. Unexpectedly, we find that the melting temperature for the formation of the 10 mer duplex that would be required for the Top1 religation reaction to occur is sensitive to the presence of DNA mismatched base pairs in the hairpin even though no mismatched base pairs are present in the 10 mer duplex region. Molecular dynamics simulations of this model system demonstrate that G-dU and G-FdU mismatched base pairs increase flexibility and affect coupling with the first 10 base pairs. We conclude that DNA mismatched base pairs adjacent to the Top1 cleavage site both decrease DNA stability and increase flexibility disfavoring formation of the DNA conformation required for Top1-mediated DNA religation.

2. Methods

2.1. Design and Synthesis of DNA Hairpin Sequences. A model Top1 cleavage site (Figure 1) was designed based upon a DNA sequence containing several A-tract motifs

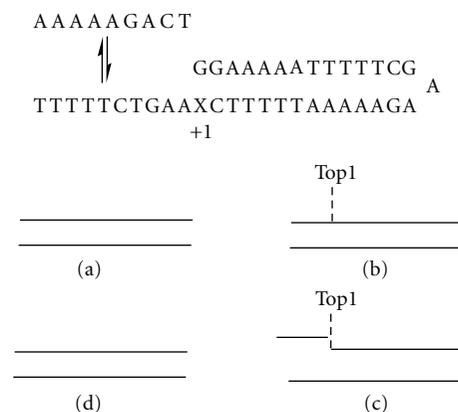


FIGURE 1: Model system used for evaluating the effects of mismatched base pairs on the stability of the intermediate required for Top1-mediated DNA religation. Top: depiction of the equilibrium between the 10 mer single-stranded DNA and the 39 mer DNA hairpin that constitutes a model system for Top1-mediated DNA religation. A G-X base pair is at the +1 position relative to the Top1 cleavage site. Bottom (a-d): depiction of the steps of Top1 binding, strand dissociation, strand reassociation, and religation.

that had been previously demonstrated to be a suitable model substrate for the Top1 cleavage/religation reaction [6]. The model Top1 cleavage site was synthesized in two pieces, a 39 mer DNA hairpin containing the GAA sequence [11] that promotes hairpin formation and a 10-nucleotide single-strand DNA complementary to the 3'-terminus of the hairpin. Upon annealing, the two sequences form a duplex with a single nick in the phosphodiester backbone of one strand corresponding to the scissile strand of the Top1 cleavage complex (Figure 1). To investigate the effects of deoxynucleotide substitutions on the stability of the model Top1 cleavage site, the 39 mer was synthesized four times, once with dC as the putative base pairing partner for the 3'-terminal dG and also with T, dU, or FdU at this site. In this manner, the base pair at the +1 site of the model Top1 cleavage complex was either the native G-C base pair or was a G-T, G-FdU, or G-dU mismatched base pair. These mismatched base pairs occur at the junction corresponding to the site of cleavage for the Top1 complex. All DNA sequences were synthesized at the University of Calgary DNA core synthesis facility and purified by gel filtration chromatography.

2.2. Temperature-Dependent UV Spectroscopy (UV Melts). Absorbance versus temperature profiles of each oligonucleotide in buffer were measured at 260 nm using a thermoelectrically controlled Aviv model 14DS UV-vis spectrophotometer (Lakewood, NJ). The temperature was scanned from 20°C to 95°C for the 39 mer DNA hairpins and from 1–100°C for the model Top1 cleavage sites at a heating rate of 0.6°C/min. DNA concentrations were 1.5–2.0 μM, and the buffer used was 10 mM sodium phosphate, pH 7.0 with 200 mM NaCl added for high-salt conditions. Shape analysis of the melting curves yielded T_m values using procedures reported earlier [12].

2.3. Molecular Modeling and Molecular Dynamics Simulations. The simulation of the four hairpins is performed using NAMD [13] with the CHARMM27 force field [14, 15], with analysis performed using CHARMM [16]. A normal, matched DNA structure was built using PREDICTOR [17] to construct 23 base pairs using the same sequence as in the experiments. The GAA hairpin was added from a high-resolution NMR structure [18]. Coordinate manipulation commands within CHARMM were used to merge the two structural elements. The resulting overall structure was minimized in CHARMM with an r -dependent dielectric of 4 r , and harmonic restraints to remove bad contacts. The minimization cycle was (1) 100 steps of steepest-descent minimization followed by 100 steps of conjugate gradient minimization both with best-fit harmonic constraints on the hairpin atoms with a mass-weighted force constant of 1 with the remaining bases fixed to relieve any bad contacts with the hairpin; (2) 100 steps of steepest-descent minimization followed by 100 steps of conjugate gradient minimization both with harmonic constraints on all atoms with a mass-weighted force constant of 10; (3) 100 steps of steepest-descent minimization followed by 100 steps of conjugate-gradient minimization both with harmonic constraints on all atoms with a mass-weighted force constant of 1. The other three hairpins were built by mutating the original matched structure and rebuilding the altered base within CHARMM followed by the same minimization protocol. Missing parameters for FdU were obtained from our previous quantum mechanical study [19] and the existing fluorine parameters within CHARMM27 supplemented by two dihedrals created using the parameters from our previous study [19] with force constants from the corresponding unperturbed dihedrals. The resulting structures were fully solvated and charge neutralized with TIP3P water in a cubic box using the visual molecular dynamics (VMD) package [20]. The simulation was performed in NAMD using [13] standard parameters: a 2.0 fs timestep using SHAKE on all bonds to hydrogen atoms, a 12 Å cutoff, Particle Mesh Ewald with a 1.0 Å grid determined by NAMD, Langevin constant pressure algorithm with a target pressure of 1.01325 bar, a piston period of 100 fs, a piston decay time of 50/ps, and a piston temperature of 300 K, all as implemented in NAMD. The simulation protocol consisted of 1000 steps of unconstrained steepest-descent minimization on the fully solvated and ionized system, followed by 250 ps of thermal equilibration to 300 K with temperature reassignment, followed by a 8–12 ns of production simulation; 10 different simulations with different random initial conditions were performed for each hairpin for a total of 40 simulations. Based on exchange between clusters after 4 ns all-atom clustering and leveling of the all-atom RMSDs (data not shown), the first 4 ns of each simulation was discarded as equilibration. The remaining 240 ns in total of simulation data was analyzed with structures saved every 2 ps. CHARMM’s analysis routines were used to calculate the root-mean-square fluctuations, covariances, and for rmsd-based clustering with a 2.5 Å cutoff. Matlab was used to perform the clustering analysis of the data with single-linkage Euclidean distance measures and an inconsistent value of 0.001.

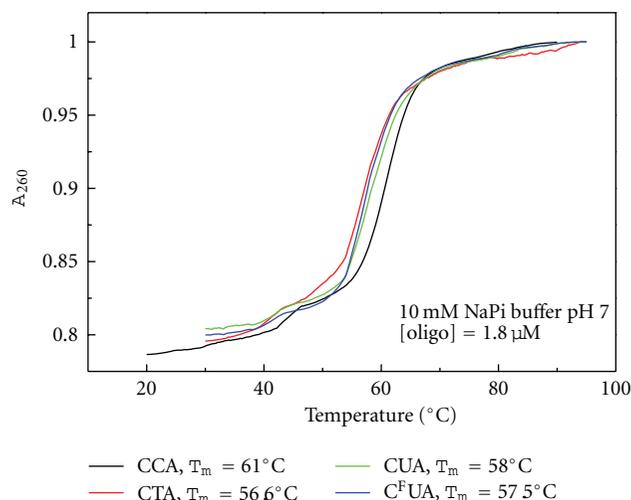


FIGURE 2: Stability of the 39 mer DNA hairpin corresponding to a model Top1 cleavage site. A single coordinated transition is observed corresponding to the thermal melting transition for the double-stranded region of the hairpin. As expected, the melting temperature is dependent upon the presence of a mismatched base pair at the terminus of the duplex region. The native sequence is most stable while C → T, C → dU, and C → FdU decrease the stability of the duplex region by 4.4°C, 3.0°C, and 3.5°C, respectively.

3. Results

3.1. UV Hyperchromicity Measurements. The effects of mismatched base pairs at the +1 position on the stability of the model Top1 cleavage site were investigated using UV hyperchromicity measurements. Initial studies focused on the stabilities of the four 39 mer DNA hairpins (Figure 2). The parental hairpin consisting of only Watson-Crick base pairs in the stem region had a T_m of 61.0°C. Introduction of a G-T mismatched base pair reduced the stability of the hairpin by 4.4°C. The G-dU and G-FdU mismatched base pairs also destabilized the hairpin decreasing the T_m by 3.0 and 3.5°C, respectively. The results demonstrate that introduction of a single mismatched base pair at the +1 site relative to Top1 cleavage destabilizes the duplex by 3–4.4°C. The results are consistent with previous studies indicating similar destabilization effects resulting from introduction of a single mismatched base pair.

We next investigated the stability of the model Top1 cleavage complex using UV hyperchromicity measurements. The four 39 mer DNA hairpins consisting of 13 base pairs with a 10 mer single-stranded overhang were annealed to the 10 mer ssDNA complementary to the overhang region (Figure 1). The single mismatched base pair corresponding to the +1 site of a Top1 cleavage complex was at the stem terminus, such that annealing of the 10 mer to the overhang extended the stem, albeit with a nicked phosphodiester backbone. The UV hyperchromicity profiles for all four hairpins in the presence of the complementary 10 mer were biphasic, as expected, with release of the 10 mer occurring first followed by melting of the 39 mer hairpins (Figure 3).

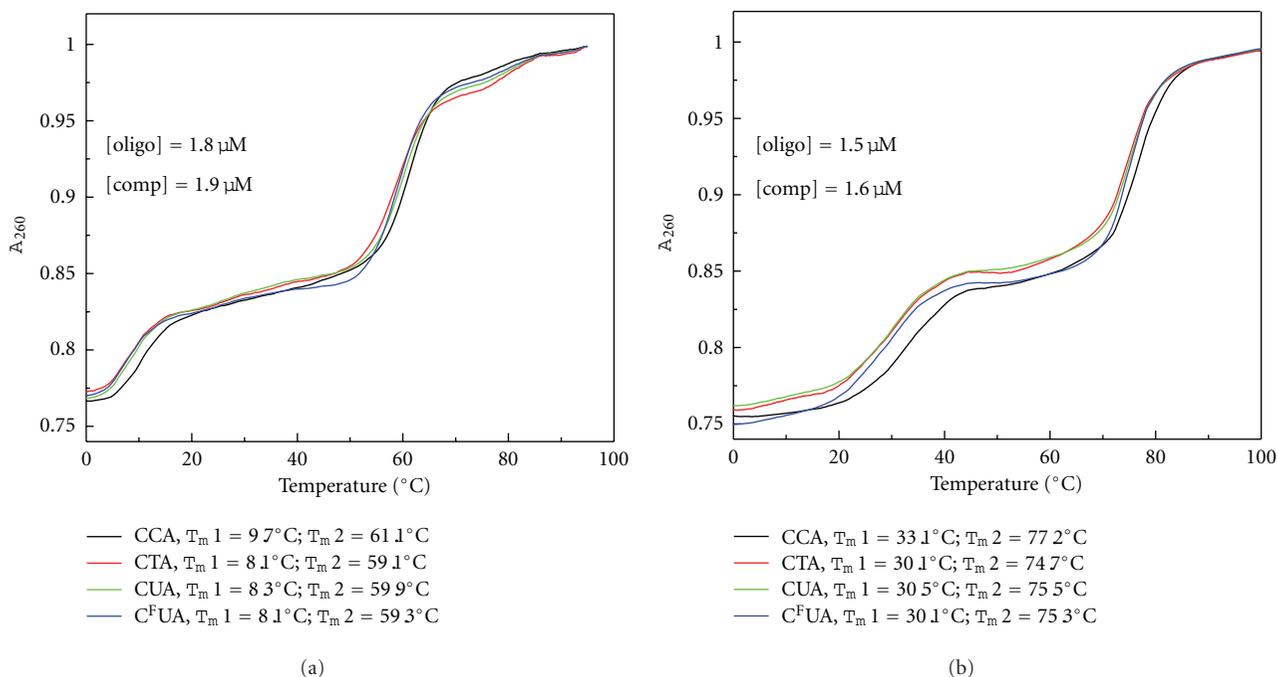


FIGURE 3: Thermal stability of the complex of the 10 mer ssDNA and the 39 mer DNA hairpin constituting a model Top1 religation complex. (a) Melting transition in 10 mM sodium phosphate buffer pH 7; (b) melting transition in the same buffer as A, but with 200 mM NaCl. A biphasic melting profile is observed in both instances. The lower melting temperature corresponds to the dissociation of the ssDNA from the 39 mer DNA hairpin, while the higher melting transition corresponds to the melting temperature for the duplex region of the hairpin. As was observed for the DNA hairpin alone (Figure 2), the melting temperature for the duplex region is sensitive to the presence of a mismatched base pair at the terminus. Surprisingly, the melting temperature for the complex between the single-stranded DNA 10 mer and the single-stranded region of the hairpin is also sensitive to the presence of a base pair mismatch in the duplex region even though the transition does not directly involve the mismatched base pair.

Under low salt conditions (10 mM phosphate buffer, pH 7), the melting temperatures for the two phases of the parental hairpin were 9.7°C and 61.1°C corresponding to release of the 10 mer and unfolding of the hairpin. The identical melting curve obtained in physiologically relevant salt conditions (10 mM phosphate buffer, pH 7, 200 mM NaCl) had a T_m of 33.1°C for release of the 10 mer and a T_m of 77.2°C for unfolding of the hairpin. Interestingly, the presence of the mismatched base pairs destabilized both components of the biphasic melting curves. For example, for the physiologically relevant salt conditions with the G-FdU base pair, the T_m for release of the 10 mer decreased from 33.1 to 30.1°C, and the T_m for unfolding of the hairpin decreased from 77.2 to 75.3°C. The decrease for the 39 mer hairpin is expected as the mismatched base pair is present in the stem of the hairpin, and the magnitude of destabilization is similar to that observed for the 39 mer hairpin alone. The destabilization of the 10 mer region is somewhat surprising as the sequence does not contain a mismatched base pair. The terminal base pair of the 10 mer, however, stacks upon the site of the mismatched base pair in the 39 mer hairpin, and the observed destabilization likely results from less efficient stacking interactions. The G-T and G-dU mismatched base pairs elicited similar degrees of destabilization as the G-FdU mismatch for both release of the 10 mer and melting of the DNA hairpin. The results indicate that the presence of

a mismatched base pair decreases the stability of the model Top1 cleavage complex disfavoring adoption of the geometry required for religation to occur. The results are consistent with decreased stability of the duplex as contributing to the less favorable religation kinetics observed for model Top1 cleavage sites containing G-FdU or other mismatched base pairs [8].

3.2. Molecular Dynamics Simulations. Comparing molecular dynamics simulations for the four different 49 mer DNA hairpins (39 mer DNA hairpin extended without a nick in the phosphodiester backbone) demonstrates that all the mismatches indeed do produce profound effects on the initial 10 base pairs that form the recognition sequence and its complement. Due to the observed effects on the thermal stability and the novel influence of different mismatches occurring outside the recognition sequence, three different atomic measures of structural fluctuations within these first 10 base pairs are used to quantify this influence. First, atomic root-mean-square fluctuations (RMSFs) are calculated for all heavy atoms and averaged on a per-base level. This quantifies the extent to which each atom fluctuates about its equilibrium position and averages these fluctuations at the base-level for comparisons to determine how the different mismatches affect both the overall atomic fluctuations of the first 10 base pairs and the pattern of fluctuations. Second,

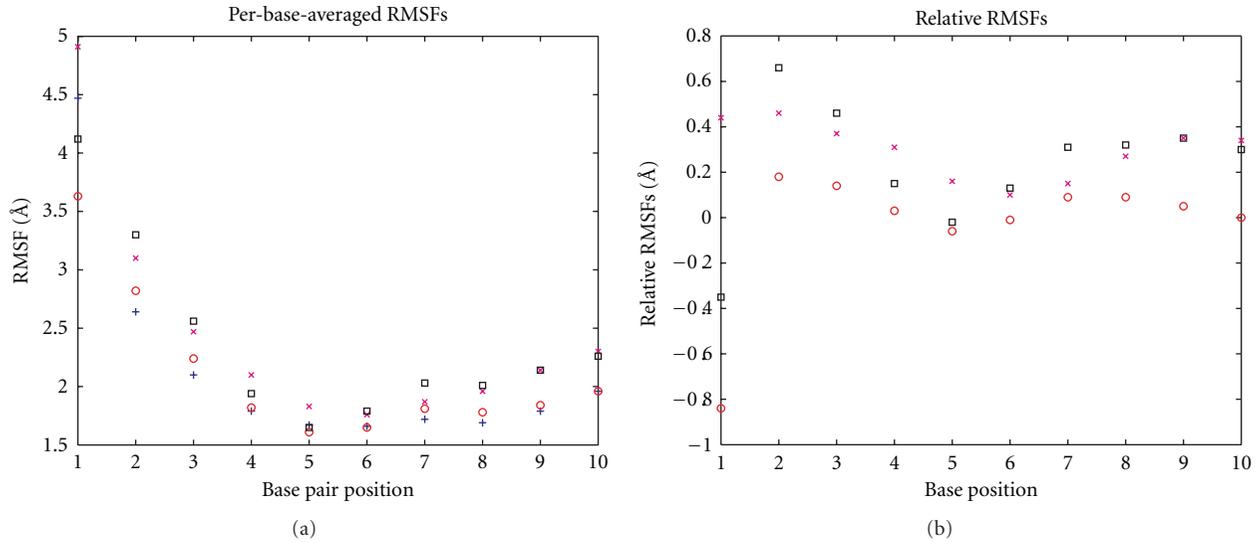


FIGURE 4: (a) Per-base-averaged RMSFs. The atomic root-mean-square fluctuations for the first ten bases for each of the DNA sequences are depicted: matched DNA as blue +s, T mismatched as red os, dU mismatched as magenta \times s, and FdU mismatched as black squares. (b) Per-base-averaged RMSFs relative to matched DNA as (a), but with the matched DNA rmsf values subtracted.

correlated fluctuations are calculated. These determine the extent to which atomic fluctuations, regardless of magnitude, are correlated or anticorrelated and are averaged at the base level for comparison among the four different simulation types. This measure, referred to as the covariance matrix, determines how the different mismatches affect the coupling within the first 10 base pairs. This method has been used by multiple research groups to analyze communication within proteins [21] and the effects of various perturbations on communication [22]. Performing the same analysis on DNA can provide similar information. The third and final measure used is cluster analysis, in which the different structural snapshots across all the simulations are clustered together to find the conformations accessed and their populations in the different hairpins.

The first noticeable effect of mismatch formation is on the RMSFs of the G-dU and G-FdU hairpins; the average RMSF over the first 10 base pairs increases by 13.5% and 12.0% for the G-dU and G-FdU hairpins, respectively, relative to the average RMSF of the matched hairpin (Table 1). Surprisingly, the G-T mismatch shows almost no overall increase in flexibility ($<1\%$). However, all the hairpins show delocalized changes in flexibility in the first 10 bases relative to the matched hairpin (Figure 4).

The covariance matrices (Figure 5) also exhibit delocalized changes in dynamics, as measured by correlated fluctuations, mostly concentrated within the chain where the mismatches occur (bases 1–10). The largest such changes occur with the FdU mismatch (Figure 5(c)), although more pairs exhibit large perturbations with the dU mismatch (Figure 5(d)), including bases in the opposite chain near the mismatch.

Both the RMSFs and covariances demonstrate that there are delocalized changes in atomic covariances and their correlations in the recognition sequences despite the mismatch

TABLE 1: DNA RMSFs (\AA)*.

Matched	2.136
G-T	2.120
G-dU	2.423
G-FdU	2.391

*The root-mean-square fluctuations in \AA for the first 10 base pairs of each sequence are averaged.

TABLE 2: Cluster populations (%).

Cluster no./Sequence	Matched	G-T	G-dU	G-FdU
1	29.0%	28.5%	14.2%	32.9%
2	37.4%	30.1%	32.7%	20.2%
3	24.8%	22.6%	27.8%	20.4%
4	8.8%	18.9%	25.3%	26.4%

occurring outside this sequence; clustering analysis sheds light on the population shifts and conformational changes that may give rise to these variations, and that may perturb the recognition and binding by Top1.

Clustering analysis on the first 10 base pairs shows that there are four different conformations that are accessible to each of the four DNA sequences. The difference lies in their populations (Table 2). The different mismatches shift population among the different conformations, with the dominate change being increasing population in the rarest conformation found in the normal, matched DNA simulation. The population of this conformation increases in the different mismatched from T, to dU and to FdU.

The actual structural rearrangements that occur during these conformational changes are actually quite modest

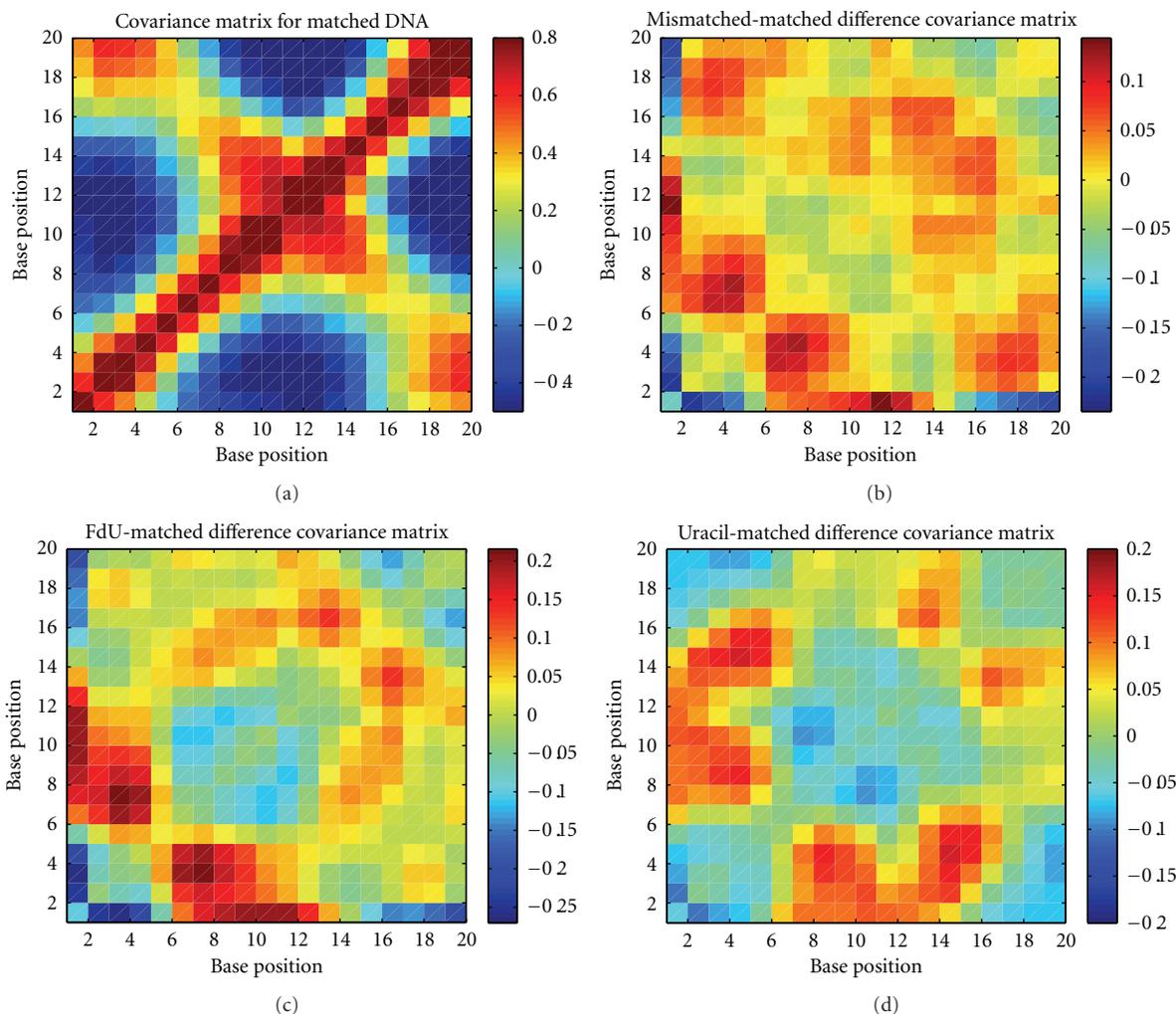


FIGURE 5: (a) Matched covariance matrix. The per-base-averaged covariance matrix plotted as a color map with red indicating high correlations of fluctuations blue indicating high anticorrelation of fluctuations, and yellow indicating uncorrelated. (b) Mismatched-matched difference covariance matrix. The matched covariance matrix (a) subtracted from the mismatched covariance matrix (not shown). (c) FdU-matched difference covariance matrix. The matched covariance matrix (a) subtracted from the FdU-mismatch covariance matrix (not shown). (d) Uracil-matched difference covariance matrix. The matched covariance matrix (a) subtracted from the uracil-mismatch covariance matrix (not shown).

(Figure 6), as is expected in a system as structured as duplex DNA. The different mismatches induce subtle shifts in the base backbone and the base interactions these shifts are especially large at the position just downstream from the mismatch (Figure 6(b)), and especially so in the conformation that is rarest in the matched DNA and dominate in the FdU mismatch (Figure 6(c)).

One issue that can also be addressed is that of the classification of the different DNA sequences, which ones are more similar, and which are more different. With three different measures, the simplest approach to this problem is to cluster the four sequences based on each of these three measures (Table 3). By RMSE, two clusters emerge, one consisting of matched and mismatched, and one of FdU and dU; by cluster population, three clusters emerge, one consisting of matched and mismatched, one of FdU, and one with dU. However, when clustering is based on covariance

matrices, although three clusters emerge, the mismatched and FdU sequences are in one cluster with the other two as singleton clusters. These results suggest that the mismatched and matched sequences are most similar, although with the mismatched more similar to FdU in one dynamical measure (covariances), and FdU is most similar to either dU or mismatched depending on the measure used.

4. Discussion

DNA mismatched base pairs occur at high levels in cells both as a consequence of base damage (e.g., cytidine deamination) [23] as well as errors during DNA replication. Cancer cells frequently have defects in DNA repair processes that may result in greater levels of DNA mismatched base pairs being present in the replicated genome. Further, treatment with fluoropyrimidine drugs such as 5FU or FdUMP [10] results

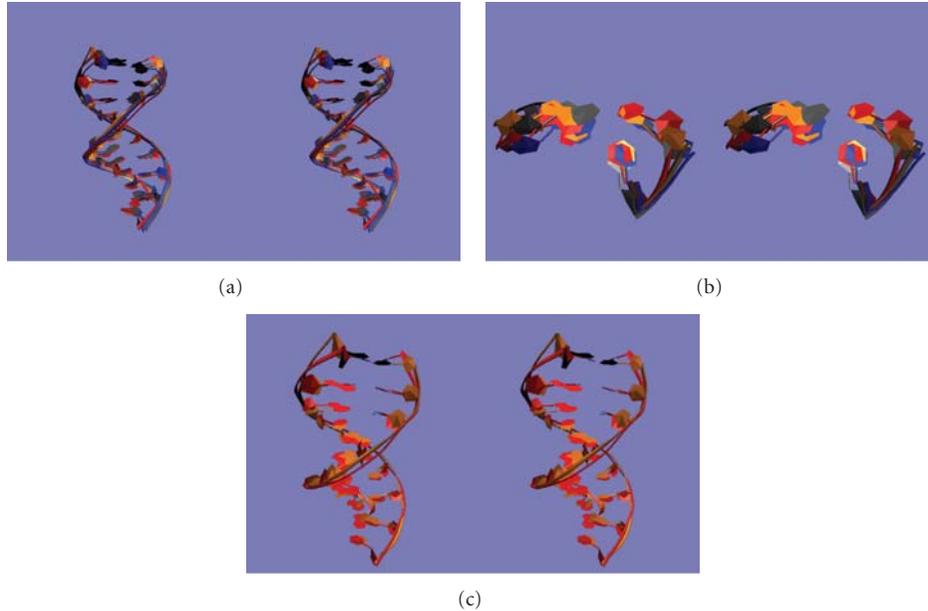


FIGURE 6: (a) Overall conformations adopted in the four clusters. The first ten base pairs of the centroids of each of the four clusters are shown in VMD's NewRibbon format. Cluster 1 (see Table 2 for numbering) is blue, cluster 2 is red, cluster 3 is gray, and cluster 4 is orange. (b) Conformations near the mismatched site in the four clusters. The two base pairs of the centroids of each of the four clusters downstream from the mismatched site are shown in VMD's New Ribbon format. Cluster 1 (see Table 2 for numbering) is blue, cluster 2 is red, cluster 3 is gray, and cluster 4 is orange. (c) Overall conformations of cluster 2 and cluster 4. The first ten base pairs of the centroids of each of the most common and the rarest clusters for matched DNA are shown in VMD's NewRibbon format. Cluster 2 (see Table 2 for numbering) is red and cluster 4 is orange.

TABLE 3: Clustering of simulations via different measures.

Measure/Sequence	Matched	G-T	G-dU	G-FdU
RMSF	1	1	2	2
Covariance matrix	1	1	2	3
Cluster population	1	2	3	2

Cluster number for each simulation for each measure.

in thymineless conditions and imbalanced deoxynucleotide pools resulting in greater preponderance of mismatched base pairs, including G-FdU mismatched base pairs, introduced during replication and subsequent attempted repair processes. The introduction of such mismatched base pairs has been demonstrated to decrease the stability of duplex DNA [23, 24]. The present work demonstrates that mismatched base pairs have a novel effect on the DNA component of a Top1 cleavage complex by destabilizing the putative religation intermediate that consists of a DNA duplex with a nicked phosphodiester backbone.

A G-dU (or G-FdU) base pair in wobble geometry has two hydrogen bonds as does an A-T Watson-Crick base pair and thus is a relatively conservative substitution although at elevated pH an ionized G-FdU base pair may form [24]. G-FdU and other mismatched base pairs not repaired prior to initiation of DNA replication interfere with the religation step of Topoisomerase 1 activity when the mismatched base pair is proximal to the site of Top1 cleavage [8]. The

structural and thermodynamic basis for the decreased Top1-mediated DNA religation due to mismatched base pairs remains incompletely understood. While direct interactions between DNA and Top1 protein may play a major role, the inherent stability of the DNA duplex likely also is a contributing factor. The present results demonstrate that G-FdU and other mismatched base pairs destabilize the interaction of the scissile strand with the non-scissile strand of a model Top1 cleavage complex. The extent of this destabilization is greater under high salt conditions similar to that which occurs in eukaryotic cells.

Molecular dynamics simulations demonstrate that the mismatched base pairs increase the flexibility of the duplex. The extent of this increase in flexibility is dependent upon the type of mismatch with G-dU and G-FdU mismatches displaying the greatest increased flexibility. These calculations reveal that the mismatched base pair causes increased atomic fluctuations up to 10 base pairs removed from the site of the mismatch. This increased flexibility makes it less likely that the scissile strand will adopt the correct conformation required for the religation reaction. Thus, the thermodynamic measurements obtained from the UV hyperchromicity data demonstrate that formation of the complex required for religation is disfavored by all DNA mismatched base pairs at the +1 site of the religation complex while the molecular dynamics simulations reveal that the G-dU and G-FdU mismatched pairs are especially potent at increasing conformational flexibility and decreasing the likelihood religation will occur. Overall, our results provide new

insights into the structural and dynamic process of Top1-mediated DNA religation and the influence of mismatched base pairs, particularly G-dU and G-FdU mismatched base pairs, at disfavoring this process.

Acknowledgment

This work was supported by Grants nos. NIH CA-102532 (W. H. Gmeiner) and NIH P30 CA-12197.

References

- [1] Y. Pommier, E. Leo, H. Zhang, and C. Marchand, "DNA topoisomerases and their poisoning by anticancer and antibacterial drugs," *Chemistry and Biology*, vol. 17, no. 5, pp. 421–433, 2010.
- [2] Y. Pommier, "DNA topoisomerase I Inhibitors: chemistry, biology, and interfacial inhibition," *Chemical Reviews*, vol. 109, no. 7, pp. 2894–2902, 2009.
- [3] A. Y. Chen and L. F. Liu, "DNA topoisomerases: essential enzymes and lethal targets," *Annual Review of Pharmacology and Toxicology*, vol. 34, pp. 191–218, 1994.
- [4] D. A. Koster, V. Croquette, C. Dekker, S. Shuman, and N. H. Dekker, "Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB," *Nature*, vol. 434, no. 7033, pp. 671–674, 2005.
- [5] L. Stewart, M. R. Redinbo, X. Qiu, W. G. J. Hol, and J. J. Champoux, "A model for the mechanism of human topoisomerase I," *Science*, vol. 279, no. 5356, pp. 1534–1541, 1998.
- [6] S. Krogh, U. H. Mortensen, O. Westergaard, and B. J. Bonven, "Eukaryotic topoisomerase I-DNA interaction is stabilized by helix curvature," *Nucleic Acids Research*, vol. 19, no. 6, pp. 1235–1241, 1991.
- [7] P. Pourquier and Y. Pommier, "Topoisomerase I-mediated DNA damage," *Advances in Cancer Research*, vol. 80, pp. 188–216, 2001.
- [8] Z. Y. Liao, O. Sordet, H. L. Zhang et al., "A novel polypyrimidine antitumor agent FdUMP[10] induces thymineless death with topoisomerase I-DNA complexes," *Cancer Research*, vol. 65, no. 11, pp. 4844–4851, 2005.
- [9] P. Pourquier, C. Gioffre, G. Kohlhagen et al., "Gemcitabine (2',2'-difluoro-2'-deoxycytidine), an antimetabolite that poisons topoisomerase I," *Clinical Cancer Research*, vol. 8, no. 8, pp. 2499–2504, 2002.
- [10] W. H. Gmeiner, S. Yu, R. T. Pon, P. Pourquier, and Y. Pommier, "Structural basis for topoisomerase I inhibition by nucleoside analogs," *Nucleosides, Nucleotides and Nucleic Acids*, vol. 22, no. 5-8, pp. 653–658, 2003.
- [11] I. Hirao, G. Kawai, S. Yoshizawa et al., "Most compact hairpin-turn structure exerted by a short DNA fragment, d(GCGMGC) in solution: an extraordinarily stable structure resistant to nucleases and heat," *Nucleic Acids Research*, vol. 22, no. 4, pp. 576–582, 1994.
- [12] L. A. Marky and K. J. Breslauer, "Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves," *Biopolymers*, vol. 26, no. 9, pp. 1601–1620, 1987.
- [13] J. C. Phillips, R. Braun, W. Wang et al., "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.
- [14] A. D. MacKerell, D. Bashford, M. Bellott et al., "All-atom empirical potential for molecular modeling and dynamics studies of proteins," *Journal of Physical Chemistry B*, vol. 102, no. 18, pp. 3586–3616, 1998.
- [15] A. D. MacKerell, N. Banavali, and N. Foloppe, "Development and current status of the CHARMM force field for nucleic acids," *Biopolymers*, vol. 56, no. 4, pp. 257–265, 2000.
- [16] B. R. Brooks, R. D. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "CHARMM: a program for macromolecular energy, minimization and dynamics calculations," *Journal of Computational Chemistry*, vol. 4, pp. 187–217, 1983.
- [17] J. Farwer, M. J. Packer, and C. A. Hunter, "PREDICTOR: a web-based tool for the prediction of atomic structure from sequence for double helical DNA with up to 150 base pairs," *In Silico Biology*, vol. 7, no. 6, pp. 595–600, 2007.
- [18] N. B. Ulyanov, W. R. Bauer, and T. L. James, "High-resolution NMR structure of an AT-rich DNA sequence," *Journal of Biomolecular NMR*, vol. 22, no. 3, pp. 265–280, 2002.
- [19] S. Ghosh, F. R. Salsbury Jr., D. A. Horita, and W. H. Gmeiner, "Zn²⁺ selectively stabilizes FdU-substituted DNA through a unique major groove binding motif," *Nucleic Acids Research*, vol. 39, no. 10, pp. 4490–4498, 2011.
- [20] W. Humphrey, A. Dalke, and K. Schulten, "VMD: visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33–38, 1996.
- [21] Y. Zhou, M. Cook, and M. Karplus, "Protein motions at zero-total angular momentum: the importance of long-range correlations," *Biophysical Journal*, vol. 79, no. 6, pp. 2902–2908, 2000.
- [22] F. R. Salsbury, M. W. Crowder, S. F. Kingsmore, and J. J. A. Huntley, "Molecular dynamic simulations of the metallo-beta-lactamase from *Bacteroides fragilis* in the presence and absence of a tight-binding inhibitor," *Journal of Molecular Modeling*, vol. 15, no. 2, pp. 133–145, 2009.
- [23] J. S. Ullman and B. J. McCarthy, "The relationship between mismatched base pairs and the thermal stability of DNA duplexes. II. Effects of deamination of cytosine," *Biochimica et Biophysica Acta*, vol. 294, no. 3, pp. 416–424, 1973.
- [24] L. C. Sowers, R. Eritja, B. Kaplan, M. F. Goodman, and G. V. Fazakerly, "Equilibrium between a wobble and ionized base pair formed between fluorouracil and guanine in DNA as studied by proton and fluorine NMR," *Journal of Biological Chemistry*, vol. 263, no. 29, pp. 14794–14801, 1988.

Research Article

Triplet Analysis That Identifies Unpaired Regions of Functional RNAs

Junji Kawakami,^{1,2} Yoshie Yamaguchi,^{1,2} and Naoki Sugimoto^{1,2}

¹Department of Nanobiochemistry, FIRST, Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

²Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

Correspondence should be addressed to Junji Kawakami, kawakami@konan-u.ac.jp

Received 15 May 2011; Accepted 28 June 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Junji Kawakami et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We developed a novel method for analyzing RNA sequences, deemed triplet analysis, and applied the method in an *in vitro* RNA selection experiment in which HIV-1 Tat was the target. Aptamers are nucleic acids that bind a desired target (bait), and to date, many aptamers have been identified by *in vitro* selection from enough concentrated libraries in which many RNAs had an obvious consensus primary sequence after sufficient cycles of the selection. Therefore, the higher-order structural features of the aptamers that are indispensable for interaction with the bait must be determined by additional investigation of the aptamers. In contrast, our triplet analysis enabled us to extract important information on functional primary and secondary structure from minimally concentrated RNA libraries. As a result, by using our method, an important unpaired region that is similar to the bulge of TAR was readily predicted from a partially concentrated library in which no consensus sequence was revealed by a conventional sequence analysis. Moreover, our analysis method may be used to assess a variety of structural motifs with desired function.

1. Introduction

In vitro selection of nucleic acids is a powerful method used to isolate novel functional molecules [1–6]. Generally, *in vitro* selection involves many rounds of selection followed by sequencing and analysis of several recovered clones and determination of a consensus sequence. These consensus sequences are often highly functional molecules that have promise as pharmaceutical and chemical agents [7–9]. However, there are several inherent biases that often affect *in vitro* selection procedures, for example, the abundance of particular motifs in the library, differential efficiency of reversetranscription or amplification by PCR due to stable secondary structures. Therefore, confirmation that a particular consensus sequence is optimal is difficult and usually requires identification of a naturally occurring, functional counterpart. For the selection of aptamers, which interact with desired targets, a consensus motif may be a sequence in a local minimum rather than an absolute minimum in an energy diagram that represents interactions between all

possible aptamers and the bait or target. Furthermore, it should be hard to identify general rules of intermolecular interactions from only such “highly evolved” motifs. To overcome this problem, the sequence changes in the library during the early stage of the selection should be analyzed. Evaluation of the sequences that are gradually concentrated from the starting library should reveal a general rule for interactions, that is, information that is indispensable for designing new artificial functional molecules.

Here, we developed a method, designated triplet analysis, to identify important sequence characteristics in the primary sequences that have been recovered from an RNA library after only several rounds of selection with a target molecule. The first step of triplet analysis is extraction of all nucleotide triplets from the primary sequence (Figure 1). For example, a sequence AGCGUCCA would be separated to six triplets: AGC, GCG, CGU, GUC, UCC, and CCA. The next step is reconstruction of the parent sequence of these triplets. If these six triplets could be extracted, one would easily reconstruct the parent sequence, AGCGUCCA. Thus, a

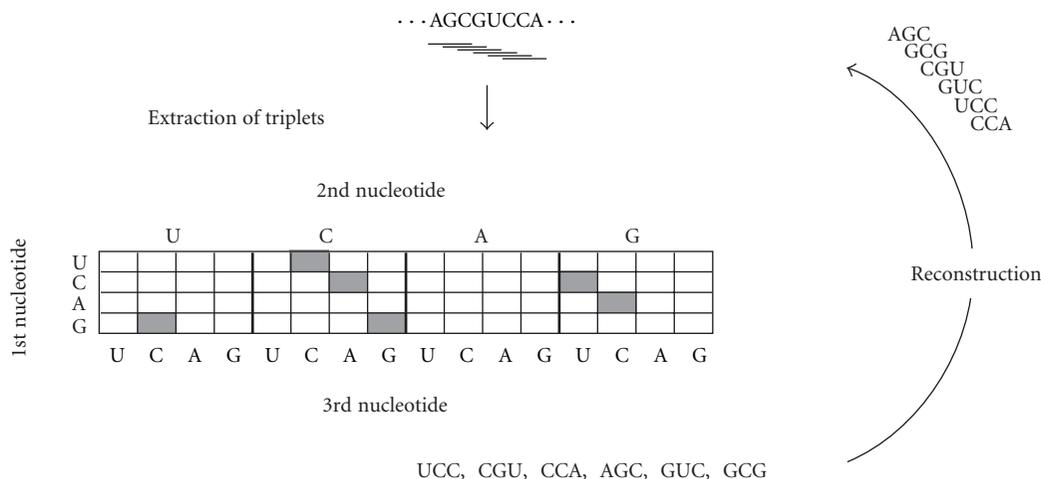


FIGURE 1: The scheme for the basic triplet analysis of a library. Frequent successive sequences could be determined by this method.

frequent successive sequence should be revealed from an analysis of frequent triplets. We chose the HIV-1 Tat protein as our experimental bait for our *in vitro* RNA selection because a naturally occurring binding RNA motif, TAR, and artificially selected binding motifs have been identified [10, 11]. Therefore, we could compare our finding with these previously identified binding motifs to judge the validity of our analysis method.

2. Materials and Methods

2.1. In Vitro Selection of RNA Aptamers. A library comprised of 78-nucleotide RNAs that each contained a region of 30 randomized nucleotides was prepared using chemically synthesized 97-bp DNAs that each had a T7 promoter sequence as templates and *in vitro* transcription. Approximately 1 nmol of RNA was generated from every 100 pmol of template DNA using the AmpliScribe T7 transcription kit (EPICENTRE). The transcripts were purified from 8% denaturing polyacrylamide gels that contained 7 M urea by means of crush and soak extraction and subsequent ethanol precipitation. The bait, Tat peptide (RKKRRGRRR), was synthesized using an Fmoc strategy on a solid support of Fmoc-NH-SAL resin (N- α -9-fluorenylmethoxy-carbonyl-super acid labile polystyrene resin, Watanabe Chem.), and this synthesis method produces an amide at the C-terminus of the bait peptide, as described previously.

The purified RNA was incubated in a binding buffer containing 2.5 mM Tris-HCl (pH 7.6), 100 mM NaCl, 2.0 mM MgCl₂ and then passed through a nitrocellulose filter (HAWP, MILLIPORE) three times to remove RNAs with filter binding properties. A sample of the RNA library (100 pmol) was incubated with 5 pmol of Tat peptide in 50 μ L of the binding buffer for 60 min at 25 C, and then the mixture was passed through another nitrocellulose filter. After washing with 1.5 mL of the binding buffer, RNAs bound to Tat on the filter were collected and amplified by subsequent reverse transcription, PCR, and transcription (RNA PCR kit, PE Applied Biosystems).

2.2. Dot Blot Analysis.

2.3. Cloning and DNA Sequencing. To analyze the binding ability of the RNA library, RNAs were labeled internally with [α -³²P] UTP (Amersham) during transcription *in vitro*. Labeled RNA (10 pmol) was mixed with 2.5 pmol of Tat peptide in 50 μ L of the binding buffer. After a 60 min incubation, the mixture was passed through a nitrocellulose filter. The filter was washed with the binding buffer, and radioactive RNAs on the filter were visualized and quantified with the BIO-Imaging Analyzer, BAS 2000 (Fuji Film).

After six rounds of selection and amplification, the ends of the double-stranded PCR products were blunted by T4 DNA polymerase (TAKARA). The blunted DNAs were purified and concentrated by ethanol precipitation and phosphorylated at their 5' end with T4 polynucleotide kinase (TOYOBO). The resulting DNAs were ligated into the *Hinc* II site of separate pUC118 plasmids using T4 DNA ligase (TAKARA). The ligated plasmid DNAs were transformed into *E. coli* MV1184, which were then cultured on LB plates containing ampicillin, IPTG (isopropyl- β -D-thiogalactoside), and X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside). After blue-white selection, 20 white colonies were randomly selected and cultured in LB medium containing ampicillin. The recombinant plasmid DNA was recovered from *E. coli* clones using the alkaline lysis procedure. Nucleotide sequences were determined using the BigDye-termination method and an ABI PRISM 310 Genetic Analyzer (PE). Of the 20 clones analyzed, 17 contained a library sequence as the insert.

3. Results and Discussion

3.1. Simple Triplet Analysis to Identify Important Primary Structure. From the library shown in Figure 2(a), RNAs that bound Tat were concentrated as described in the previous section. The binding ability of the library RNAs on each round was analyzed by dot blot analysis, and after six rounds of selection, accumulation of Tat-binding RNAs on the

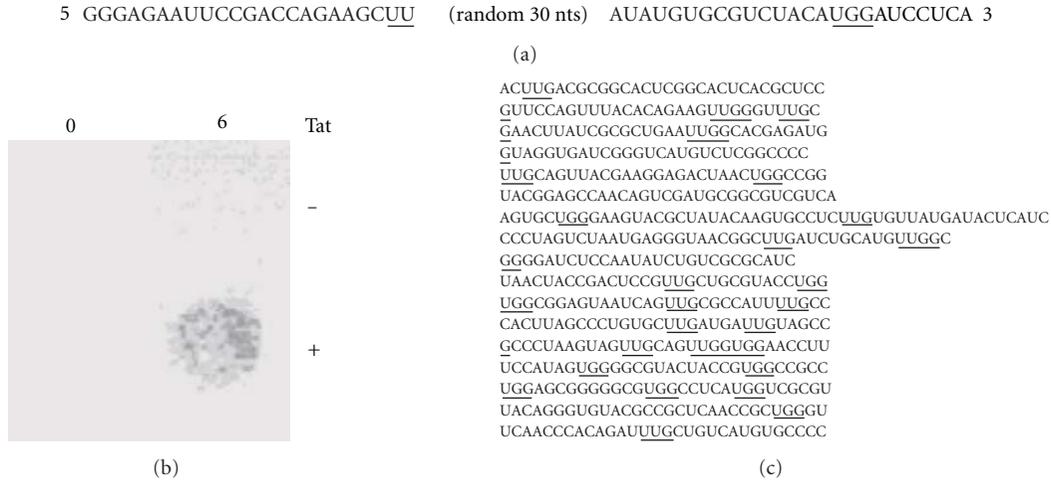


FIGURE 2: The RNA library used in this study. (a) The primary sequence of the constant region and the location of the 30 random nucleotides within the 78-nucleotide RNAs that compose the library are indicated. (b) Tat-binding ability of the libraries before selection (0) and after six rounds of the selection (6). Dot blot analysis was carried out with 10 pmol 32 P-labeled RNA and 2.5 pmol Tat peptide in 50 μ L of the binding buffer. (c) Primary sequence of the 30-nucleotide random region from 17 clones of the library after six rounds of the selection. The frequent triplets after the triplet subtraction, UUG and UGG, are underlined.

TABLE 1: Number and frequency of each triplet obtained from the primary sequences of the clones after the selection^(a).

	Second		U				C			A			G					
	First	Third	U	C	A	G	U	C	A	G	U	C	A	G	U	C	A	G
Number of triplet	U		5	1	5	16	6	5	9	8	4	12	6	6	9	14	8	15
	C		7	10	7	10	7	8	6	7	8	7	5	8	8	13	4	10
	A		4	8	4	9	11	6	5	8	4	8	0	5	14	4	4	4
	G		11	9	9	11	10	14	7	12	9	3	6	6	8	13	7	13
Frequency of triplet	U		0.29	0.06	0.29	0.94	0.35	0.29	0.53	0.47	0.24	0.71	0.35	0.35	0.53	0.82	0.47	0.88
	C		0.41	0.59	0.41	0.59	0.41	0.47	0.35	0.41	0.47	0.41	0.29	0.47	0.47	0.76	0.24	0.59
	A		0.24	0.47	0.24	0.53	0.65	0.35	0.29	0.47	0.24	0.47	0.00	0.29	0.82	0.24	0.24	0.24
	G		0.65	0.53	0.53	0.65	0.59	0.82	0.41	0.71	0.53	0.18	0.35	0.35	0.47	0.76	0.41	0.76

^(a) Abundant triplets with frequency above 0.75 are indicated by boldface.

filter increased as shown in Figure 2(b). The RNAs which recovered after six rounds of selection were cloned using the *E. coli* MV1184/pUC118 system and sequenced. The primary sequences of 17 clones are shown in Figure 2(c).

Apparently, there were no regions of obvious primary sequence similarity shared among these sequences, and a conventional sequence alignment method [12] did not reveal any consensus sequence (data not shown). All triplets in the primary sequences of the 30-nucleotide randomized regions were extracted. Of all 64 possible triplets, 23 triplets had frequencies (the number of triplets divided by the number of clones) above 0.5 (Table 1). In other words, these 23 triplets occurred in a half of the clones stochastically. Notably, eight triplets, UUG, UGC, UGG, CGC, AGU, GCC, GGC, and GGG had, frequencies over 0.75 (bold values in Table 1). Successive sequences of over five nucleotides available from these triplets were UUGCC, UUGGC, and UUGGG. However, these pentanucleotide sequences were found in only one or two of the 17 clones though the eight triplets should potentially be an important component of

a Tat-binding RNA. This finding indicated that the longer consensus sequence had not been concentrated within six rounds of selection and that the process was in a relatively early stage of selection.

3.2. Triplet Subtraction Analysis to Identify the Important Secondary Structure. Further analysis, called triplet subtraction, was carried out as shown in Figure 3 to obtain detailed information, including RNA secondary structural features. In this analysis, surpluses of triplets are defined as the number of a triplet after taking away the number of all complementary triplets. For example, the numbers of AGC, GCG, and CGU subtracted by the numbers of GCU, CGC, and ACG, respectively, were calculated and listed. If a consensus sequence is in a fully matched stem region, there is no remaining triplet after the subtraction. On the other hand, if a consensus region is unpaired (i.e., in a bulge or loop), only the triplets concerned with the unpaired region are revealed after the subtraction. For example, six triplets—GCG, CGU, GUC, UCC, GGG, and

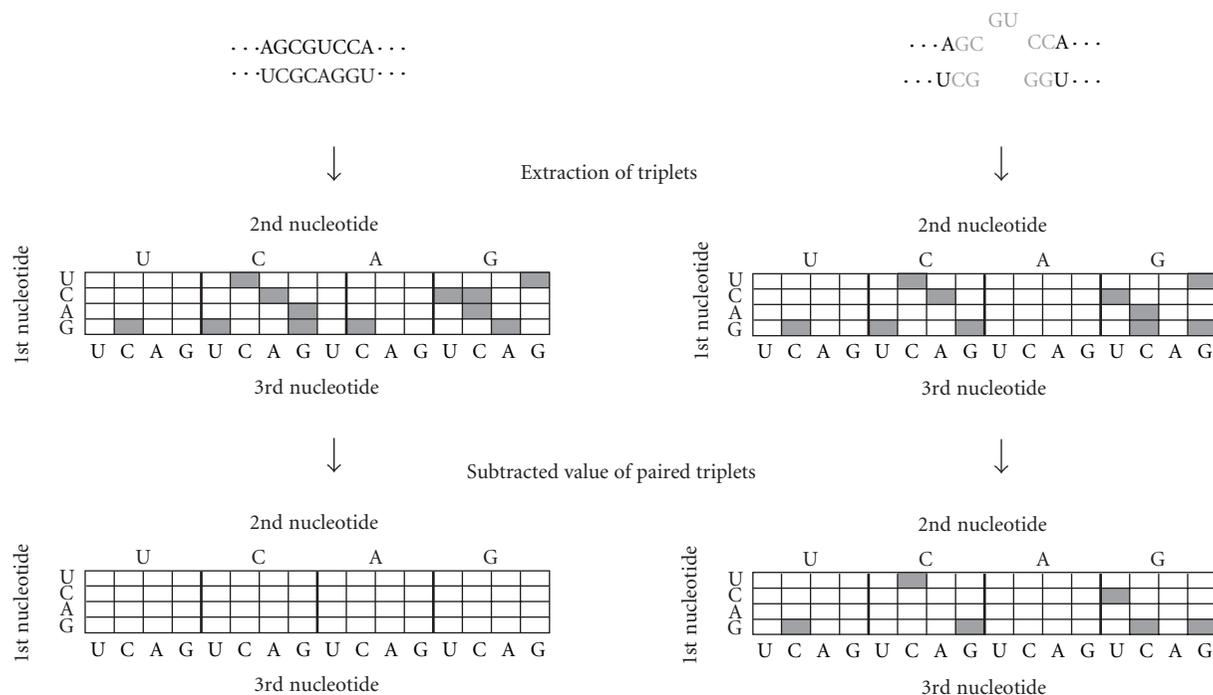


FIGURE 3: The scheme for the triplet subtraction from a library. Only the triplets from unpaired regions could be determined.

GGC—should be revealed from a paired region with a GU bulge, AGCGUCCA/UGGGCU, as shown in Figure 3. Thus, one can easily predict a frequent sequence around an unpaired region that might contribute to the function of an RNA. The results of the triplet subtraction analysis are summarized in Table 2 and Figure 4. As listed in Table 2, only two triplets UUG and UGG had a frequency over 0.5 after the subtraction. This value of 0.5 was surprisingly high because all the triplet frequency should be zero for a truly random library. Though UUU and GGG were the most frequent type of all NUU and GGN (where N indicates U, C, A, or G) triplets, respectively, the frequencies of these triplets were lower than those of UUG or UGG. Therefore, the most probable consensus sequence predicted from these triplets is UUGG. As listed in Table 2, many NNU and GNN might exist in unpaired region and the pattern of these triplets was widely spread, whereas UUG and UGG stood out among the NUN and NGN sequences, respectively. These findings indicated that the second U and/or the third G of the UUGG would be unpaired bases. The frequency of GGN triplets was moderate compared with that of UUG triplets, and, therefore, it was not clear whether the third G was unpaired. Nevertheless, this G might be an unpaired base because CCN triplets seemed not to be concerned with an unpaired region. Additionally, neither AAN nor ANN triplets seemed to occur in unpaired regions of secondary structure. Similar features were not found for NNC. These findings strongly indicated that an unpaired region closed by a UA base pair at their 5' end would not occur in motifs that bound to Tat, that is, the first UU of the UUGG should be unpaired bases. Ultimately, we could determine

that a sequence, NUUGG/CN with unpaired UUG, was an important structural motif that should have significant role in an RNA-Tat interaction (Figure 5(a), left). Possibly, a small portion of Tat-binding RNAs may contain an NUUGG/CCN sequence in which the UU doublet is unpaired as shown in Figure 5(a), right. This conclusion did not conflict with the results from a prediction of the secondary structure of RNA based on our thermodynamic parameters (data not shown).

3.3. Validation of the Triplet Analysis. TAR, the naturally occurring Tat-binding RNA motif, has a pyrimidine bulge, UCU, UUU, or UU, closed by a GC base pair at its 3' end as shown in Figure 5(b) [14–17]. The arginine side residues in Tat stack with the first U of the bulge and make hydrogen bonds to the GC base pair next to the 3' end of the bulge. These interactions along with electrostatic interactions at phosphate backbone are indispensable for the specific Tat-TAR interaction. Therefore, the essential structural motif of TAR that mediates binding with Tat is NNU(Y)UG(A)/(U)CNN with U(Y)U bulge, where Y indicates a pyrimidine nucleotide [11, 18]. Similar structural features were seen in RNA aptamers selected with arginine as a target. A general motif of these arginine aptamers was GUNGA/UCC with a UN bulge as shown in Figure 5(c) [13, 19]. Our UUGG motif closely resembled these indispensable structural features of TAR and arginine aptamers; therefore, we were able to confirm the validity of our triplet analysis method. The UUGG motif was revealed from the early stage of Tat-binding RNA selection, and therefore, the UUGG motif might be the ancestor of other Tat-binding RNAs and have fundamental property as a binding motif.

TABLE 2: Frequency of each triplet after subtraction^(a).

Second		U				C				A				G			
First	Third	U	C	A	G	U	C	A	G	U	C	A	G	U	C	A	G
U		0.29	—	—	0.65	0.12	—	0.06	0.24	—	0.18	0.06	—	0.24	0.41	—	0.53
C		0.12	0.24	0.06	0.12	0.18	—	—	—	—	—	—	—	—	0.06	—	0.18
A		—	—	—	0.06	—	—	—	—	—	—	—	—	0.18	—	—	—
G		0.18	0.35	—	0.24	0.35	0.06	—	—	0.06	—	0.29	—	0.12	—	0.12	0.29

^(a)Dash (—) indicates the value less than or equal to zero.

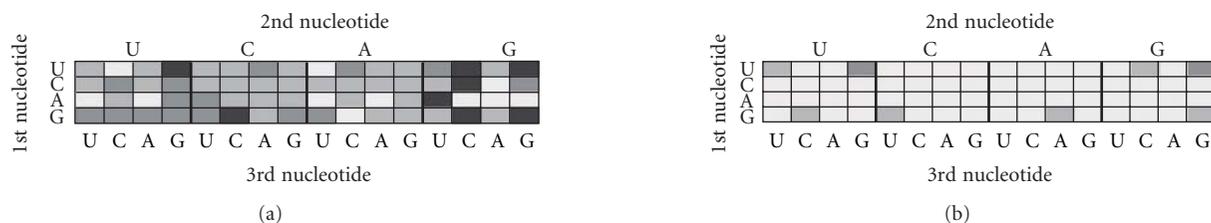


FIGURE 4: The frequency of the triplets before (a) and after (b) subtraction. The changes in gray scale (from light to dark) indicate changes in frequency (from 0–25%, 26–50%, 51–75%, to 76–100%).

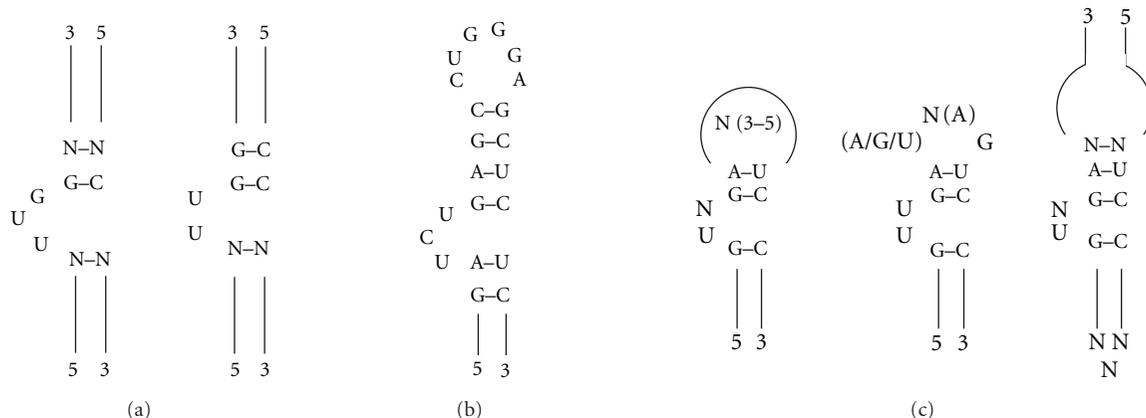


FIGURE 5: Secondary structures of Tat and/or arginine binding RNAs. (a) UUGG motif obtained in the present study. (b) Tat-binding motif of TAR [12]. (c) Several motifs of arginine aptamers from *in vitro* selection experiments [13].

In summary, we developed triplet analysis method to analyze a library. From our analysis of a small number of clones which recovered from a library after only a few rounds of selection, we could determine an important sequence motif though no consensus sequence was revealed by conventional analysis. Furthermore, unpaired regions in the selected RNAs, that may play important roles in many functional nucleic acids, could be readily predicted using this method. The resulting UUGG motif, NUUGG/CN with an unpaired UUG, resembled the essential features of Tat-binding RNAs. This motif may be a fundamental sequence of Tat- and arginine-binding motifs, and researchers may be able to construct a Tat-binding RNA library based on this information for obtaining “higher evolved” motifs.

Acknowledgments

This work was supported in part by Grant-in-Aid for Scientific Research, the “Core Research” Project (2009–2014), and the “Academic Frontier” Project (2004–2009)

from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

- [1] A. D. Ellington and J. W. Szostak, “In vitro selection of RNA molecules that bind specific ligands,” *Nature*, vol. 346, no. 6287, pp. 818–822, 1990.
- [2] C. Tuerk and L. Gold, “Systemic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase,” *Science*, vol. 249, no. 4968, pp. 505–510, 1990.
- [3] D. L. Robertson and G. F. Joyce, “Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA,” *Nature*, vol. 344, no. 6265, pp. 467–468, 1990.
- [4] J. Tang and R. R. Breaker, “Rational design of allosteric ribozymes,” *Chemistry and Biology*, vol. 4, no. 6, pp. 453–459, 1997.
- [5] B. J. Tucker and R. R. Breaker, “Riboswitches as versatile gene control elements,” *Current Opinion in Structural Biology*, vol. 15, no. 3, pp. 342–348, 2005.

- [6] D. S. Wilson and J. W. Szostak, "In vitro selection of functional nucleic acids," *Annual Review of Biochemistry*, vol. 68, pp. 611–647, 1999.
- [7] J. Liu, Z. Cao, and Y. Lu, "Functional nucleic acid sensors," *Chemical Reviews*, vol. 109, no. 5, pp. 1948–1998, 2009.
- [8] S. M. Nimjee, C. P. Rusconi, and B. A. Sullenger, "Aptamers: an emerging class of therapeutics," *Annual Review of Medicine*, vol. 56, pp. 555–583, 2005.
- [9] D. López-Colón, E. Jiménez, M. You, B. Gulbakan, and W. Tan, "Aptamers: turning the spotlight on cells," *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, vol. 3, no. 3, pp. 328–340, 2011.
- [10] R. Yamamoto, M. Katahira, S. Nishikawa, T. Baba, K. Taira, and P. K. R. Kumar, "A novel RNA motif that binds efficiently and specifically to the Tat protein of HIV and inhibits the trans-activation by Tat of transcription in vitro and in vivo," *Genes to Cells*, vol. 5, no. 5, pp. 371–388, 2000.
- [11] A. Marozzi, R. Meneveri, M. Giacca, M. I. Gutierrez, A. G. Siccardi, and E. Ginelli, "In vitro selection of HIV-1 TAR variants by the Tat protein," *Journal of Biotechnology*, vol. 61, no. 2, pp. 117–128, 1998.
- [12] J. P. Davis, N. Janjić, B. E. Javornik, and D. A. Zichi, "Identifying consensus patterns and secondary structure in SELEX sequence sets," *Methods in Enzymology*, vol. 267, pp. 302–312, 1996.
- [13] J. Tao and A. D. Frankel, "Arginine-binding RNAs resembling TAR identified by in vitro selection," *Biochemistry*, vol. 35, no. 7, pp. 2229–2238, 1996.
- [14] M. Guyader, M. Emerman, and P. Sonigo, "Genome organization and transactivation of the human immunodeficiency virus type 2," *Nature*, vol. 326, no. 6114, pp. 662–669, 1987.
- [15] M. A. Muesing, D. H. Smith, and D. J. Capon, "Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein," *Cell*, vol. 48, no. 4, pp. 691–701, 1987.
- [16] B. Berkhout, R. H. Silverman, and K. T. Jeang, "Tat trans-activates the human immunodeficiency virus through a nascent RNA target," *Cell*, vol. 59, no. 2, pp. 273–282, 1989.
- [17] C. Dingwall, I. Ernberg, M. J. Gait et al., "HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure," *EMBO Journal*, vol. 9, no. 12, pp. 4145–4153, 1990.
- [18] J. Tao and A. D. Frankel, "Specific binding of arginine to TAR RNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 7, pp. 2723–2726, 1992.
- [19] Y. Yang, M. Kochoyan, P. Burgstaller, E. Westhof, and M. Famulok, "Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy," *Science*, vol. 272, no. 5266, pp. 1343–1347, 1996.

Research Article

Synthesis of Specifically Modified Oligonucleotides for Application in Structural and Functional Analysis of RNA

Nico Rublack,¹ Hien Nguyen,¹ Bettina Appel,¹ Danilo Springstube,¹
Denise Strohbach,^{1,2} and Sabine Müller¹

¹Ernst-Moritz-Arndt-Universität Greifswald, Institut für Biochemie, Felix-Hausdorff-Straße 4, 17487 Greifswald, Germany

²Biomers.net GmbH, Söflinger Straß 100, 89077 Ulm, Germany

Correspondence should be addressed to Sabine Müller, smueller@uni-greifswald.de

Received 14 May 2011; Revised 11 July 2011; Accepted 21 July 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Nico Rublack et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, RNA synthesis has become an essential tool not only in the field of molecular biology and medicine, but also in areas like molecular diagnostics and material sciences. Beyond synthetic RNAs for antisense, aptamer, ribozyme, and siRNA technologies, oligoribonucleotides carrying site-specific modifications for structure and function studies are needed. This often requires labeling of the RNA with a suitable spectroscopic reporter group. Herein, we describe the synthesis of functionalized monomer building blocks that upon incorporation in RNA allow for selective reaction with a specific reporter or functional entity. In particular, we report on the synthesis of 5'-O-dimethoxytrityl-2'-O-*tert*-butyldimethylsilyl protected 3'-O-phosphoramidites of nucleosides that carry amino linkers of different lengths and flexibility at the heterocyclic base, their incorporation in a variety of RNAs, and postsynthetic conjugation with fluorescent dyes and nitroxide spin labels. Further, we show the synthesis of a flavine mononucleotide-N-hydroxy-succinimidyl ester and its conjugation to amino functionalized RNA.

1. Introduction

Over the past two decades, RNA synthesis has become a very active field. Synthetic RNAs are required for a large number of applications. Antisense oligonucleotides, ribozymes, aptamers, and siRNAs are required for medicinal diagnostic and therapy as well as for a wide variety of biochemistry and molecular biology studies [1, 2]. Furthermore, aptamers and reporter ribozymes are designed and applied in environmental diagnostics [3]. A number of other disciplines make use of synthetic oligoribonucleotides, for example, the field of material science, where novel materials from nanoparticle oligonucleotide conjugates are generated [4]. In parallel, RNA synthesis has developed to a degree that allows the synthesis of RNA oligonucleotides of any desired sequence from microgram to multigram scale. In addition to 2'-O-TBDMS chemistry, which may be considered as the standard procedure for RNA preparation, novel strategies mainly focussing on different 2'-O-protecting groups such as 2'-O-TOM, 2'-O-ACE or 2'-O-TC are available for laboratory use applying the specific monomer building blocks, or have been commercialized for custom RNA synthesis [5, 6].

Research in our laboratory is devoted to the chemistry and biochemistry of RNA with a strong focus on RNA aptamers and ribozymes. Therefore, we synthesize natural and modified RNA strands required for the design of functional RNAs that we want to investigate. Furthermore, we also prepare RNA molecules carrying site-specific modifications such as fluorescent dyes or other reporter groups to be used for studies into RNA structure. A number of monomer building blocks for the synthesis of modified RNA are commercially available. For example, amino linkers for 5'- and/or 3'-terminal labeling can be obtained ready for use in RNA assembly (Figure 1). Furthermore, 5'-O-DMT-2'-O-TBDMS protected 3'-O-phosphoramidites of 2'-aminouridine (3) and 4-thiouridine (4) (Figure 1) are available and after incorporation in RNA can be used for the specific attachment of desired molecular entities.

However, monomer building blocks of the purine nucleosides with functionalities suitable for postsynthetic conjugation are completely missing, and also in the pyrimidine series, the few existing derivatives of uridine do not offer much variety. Therefore, a currently very active part of our

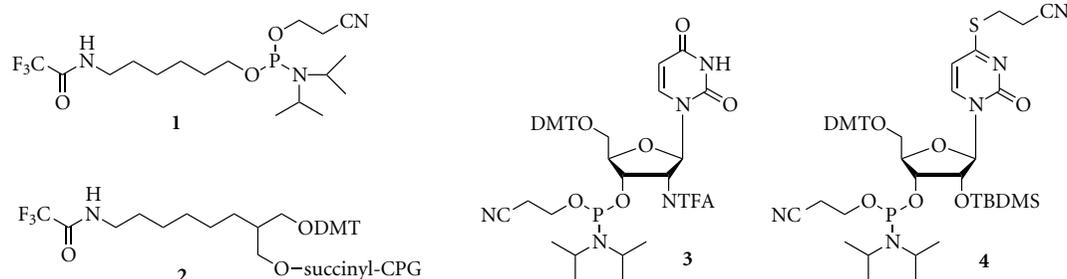


FIGURE 1: Structures of the commercially available building blocks used in our RNA functionalization studies.

research is focussed on the synthesis of 5'-O-DMT-2'-O-TBDMS protected 3'-O-phosphoramidites of nucleosides that carry functionalized linkers of different lengths and flexibility at the heterocyclic base, in particular at C5 of the pyrimidines as well as at C8 and C2 of the purines (Figure 2). Moreover, we currently work at the synthesis of modified nucleosides as they appear in tRNA as well as of isotope enriched nucleosides and their incorporation in specific RNA sequences (which will be reported elsewhere).

Herein, we focus on the synthesis of amino-modified monomers, their incorporation into RNA, and the postsynthetic labeling with fluorescent dyes. Furthermore, we report on strategies for preparation of specifically spin labeled RNAs for EPR studies, and on the preparation of a flavinmononucleotide (FMN) derivative that is postsynthetically attached to an amino-modified hairpin ribozyme derivative. The modified RNAs described herein are used in collaboration with spectroscopists for structural analysis of functional RNAs, such as RNA thermometers [7] and RNA four-way junctions [8] or riboswitches [9], or in RNA engineering projects like the design of an FMN-dependent aptazyme [10].

2. Materials and Methods

2.1. General Information. All reactions were carried out in dry solvents under argon atmosphere. All solvents and reagents were purchased from commercial sources and used as supplied unless otherwise stated. The solvents used in Pd coupling reactions were freed from oxygen. All products were visualized on TLC plates (aluminium sheets coated with silica gel 60 F 254, 0.2 nm thickness) at 254 nm ultraviolet light. The linker molecules L1, L2, and L3 (Figure 3) were synthesized according to the literature: the alkinyl linker L1 was synthesized according to Cruickshank and Stockwell [11], and alkenyl linkers were synthesized as described by Dey and Sheppard [12] and McKeen et al. [13]. All modified phosphoramidites were freshly prepared prior to use in RNA synthesis. Progress of reaction was monitored by TLC (hexane/ethyl acetate 1:1 containing 1% NEt₃). Phosphoramidites were coevaporated several times with dry DCM and kept under vacuum over P₂O₅ in a desiccator overnight. A 0.1 M solution of the modified phosphoramidite in anhydrous acetonitril was freshly prepared right before use in solid-phase RNA synthesis.

2.2. MALDI-TOF Analysis. For nucleoside derivatives: 1–2 mg sample was dissolved in 100 μL methanol or water. 1.5 μL of the sample solution and 1.5–3.0 μL of the matrix solution (30–40 mg 3-hydroxypicolinic acid (3-HPA) in 500 μL MeCN/H₂O 1:1, vigorously mixed for 30 s) were mixed, and 1–2 μL of that mixture was loaded on a MALDI-plate, dried, and allowed to crystallize. The sample was analyzed on a Bruker microflex mass spectrometer. The program was set as reflective positive mode with 40–50 shoots/measurement. The laser intensity was adjusted such that the intensity of the desired signals was about 1000.

For RNA: RNA samples were desalted by gel filtration on Sephadex G25 fine (GE Healthcare). After desalination, samples were dissolved in autoclaved micropore water to a concentration of about 100–500 pmol/μL. 1 μL of this RNA solution was mixed with 1–2 μL of matrix solution (30–40 mg 3-hydroxypicolinic acid (3-HPA) in 500 μL MeCN/H₂O 1:1, vigorously mixed for 30 s) and treated with cation-exchange beads (Dowex 50 WX8, NH₄⁺-form, 100–200 mesh, ServaFeinbiochemica). Therefore, 4–7 μL suspension of the beads in autoclaved micropore water was pipetted in a 250 μL tube, and water was removed, followed by addition of the sample and matrix solutions. The mixture was left at rt for 10 min then loaded on the MALDI-plate and dried on air. The sample was measured using the linear negative (or positive) mode with 40–50 shoots/measurement. In addition to the main peak [M–1][–] (negative mode) or [M+1]⁺ (positive mode), occasionally the signals corresponding to [M–2]/2 were observed.

2.3. Synthesis of Amino-Modified Monomer Building Blocks. Amino linker-modified uridine derivatives were synthesized as described [14]. The detailed synthesis of linker-modified adenosine derivatives with focus on the problem of double bond isomerization will be reported elsewhere (H. Nguyen, B. Appel, and S. Müller, manuscript in preparation).

2.4. 8-(3-Trifluoroacetamidoprop-1-ynyl)guanosine 16. 8-Bromoguanosine-hydrate **15** (98%) (1448.4 mg, 4.0 mmol) was dissolved in dry DMF (32 mL). Traces of dissolved oxygen in the solution were removed by iterative cycles of reduced pressure and passing in argon (3 times). Then, Pd(Ph₃P)₄ (464 mg, 0.4 mmol, 0.1 eq.) was added. The solution was again kept under vacuum and then saturated with argon. DIPEA (816 μL, 4.8 mmol, 1.2 eq., freshly

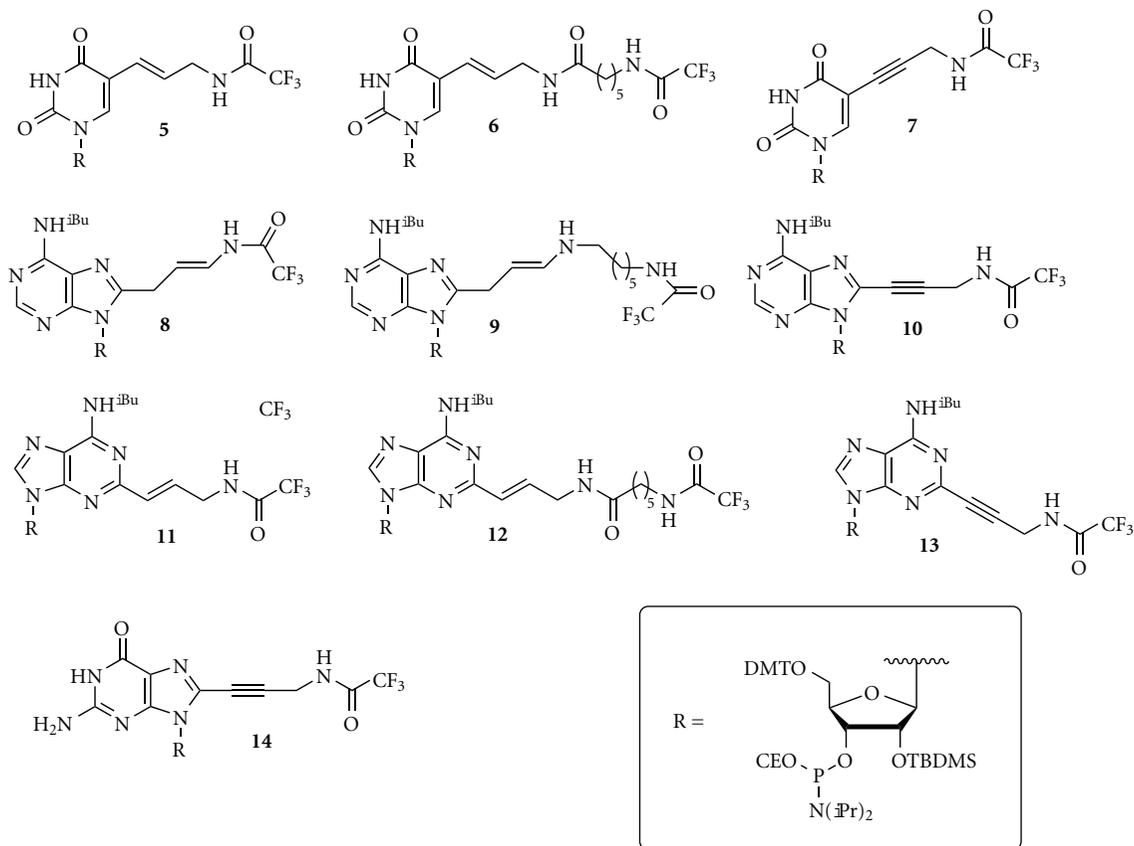


FIGURE 2: Amino-modified monomer building blocks used for RNA synthesis and functionalization.

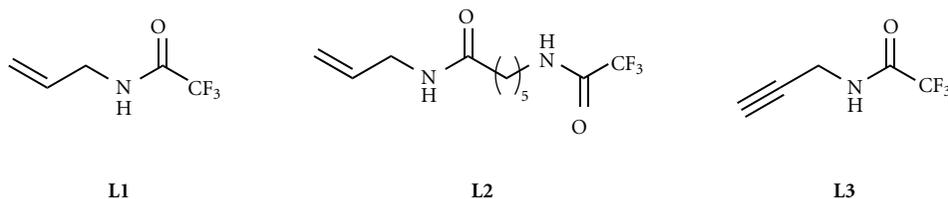


FIGURE 3: Amino linkers used for preparation of amino-modified nucleosides.

distilled over CaH_2) was added via a syringe, followed by the addition of *N*-propynyltrifluoroacetamide L3 (1208 mg, 8.0 mmol, 2.0 eq.) and CuI (152 mg, 0.8 mmol, 0.2 eq.). The mixture was stirred at 40°C for 20 hours. The solvent was then removed to give a brown viscous oil which was dissolved in DCM/MeOH 1:1 (v/v) and filtered to remove the precipitate. The solid residue was purified by silica gel column chromatography (DCM/MeOH 90:10 to 84:16 (v/v)) to give a pale yellow solid. Yield: 2765.4 mg (6.4 mmol, 80%). R_f (SiO_2 ; DCM/MeOH 60:40 (v/v)): 0.80. ^1H NMR (300 MHz, DMSO-d_6): δ (ppm) = 3.50 (1H, m, H-5'), 3.64 (1H, m, H-5''), 3.84 (1H, m, H-4'), 4.12 (1H, m, H-3'), 4.36 (2H, d, J 5.4, CH_2), 4.88 (2H, m, H-2' and OH), 5.04 (1H, d, J 4.8, OH), 5.40 (1H, d, J 6.0, OH), 5.76 (1H, d, J 6.3, H-1'), 6.57 (2H, br s, NH_2), 10.20 (1H, t, J 5.4, NH), 10.90 (1H, br s, NH). ^{13}C NMR (300 MHz, DMSO-d_6): δ (ppm) = 29.40, 62.06, 70.47, 70.85, 72.75, 85.77, 88.38,

89.12, (110.03, 113.84, 117.65 and 121.47, q, J 288.02, CF_3), 117.24, 129.04, 150.90, 154.12, (155.54, 156.03, 156.51 and 157.00, q, J 36.76, $(\text{C}=\text{O})\text{CF}_3$), 156.24. MALDI-TOF: $\text{C}_{15}\text{H}_{15}\text{F}_3\text{N}_6\text{O}_6$, calculated 432.10, found 432.82 $[\text{M}+\text{H}]^+$.

2.5. *N*-2-Isobutryryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosine 17. 8-(3-Trifluoroacetamido-prop-1-ynyl)guanosine **16** (216.5 mg, 0.50 mmol) was coevaporated with anhydrous pyridine ($3 \times 4\text{ mL}$), and then dissolved in anhydrous pyridine (2 mL). The resulting solution was protected from moisture (drying tube), purged with argon and placed on ice. To the ice cold solution, TMS-Cl (0.55 mL, 4.1 mmol, 8.2 eq.) was added dropwise via a syringe. The ice bath was then removed and the mixture was stirred for 2 hours. The solution was cooled on ice and isobutyric anhydride (0.13 mL, 1.1 mmol, 2.2 eq.) was added dropwise via a syringe and the

ice bath was removed. After stirring for another 2 hours at room temperature, the reaction was placed again on ice and ice cold water (20 mL) was slowly added, followed after 15 minutes by concentrated ammonia solution (1.5 mL) to get a final 2.5 M concentration of ammonia. The mixture was kept on ice for 30 minutes, and then evaporated to dryness. The residue was coevaporated with toluene (3 × 5 mL) to remove traces of water, resuspended in MeOH and filtered to remove the precipitate. The filtrate was then concentrated, dissolved in a small amount of MeOH, absorbed on silica gel and purified by column chromatography (DCM/MeOH 95:5 to 91:9 (v/v)) to give *N*-2-isobutyryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosine as a yellow solid. Yield: 183.3 mg (0.37 mmol; 73%). R_f (SiO₂; DCM/MeOH 80:20 (v/v)): 0.63. ¹H NMR (300 MHz, DMSO-*d*₆): δ (ppm) = 1.13 (6H, d, J 6.9, CH(CH₃)₂), 2.79 (1H, sept, J 6.9, CH(CH₃)₂), 3.53 (1H, m, H-5'), 3.68 (1H, m, H-5''), 3.84 (1H, m, H-4'), 4.17 (1H, br s, H-3'), 4.40 (2H, s, CH₂), 4.81 (1H, br s, H-2'), 4.91 (1H, br s, OH), 5.09 (1H, br s, OH), 5.50 (1H, br s, OH), 5.86 (1H, d, J 6.0, H-1'), 10.21 (1H, br s, NH), 11.62 (1H, br s, NH), 12.16 (1H, br s, NH). MALDI-TOF: C₁₉H₂₁F₃N₆O₇, calculated 502.14, found 502.84 [M+H]⁺.

2.6. 5'-*O*-(4,4'-Dimethoxytrityl)-*N*-6-isobutyryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosine **18**. *N*-2-Isobutyryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosine **17** (1174.4 mg, 2.4 mmol) was coevaporated with anhydrous pyridine (3 × 3 mL) and dissolved in anhydrous pyridine (10.0 mL). DMT-Cl (1219.8 mg, 3.6 mmol, 1.5 eq.) was dissolved in anhydrous pyridine (3.0 mL) and added dropwise to the ice cold solution of the nucleoside over 2.5 hours. After 6 hours at ice cold temperature, TLC analysis (DCM/MeOH 95:5(v/v)) showed that the starting material was completely consumed. Methanol (1.0 mL) was added to quench the unreacted DMT-Cl, and after 10 minutes, the solution was concentrated to dryness *in vacuo*. The residue was dissolved in DCM (100 mL), washed with a 5% aqueous solution of NaHCO₃ (2 × 20 mL) and water (2 × 20 mL). The organic layer was dried over Na₂SO₄ and concentrated to give a yellow gum which was purified by column chromatography (DCM/MeOH 97:3 to 95:5 (v/v) containing 1% triethylamine) to give a pale yellow foam. Yield: 1671.6 mg (2.08 mmol, 87%). R_f (DCM/MeOH 90:10 (v/v)): 0.44. ¹H NMR (300 MHz, DMSO-*d*₆): δ (ppm) = 1.11 (6H, 2 × d, J 6.9 and 6.6, CH(CH₃)₂), 2.72 (1H, sept, J 6.9, CH(CH₃)₂), 3.11 and 3.45 (2H, m, H-5' and H-5''), 3.69 and 3.71 (6H, 2 × s, 2 × OCH₃), 4.04 (1H, m, H-4'), 4.31 (3H, br s, H-3' and CH₂), 4.91 (1H, m, H-2'), 5.04 (1H, br s, OH), 5.61 (1H, br s, OH), 5.94 (1H, d, J 4.5, H-1'), 6.73 (4H, 2 × d, J 8.7 and 9.0, aromatic), 7.23 (9H, m, aromatic), 10.85 (3H, br s, 3 × NH). ¹³C NMR (300 MHz, DMSO-*d*₆): δ (ppm) = 18.71, 18.94, 29.33, 34.80, 54.86, 54.91, 64.72, 70.49, 71.36, 72.06, 84.21, 85.21, 89.91, 90.49, (110.02, 113.82, 117.65 and 121.47, q, J 288.5, CF₃), 112.73, 112.82, 120.83, 126.45, 127.47, 127.78, 129.64, 129.80, 131.52, 135.58, 144.87, 148.24, 154.33, (155.55, 156.05, 156.53 and 157.03, q, J 37.0, (C=O)CF₃), 157.88, 157.95, 180.09. MALDI-TOF: C₄₀H₃₉F₃N₆O₉, calculated 804.27, found 805.24 [M+H]⁺.

2.7. 5'-*O*-(4,4'-Dimethoxytrityl)-2'-(*tert*-butyldimethylsilyl)-*N*-2-isobutyryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosine **19**. 5'-*O*-(4,4'-Dimethoxytrityl)-*N*-2-isobutyryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosine **18** (1623 mg, 2.02 mmol) was coevaporated with anhydrous pyridine (2 mL × 2), dry DCM (2 × 5 mL), kept under vacuum overnight, and then dissolved in anhydrous THF (20 mL). Anhydrous pyridine (606 μ L, 7.47 mmol, 3.7 eq.) and AgNO₃ (505 mg, 3.0 mmol, 1.5 eq.) were added. The reaction mixture was stirred for 15 minutes to give a white emulsion, and TBDMS-Cl (528 mg, 3.43 mmol, 1.7 eq.) was added. The reaction was protected from moisture and stirred at room temperature until TLC (SiO₂, hexane/ethyl acetate 1:1 (v/v)) showed no starting material left (8.5 hours). The reaction mixture was diluted with ethyl acetate (50 mL), filtered to remove the precipitate and washed with a saturated solution of NaHCO₃ (2 × 20 mL). The organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to give a white gum which was then purified by silica gel column chromatography (hexane/ethyl acetate 85:15 to 40:60 (v/v)) to give a pale yellow foam. Yield: 482.3 mg (0.53 mmol, 26%). R_f (SiO₂; hexane/ethyl acetate 1:1 (v/v) containing 4% MeOH): 0.41. ¹H NMR (300 MHz, DMSO-*d*₆): δ (ppm) = -0.14 and -0.02 (6H, 2 × s, Si(CH₃)₂), 0.76 (9H, s, SiC(CH₃)₃), 1.09 (6H, 2 × d, J 6.6 and 6.9, CH(CH₃)₂), 2.70 (1H, sept, J 6.6 and 6.9, CH(CH₃)₂), 3.19 (1H, m, H-5'), 3.51 (1H, m, H-5''), 3.67 and 3.69 (6H, 2 × s, 2 × OCH₃), 4.09 (1H, m, H-4'), 4.20 (1H, br m, H-3'), 4.38 (2H, br d, J 5.1, CH₂), 4.79 (1H, m, H-2'), 5.95 (1H, d, J 4.8, H-1'), 6.75 (4H, 2 × d, J 8.7, aromatic), 7.27 (9H, m, aromatic), 10.16 (1H, t, J 5.1 and 5.4, NH), 11.38 (1H, br s, NH), 12.14 (1H, br s, NH). ¹³C NMR (300 MHz, DMSO-*d*₆): δ (ppm) = -5.40, -4.91, 17.81, 18.63, 18.88, 25.47, 29.24, 34.77, 54.89, 54.93, 64.49, 70.16, 71.95, 73.74, 84.26, 85.30, 90.66, (109.97, 113.79, 117.61 and 121.43, q, J 288.40, CF₃), 112.82, 112.87, 126.51, 127.55, 127.76, 129.70, 129.78, 135.50, 135.58, 144.86, 148.52, 154.10, (155.54, 156.03, 156.52 and 157.01, q, J 37.0, (C=O)CF₃), 157.93, 157.98, 180.07. MALDI-TOF: C₄₆H₅₃F₃N₆O₉Si, calculated 918.36, found 919.19 [M+H]⁺.

2.8. 5'-*O*-(4,4'-Dimethoxytrityl)-2'-(*tert*-butyldimethylsilyl)-*N*-2-isobutyryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosinephosphoramidite **14**. 5'-*O*-(4,4'-Dimethoxytrityl)-2'-(*tert*-butyldimethylsilyl)-*N*-2-isobutyryl-8-(3-trifluoroacetamidoprop-1-ynyl)guanosine **19** (67.96 mg, 0.11 mmol) was coevaporated with anhydrous pyridine (2 × 5 mL), dry DCM (2 × 5 mL), kept under vacuum overnight, and then dissolved in anhydrous DCM (2 mL). To the resulting solution, DIPEA (75 μ L, 0.44 mmol, 4 eq.) was added, followed by 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (37 μ L, 0.16 mmol, 1.5 eq.). The solution was stirred at room temperature. The reaction was followed by TLC (EtOAc:hexane 1:1 containing 1% TEA). After 2.5 hours, the reaction was stopped and washed with a saturated solution of NaHCO₃ (10 mL). The organic layer was dried over Na₂SO₄ and filtered, and the solvent was removed under reduced pressure. The residue was purified by silica gel column chromatography. The column was

packed with DCM/TEA (98:2), the residue was dissolved in 1–2 mL DCM/TEA (98:2), and the column was loaded. After 30 mL of DCM/TEA (98:2), the solvent was changed to DCM/TEA/MeCN (49:1:50, 50 mL). The fractions containing the desired product were collected, and the solvent was removed under reduced pressure. The product was lyophilized overnight. ^{31}P -NMR (300 MHz, DMSO- d_6), two diastereomers: δ (ppm) = 148.98, 149.71.

2.9. RNA Synthesis. Oligoribonucleotides were synthesized by the phosphoramidite method on a Pharmacia Gene Assembler Plus, at 1 μmol scale as described elsewhere [15]. Standard PAC-phosphoramidites as well as CPG supports were obtained from ChemGenes or Link Technologies. BMT (emp Biotech) was used as activator [16]. The modified nucleoside phosphoramidites were coevaporated three times with dry dichloromethane, kept under vacuum for 2 hours to remove traces of solvents, stored in vacuum over P_2O_5 overnight, and used in oligo synthesis as 0.1–0.15 M solution in acetonitrile. The solutions of phosphoramidites, BMT, and MeCN were kept over molecular sieve 0.3 nm. The coupling time for all natural and modified building blocks was 5 min. All syntheses were carried out “trityl-off”. The obtained RNA was cleaved from the support and deprotected using NH_3 (30%) and ethanolic methylamine in a 1:1 mixture and TEA3HF as described [17] and purified by gel electrophoresis using 10% denaturing polyacrylamide gels. Elution was carried out using 0.5 M LiOAc followed by EtOH precipitation. Oligonucleotides were analyzed by PAGE and MALDI MS.

2.10. Postsynthetic Labeling of Oligoribonucleotides with Alexa488, Cy5 and ATTO647N. 5 to 10 nmol amino-modified RNA were dissolved in 15–25 μL 0.1 M Borax buffer, at pH 9.4 for Alexa488 coupling reactions, and at pH 8.7 for Cy5 and ATTO647N coupling reactions. The mixture was added to a solution of 100 μg dye in 5 μL DMF (the solution of the dye in DMF was freshly prepared just before use). The coupling reaction was carried out at room temperature overnight in the dark. Excess of dye was removed via gel filtration (Sephadex G25 fine, GE Healthcare). The Alexa488 and Cy5 labeled RNAs were purified by denaturing PAGE. The sample was denatured at 90°C for 3 minutes, immediately subjected onto a 15% polyacrylamide gel (200 \times 150 \times 1.5 mm), and electrophoresis was run at 400 V for 5.5 hours in the dark. The band corresponding to the labeled RNA was cleaved out; the RNA was eluted with 0.5 M lithium acetate and precipitated from ethanol at -20°C (yield 15–37% for Alexa488, 26–54% for Cy5). The ATTO647N labeled RNA was purified by reverse-phase HPLC. The RNA was solved in 1 mL water and filtered through a 0.45 μL filter before subjecting onto the column. Column EC 250/4 Nucleodur 100-5 C18 ec (Macherey-Nagel), flow rate 0.5 mL/min, Buffer A 0.1 M TEAAc (pH 7), 5% MeCN, Buffer B 0.1 M TEAAc (pH 7), 70% MeCN, Gradient 0% B for 5 min, 10–55% B in 76 min, 55–100% B in 19 minutes (yield 40–50%).

2.11. Hybridization of Single-Stranded RNAs for Structural Studies on RNA Thermometer

Condition A. For the hybridization of unlabeled single-stranded oligonucleotides, D1L3 (300 pmol) and A1L3 (300 pmol) were dissolved in 50 mM Tris buffer, pH 7.4 in the presence of 10 mM MgCl_2 in a total volume of 100 μL . The solution was then well mixed, incubated at 78°C for 3 minutes, centrifuged to collect the vapor on the cap of the eppendorf tube, wrapped in aluminium film, and slowly cooled down to room temperature. The hybrid was analyzed by 15% native polyacrylamide gel.

Condition B. Alexa Fluor 488-labeled D1L3 (25 pmol) and Cy5-labeled A1L3 (25 pmol), were dissolved in 20 mM KH_2PO_4 - K_2HPO_4 buffer, pH 6.5 in the presence of 10 mM MgCl_2 and 100 mM KCl in a total volume of 50 μL . The components were then mixed, centrifuged and incubated at 78°C for 3 minutes. The mixture was slowly cooled down to room temperature in aluminium film. The hybridization was checked by 15% native polyacrylamide gel.

2.12. 5'-O-(4,4'-Dimethoxytrityl)riboflavin 21. To a suspension of riboflavin **20** (1.5 g, 4 mmol) in dry pyridine TEA (2 mL, 14.4 mmol), DMAP (24 mg, 0.2 mmol), and 4,4'-dimethoxytritylchloride (1.7 g, 5.2 mmol) were added under argon. After 4 hours of stirring at room temperature in the dark, the reaction was quenched with dry methanol (5 mL). The solvent was evaporated under reduced pressure. The residue was dissolved in DCM (100 mL) and filtered. The clear yellow filtrate was dried with Na_2SO_4 and the solvent was removed. Purification of the crude product by column chromatography with DCM/MeOH (98:2) gave the tritylated product **21** as an orange solid. Yield: 1.4 g (2.1 mmol, 53%). ^1H NMR (300 MHz, DMSO- d_6): δ (ppm) = 2.39 (s, 3 H, CH_3), 2.41 (s, 3 H, CH_3), 3.20–3.17 (m, 2 H), 3.72 and 3.73 (2 \times s, 2 \times 3 H, 2 \times OCH_3), 3.92 (br s, 1 H), 4.23 (br s, 1 H), 4.62 (m 1 H), 4.98 (m, 1 H), 6.85 (m, 4 H, DMT), 7.46–7.16 (m, 9 H, DMT), 7.83 (s, 1 H), 7.90 (s, 1 H), 11.34 (s, 1 H, NH). ESI-MS: $\text{C}_{38}\text{H}_{38}\text{N}_4\text{O}_8$, calculated 678.27, found 677.26 $[\text{M}-\text{H}]^-$.

2.13. 2',3',4'-Tri-O-acetyl-5'-O-(4,4'-dimethoxytrityl)riboflavin 22. 5'-O-(4,4'-Dimethoxytrityl)-riboflavin **21** (1.4 g, 2.06 mmol) in dry pyridine (200 mL) was cooled in an ice bath under argon atmosphere. After dropwise addition of acetic anhydride (7.6 mL, 80 mmol), the solution was stirred in the dark at room temperature overnight. Then, methanol (5 mL) was added, and the solvent was evaporated. Column chromatography with DCM/MeOH (98:2) gave the title compound **22** as orange foam. Yield: 1.5 g (1.86 mmol, 90%). ^1H NMR (300 MHz, DMSO- d_6): δ (ppm) = 1.49 (s, 3 H, acetyl- CH_3), 2.02 (s, 3 H, acetyl- CH_3), 2.27 (s, 3 H, acetyl- CH_3), 2.39 (s, 3 H, CH_3), 2.46 (s, 3 H, CH_3), 3.12–3.05 (m, 1 H), 3.29–3.22 (m, 1 H), 3.73 (2 \times s, 2 \times 3 H, 2 \times OCH_3), 5.16–4.79 (br m, 2 H), 5.24 (m, 1 H), 5.54 (m, 2 H), 6.88 (m, 4 H, DMT), 6.37–7.20 (m, 9 H, DMT), 7.64 (s, 1 H), 7.89 (s, 1 H), 11.40 (s, 1 H, NH), ESI-MS: $\text{C}_{44}\text{H}_{44}\text{N}_4\text{O}_{11}$, calculated 804.30, found 803.29 $[\text{M}-\text{H}]^-$.

2.14. *2',3',4'-Tri-O-acetylriboflavin* 23. *2',3',4'-Tri-O-acetyl-5'-O-(4,4'-dimethoxytrityl)riboflavin* 22 (500 mg, 0.6 mmol) was dissolved in 20 mL nitromethane. To this solution, anhydrous ZnBr₂ (1.25 g, 5.6 mmol) was added. After 2 minutes, the reaction was quenched by addition of ammonium acetate (50 mL, 1 M in water). DCM (50 mL) were added, the organic layer was separated and washed with brine (2 × 20 mL), followed by evaporation under reduced pressure. The crude product was extracted several times with ice cold diethyl ether to remove residual DMT. In this way, compound 23 (250 mg, 0.5 mmol) were obtained as a yellow solid. The product was used for the next reaction without further purification. MALDI-TOF: C₂₃H₂₆N₄O₉, calculated 502.17, found 503.23 [M+H]⁺.

2.15. *3-(Tetraethylene glycol)propionic Acid tert-Butyl Ester* 25. Tetraethylene glycol 24 (38.8 g, 200 mmol) was dissolved in dry THF (80 mL). Under argon atmosphere, sodium (0.2 g, 8.7 mmol) was added. After all of the sodium had dissolved, *tert*-butylacrylate (7.7 g, 60 mmol) were added and the mixture was stirred at room temperature for 23 hours. The slightly yellow solution was neutralized with 1 M HCl and the solvents were evaporated under reduced pressure. The residue was redissolved in brine (80 mL) and extracted with ethyl acetate (3 × 50 mL). The combined organic layers were washed with water and dried with sodium sulfate. After evaporation of the solvents, the oil was subjected to column chromatography on silica gel (DCM/MeOH 98:2) to give 25 as colorless oil. Yield: 6.65 g (20.6 mmol, 34%). ¹H NMR (300 MHz, CDCl₃): δ (ppm) = 1.45 (s, 9H, *tert*-butyl), 2.50 (t, 2H, CH₂-COO-*tert*-Bu), 3.74–3.60 (m, 18H, OCH₂CH₂O). MALDI-TOF: C₁₅H₃₀O₇, calculated 322.20, found 345.16 [M+Na]⁺.

2.16. *3-(Tetraethylene glycol)propionic Acid tert-Butyl Ester Phosphoramidite* 26. *3-(Tetraethylene glycol)propionic acid tert-butyl ester* 25 (1 g, 3.1 mmol) was coevaporated three times with dichloromethane and then dissolved in 25 mL DCM under argon. To this solution, freshly distilled DIPEA (2.5 mL) and 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (0.85 mL, 3.6 mmol) were added. After 1.5 hours, the reaction was stopped by addition of 1 mL methanol. The mixture was diluted with ethylacetate (20 mL) and TEA (1.5 mL) and washed with a saturated solution of NaHCO₃ (25 mL) and brine (25 mL). The organic layer was dried with Na₂SO₄, and the solvents were removed under reduced pressure. The crude product was coevaporated three times with dry DCM and stored overnight in an evacuated desiccator with calcium chloride. According to TLC analysis, the phosphoramidite was sufficiently pure to be used for the next reaction without further purification. ³¹P NMR (300 MHz, CDCl₃), two diastereomers: δ (ppm) = 148.45, 149.37.

2.17. *5'-O-(3-(Tetraethylene glycol)propionic Acid tert-Butyl Ester-β-cyanoethylphosphoryl)-2',3',4'-tri-O-Acetylriboflavin* 27. All starting materials were coevaporated separately with dry DCM. The riboflavine derivative 23 (100 mg, 0.2 mmol) was dissolved in dry MeCN (15 mL). To this

solution, phosphoramidite 26 (136 mg, 0.26 mmol) in 2 mL dry MeCN and BMT (192 mg, 1 mmol) in 3 mL MeCN were added. The reaction mixture was stirred for 1 hour at room temperature. A solution of 0.2 M iodine in a mixture of THF/pyridine/H₂O (2:1:1) was added. After ten minutes, sodium bisulfite (5% in water) was added until the brown color had vanished. The mixture was diluted with DCM (40 mL). The organic layer was washed with sodium bicarbonate and brine and finally dried over sodium sulfate. The solvents were evaporated under reduced pressure. The residue was purified by reversed phase HPLC. Column VP125 Nucleosil 120-10 C18 ec (Macherey-Nagel), flow rate 0.5 mL/min, Eluent A 10% MeOH, Eluent B 70% MeOH, Gradient 0–70% B in 5 CV, 70–100% B in 3 CV, 100% B 3 CV. Yield: 59.5 mg (63 nmol, 32%). ¹H NMR (300 MHz, CDCl₃), two diastereomers: δ (ppm) = 1.44 (s, 9H, *tert*-butyl), 1.79 (s, 3H), 2.21 (2 × s, 3H), 2.35 and 2.33 (2 × s, 3H), 2.44 (s, 3H), 2.50 (t, 2H, CH₂-COO-*tert*-Bu), 2.57 (s, 3H), 2.81 (m, 2H, CH₂-CN), 3.74–3.60 (m, 18H, OCH₂CH₂O), 4.40–4.21 (m, 6H, P-O-CH₂), 5.33–4.89 (br, 2H), 5.41 (m, 1H), 5.50 (m, 1H), 5.65 (m, 1H), 7.60 and 7.59 (2 × s, 1H), 8.02 (s, 1H), 8.80 and 8.76 (2 × s, 1H, NH). MALDI-TOF: C₄₁H₅₈N₅O₁₈P, calculated 939.35, found 962.62 [M+Na]⁺.

2.18. *5'-O-(3-(Tetraethylene glycol)propionic Acid-β-cyanoethylphosphoryl)-2',3',4'-tri-O-Acetylriboflavin* 28. The fully protected coupling compound 27 (25 mg, 26.6 nmol) was treated with 1 mL DCM/TFA (1:1) for 1 hour at room temperature. Subsequently, the solvents were evaporated under reduced pressure, and the residue was purified by reversed phase HPLC to give the free acid as an orange solid. Column VP125 Nucleosil 120-10 C18 ec (Macherey-Nagel), flow rate 0.5 mL/min, Eluent A 10% MeOH, Eluent B 70% MeOH, Gradient 0–70% B in 5 CV, 70–100% B in 3 CV, 100% B 3 CV. Yield 18.8 mg (21.3 nmol, 80%). ¹H NMR (300 MHz, DMSO-*d*₆), two diastereomers: δ (ppm) = 1.61 (s, 3H), 2.20 (s, 3H), 2.25 (s, 3H), 2.41 (s, 3H), 2.43 (t, 2H, CH₂-COOH), 2.51 (s, 3H), 2.93 (m, 2H, CH₂-CN), 3.63–3.48 (m, 18H, OCH₂CH₂O), 4.41–4.10 (m, 6H, P-O-CH₂), 4.83 (d, 1H), 5.18–5.00 (br, 1H), 5.32 (m, 1H), 5.48 (m, 2H), 7.73 (s, 1H), 7.90 (s, 1H), 11.40 (s, 1H, NH). MALDI-TOF: C₃₇H₅₀N₅O₁₈P, calculated 883.29, found 906.43 [M+Na]⁺.

2.19. *5'-O-(3-(Tetraethylene glycol)propionic Acid Phosphoryl)riboflavin* 29. In 4 mL of a 30% ammonia solution in methanol, the riboflavine derivative 28 (15 mg, 17 nmol) was dissolved and stirred at room temperature. After 30 min, MALDI analysis showed complete conversion to product. The solvents were evaporated, and the solid was taken up in 1 mL deionized water and subjected onto a short DOWEX column (H⁺ form, 4 × 1.5 cm). The product was eluted with water and dried *in vacuo*. Yield: 11 mg (15.6 nmol, 92%). ¹H NMR: (600 MHz, DMSO-*d*₆): δ (ppm) = 2.40 (s, 3H), 2.43 (t, 2H, CH₂-COOH), 2.49 (s, 3H), 3.61–3.47 (m, 16H, OCH₂CH₂O), 3.64 (m, 1H), 3.86 (m, 1H), 3.90 (m, 1H), 4.00 (m, 2H), 4.11 (m, 1H), 4.24 (m, 1H), 4.65 (d, 1H),

4.89 (m, 1 H), 7.90 (s, 2 H), 11.35 (s, 1 H, NH). MALDI-TOF: $C_{28}H_{41}N_4O_{15}P$, calculated 704.23, found 705.32 $[M+H]^+$.

2.20. Introduction of Spin Labels to RNA Containing 4-Thiouridine. The modified nucleotide 4-thiouridine was incorporated at preselected positions of RNAs, to be used for postsynthetic spin labeling with 1-oxyl-2,2,5,5-tetramethylpyrrolidine-3-methylmethane-thiosulfonate or *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide. RNA samples were incubated for 2 hours at room temperature with a 200-fold excess of dithiothreitol (DTT) in 50 μ L of 100 mM sodium phosphate buffer (pH 8.0). The reducing agent was removed by demineralization using centrifugal filters, followed by lyophilization of the RNA. Spin labeling was carried out overnight by incubation of a 100 μ M solution of the respective RNA in 90 mM sodium phosphate buffer (90% v/v) (pH 8.0) and DMF (10% v/v) with a 10-fold excess of 1-oxyl-2,2,5,5-tetramethylpyrrolidine-3-methylmethanethiosulfonate or *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide, respectively. Unbound nitroxide reagent was removed by demineralization and lyophilization as described above. The spin label efficiency ranged from 10% to 100%.

2.21. Introduction of Spin Labels to RNA Containing 2'-Aminouridine. RNAs were synthesized with 2'-aminouridine at predefined positions and subsequently spin labeled.

2.21.1. Reaction with (1-Oxyl-2,2,5,5-tetramethylpyrrolidine-3-carboxylat)-*N*-Hydroxysuccinimidyl Ester. A solution of the RNA in 80 mM borax buffer (pH 8.5) (90% v/v), DMF (10% v/v) containing 10 mM NaCl was incubated for 5 hours at 37°C with a 500-fold excess of (1-oxyl-2,2,5,5-tetramethylpyrrolidine-3-carboxylat)-*N*-hydroxysuccinimidyl ester. Side products and unreacted reagents were extracted with 300 μ L chloroform at room temperature followed by precipitation of the labeled RNA from ethanol. The pellet was resolved in water, the solution was demineralized using centrifugal filters, and finally, the sample was lyophilized. Yields were between 70% and 80%.

2.21.2. Reaction with 4-Isocyanato-2,6-tetramethylpiperidyl-*N*-oxid. A solution of the RNA in 50 mM borax buffer (pH 8.5) (50% v/v), DMF (20% v/v) and formamide (30% v/v) was cooled to -8°C and incubated for 2 hours with a 15-fold excess of 4-isocyanato-2,6-tetramethylpiperidyl-*N*-oxid. Side products and unreacted reagent were extracted with 300 μ L chloroform at room temperature followed by precipitation of the labeled RNA from ethanol. The pellet was resolved in water, the solution was demineralized using centrifugal filters, and finally, the sample was lyophilized. Yields were between 75% and 100%.

3. Results and Discussion

3.1. Synthesis of Natural and Modified RNA with Commercially Available Building Blocks. The coupling reaction of RNA phosphoramidites is less facile than that of DNA

phosphoramidite building blocks; hence, the coupling yields normally are slightly poorer. Thus, chemical synthesis of RNA is limited by the length of the oligomers. Up to now, we, therefore, preferred to synthesize nonmodified RNA molecules which are longer than 70 bases by *in vitro* transcription techniques as described [17]. Shorter nonmodified and modified RNAs in our laboratory are prepared by solid phase synthesis. Our procedure is based on 2'-*O*-*tert*-butyldimethylsilyl (TBDMS) nucleoside protection and involves sequential couplings of β -cyanoethyl-(*N,N'*-diisopropyl)phosphoramidites of 5'-*O*-dimethoxytrityl-2'-*O*-TBDMS nucleosides essentially by the method as described [17]. For assembly of the oligoribonucleotides, we predominantly use CPG or polystyrene beads to which the first ribonucleoside derivative is attached via its 3'-OH group as a succinate linkage. The standard synthesis is on the micro-mole scale, and for coupling of phosphoramidites, we use 5-benzylmercaptotetrazole activation [16] and coupling times of 5 min. Under these conditions, the yields in couplings are usually in the 99% range. We have synthesized a variety of amino- and 4-thiouridine-modified RNAs (Tables 1, 2, and 3) that were postsynthetically functionalized with fluorescent dyes, spin labels, or flavine mononucleotide as described below.

The 5'- and 3'-amino modifiers, as well as 4-thiouridine were introduced during RNA solid phase synthesis using commercially available building blocks (Figure 1). While for the 5'-amino modifier, the standard coupling conditions were used, 4-thiouridine was introduced by two consecutive couplings of 5 minutes each using fresh amidite solution. After chain assembly was completed, cleavage from the solid support and removal of base-protecting groups were accomplished by treatment with aqueous NH_3 (30%) and ethanolic methylamine in a 1:1 mixture for 2 h at room temperature. We found this protocol advantageous over the usage of NH_3/MeOH , as we had applied previously [18]. Finally, 2'-*O*-TBDMS groups were removed by fluoride ion treatment. While 1 M tetrabutylammonium fluoride was conventionally used for this purpose, we prefer the use of triethylamine trihydrofluoride first suggested by Gasparutto et al. [19]. In combination with polar aprotic solvents and at elevated temperatures, this reagent offers advantages in speed and reliability [20]. Therefore, we use triethylamine trihydrofluoride/DMF (3:1) at 55°C for 1.5 hours for efficient removal of the 2'-*O*-TBDMS groups. Following deprotection, oligonucleotides were desalted by precipitation with *n*-butanol and purified by electrophoresis on denaturing polyacrylamide gels or by HPLC. The yields of oligonucleotides shown in Tables 1–3 were in the range of 1 to 400 nmol after purification.

3.2. Synthesis of Amino-Modified RNA Building Blocks and Incorporation in RNA. Within a collaborative research project, the three-dimensional architecture of structured RNAs, for example, RNA thermometers [7] or RNA four-way junctions [8], is investigated by quantitative fluorescence resonance energy transfer (FRET) measurements. For absolute distance measurements by FRET, the microenvironment of the probe attached to the sample is very important. However,

TABLE 1: Sequences of synthesized linker-modified RNA oligonucleotides. The bold letters (U/A) mark the position of the modification (see Figure 2).

Name	Length	Sequence	Mod.	Dye
A1L1	25 nts	UGG CAA GCU CGC AGU CGG CAC CGC C	5	Cy5
A1L2	25 nts	UGG CAA GCU CGC AGU CGG CAC CGC C	6	Cy5, ATTO 647N
A1L3	25 nts	UGG CAA GCU CGC AGU CGG CAC CGC C	7	Cy5
A2L1	25 nts	U GG CAA GCU CGC AGU CGG CAC CGC C	5	Cy5
A2L2	25 nts	U GG CAA GCU CGC AGU CGG CAC CGC C	6	Cy5, ATTO 647N
A2L3	25 nts	U GG CAA GCU CGC AGU CGG CAC CGC C	7	Cy5
D1L1	25 nts	GGC G GU GCC GAC UGC GAG CUU GCC A	5	Alexa 488
D1L2	25 nts	GGC G GU GCC GAC UGC GAG CUU GCC A	6	Alexa 488
D1L3	25 nts	GGC G GU GCC GAC UGC GAG CUU GCC A	7	Alexa 488
L2-TEST	8 nts	GGA AUU CC	6	—
2-L3-A	8 nts	GGA AUU CC	7	ATTO 647N
8-L1-A	8 nts	UUA GUA CU	8	—
2-L3-A	8 nts	GGA AAU CC	10	ATTO 647N
RNA-25	32 nts	CCC CAC GUC AAG GCG UGG UGG CCG AAG GUC GG	6	Cy5
RNA-26	32 nts	CCG ACC U UC GGC CAC CUG ACA GUC CUG UGG GG	6	Alexa 488
4U-Ther.	60 nts	GGC G GU GCC GAC UGC GAG CUU GCC AUG UUG AAC UUU UGA AUA GUG AUU CAG GAG GUU AAU	6	Alexa 488
4U-Ther.	60 nts	GGC G GU GCC GAC UGC GAG CUU GCC AUG UUG AAC UUU UGA AUA GUG AUU CAG GAG GUU AAU	6	Alexa 488
4U-Ther.	60 nts	GGC G GU GCC GAC UGC GAG CUU GCC AUG UUG AAC UUU UGA AUA GUG AUU CAG GAG GUU AAU	6	Alexa 488
4U-Ther.	60 nts	GGC G GU GCC GAC UGC GAG CUU GCC AUG UUG AAC UUU UGA AUA GUG AUU CAG GAG GUU AAU	6	Alexa 488

the commercial availability of building blocks for internal labeling of RNA is rare. Therefore, often molecules with a deoxy sugar moiety at the position of the internal label are applied for structural studies. In order to provide a set of nucleotides that upon incorporation in RNA can be used for RNA functionalization at a desired internal site, we started an effort to synthesize various ribonucleoside phosphoramidite building blocks with amino linkers of different lengths and flexibility attached to the heterocyclic base. We choose one short (L1) and one long (L2) rather flexible linker (Figure 3) coupled to the nucleobases by Heck chemistry [12, 13] and a short alkynyl linker (L3) (Figure 3) that was introduced by Sonogashira coupling [11].

In a first row of experiments, we have attached the three different linkers at C5 of uridine by palladium catalyzed cross-coupling reactions as described [14] (see 5, 6, and 7 in Figure 2).

Furthermore, adenosine derivatives with alkenyl linkers L1 and L2 at C8 or C2 were prepared under analogous conditions as used for uridine derivatives 5 and 6. However, a striking difference was the formation of double bond isomers in the Heck coupling reaction. Even though this phenomenon

is well known in the field of palladium catalyzed cross-coupling reactions, we were surprised observing a strong influence of the nature of the catalysts on formation of either isomer. Thorough investigation of a variety of catalysts and reaction conditions allowed us to tune reaction towards one of the two possible isomers by the choice of the catalysts (H. Nguyen, B. Appel, and S. Müller, manuscript in preparation). The specific amino-modified adenosine derivatives shown in Figure 2 were obtained with $\text{Na}_2[\text{PdCl}_4]$ (8, 9, and 12) or with the catalyst palladacycle (11) [21].

The phosphoramidite building blocks 7, 10, 13, and 14 carrying the propargylamino linker (Figure 2) were prepared by Pd-catalyzed Sonogashira cross-coupling. Due to the milder conditions compared with Heck chemistry, successful Sonogashira reaction could be carried out at fully protected nucleosides. Protection is not essential; however, we found introduction of sugar and base protecting groups prior to the cross-coupling reaction advantageous, since the overall yield in this case was higher. The synthesis starts with the introduction of the DMT- and TBDMS protecting group via standard procedures [22, 23]. Then, the propargylamino linker L3 was coupled to the nucleobase by Sonogashira

TABLE 2: Sequences of spin labelled RNAs. The bold U marks the position of 4-thiouridine, and the underlined U the position of 2'-aminouridine.

Name	Sequence	Spin label	Yield (%)
VW1	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	25
		TMP-IAA	23
VW2	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	10
		TMP-IAA	63
VW3	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	20
		TMP-IAA	80
VW4	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	77
VW5	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	25
VW6	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	100
VW7	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	61
VW8	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	13
VW9	GGG CUA AAA CAU ACC AGA UUU CGA UCU GGA GAG GUG AAG AAU ACG ACC ACCUAG CUC	TPM-MTS	45
VW10	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	89
VW11	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	89
VW12	GGG CUA AAA CAU ACC AGA UUUCGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC	TPM-MTS	10
HPAS11	ACC AGA GAA ACA GCC U <u>U</u> A GGA UAU GCU GG	TMC-NHS	81
		TEMPO-NCO	83
HPAS12	CCA GCA GAA GGA CG <u>U</u> CGU AUA UUA CCU GGU	TMC-NHS	80
		TEMPO-NCO	77
		TEMPO-NCO	97
HPAS13	ACC AGA GAA ACA ACC U <u>U</u> G GC	TMC-NHS	82
		TEMPO-NCO	97
HPAS14	GGC AAG G <u>U</u> C GUA UAU UAC CUG GU	TMC-NHS	85
		TEMPO-NCO	75
HPAS15	ACC AGA GAA ACA GAC U <u>U</u> G GC	TMC-NHS	72
		TEMPO-NCO	100
HPAS16	GCC AAG <u>U</u> CG UAU AUU ACC UGG U	TMC-NHS	80
		TEMPO-NCO	82

reaction, and finally, the phosphoramidite was prepared as described above for the C5-propargylamino uridine derivative.

Adenosine derivatives with the propargyl amino linker at C8 and C2 were prepared following the same route. However, as an additional step, the exocyclic amino group of the base was protected prior to introduction of the other protecting groups and of the linker unit.

The propargylamino linker was also attached to C8 of guanosine starting from 8-bromoguanosine. The exocyclic amino group was protected as an amide using isobutyric anhydride. The DMT and the TBDMS group were introduced according to standard methods [22, 23] as described above, and finally, the phosphoramidite was prepared with 4 eq. of ethyl diisopropyl amine and 1.5 eq. of 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (Figure 4).

TABLE 3: Sequences of two-stranded or three-stranded amino modified aptazymes. APU: 5-(3-Aminoprop-1-enyl)uridine. AMU: 2'-amino-uridine.

Name	Length	Sequence	Mod. X
HPAT2-L12	72 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAU AXG CUU CGG CAG AAG GAC GUC GUA UAU UAC CUG GUA	APU
HPAT2-L14	72 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAX AUG CUU CGG CAG AAG GAC GUC GUA UAU UAC CUG GUA	APU
HPAT2-U12	72 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAU AXG CUU CGG CAG AAG GAC GUC GUA UAU UAC CUG GUA	AMU
HPAT2-U14	72 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAX AUG CUU CGG CAG AAG GAC GUC GUA UAU UAC CUG GUA	AMU
HPAT1-AAM	43 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAX AUG CUG G	AMU
HPAT1-AHI	43 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAX AUG CUG G	APU
HPAT1-BAM	43 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAU AXG CUG G	AMU
HPAT1-BHI	43 nts	AAA GAG AGA AGU GAA CCA GAG AAA CAG CCU UAG GAU AXG CUG G	APU
HPAT2	31 nts	CCA GCA GAA GGA CGU CGU AUA UUA CCU GGU A	—

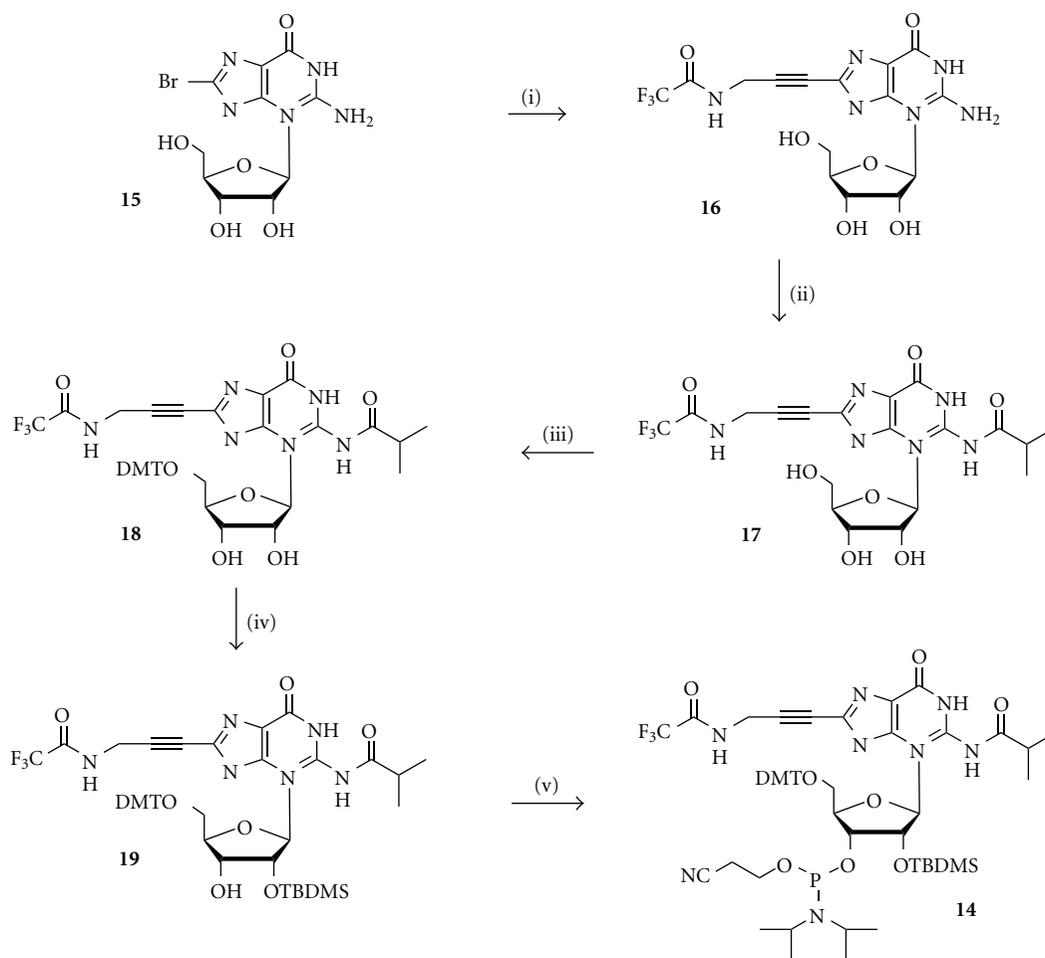


FIGURE 4: Synthesis of the guanosine-phosphoramidite. (i) 8-bromoguanosine **15**, 2.0 eq. L3, 0.1 eq. Pd(PPh₃)₄, 0.2 eq. CuI, 1.2 eq. DIPEA, 40°C, 20 h (80%); (ii) **16**, pyridine, 8.2 eq. TMS-Cl, 0°C, 2 h, 2.2 eq. isobutyric anhydride, rt, 2 h, NH₃ up to 2.5 M, 0°C; (iii) **17**, pyridine, 1.5 eq. DMT-Cl, 0°C, 2.5 h; (iv) **18**, THF, 3.7 eq. pyridine, 1.5 eq. AgNO₃, 15 min, 1.7 eq. TBDMS-Cl, rt, 8.5 h; (v) **19**, DCM, 4 eq. DIPEA, 1.5 eq. 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, rt, 2.5 h.

Several of the prepared phosphoramidite building blocks were successfully incorporated into oligonucleotides using standard solid phase chemistry (Table 1) [17]. The modified phosphoramidites were thoroughly dried by repetitive coevaporating with dichloromethane. After solving the amidite in dry acetonitrile to a concentration of 0.1 to 0.15 M, molecular sieves were added and the solution was kept at room temperature for one hour prior to use. Amino-modified phosphoramidites were coupled for 5 min, as were the standard building blocks. The coupling yield was in the 99% range, comparable with those of the unmodified phosphoramidites. After synthesis, oligonucleotides were deprotected under standard conditions [17], leading also to removal of the TFA group protecting the aliphatic amine of the linker. Purification was carried out by gel electrophoresis, followed by MALDI-TOF MS analysis. Table 4 shows some example masses of MALDI-TOF-MS determination for some synthesized RNA oligonucleotides.

3.3. Postsynthetic Labeling. The amino-modified RNAs were subsequently used for attachment of various dyes. Due to selective reaction of activated carboxylic acids with aliphatic amines, the fully deprotected RNA was used without the danger of significant side reactions at the sugar hydroxyl groups or aromatic amino groups of the bases. In majority, activate esters of Cy5, ATTO647N and Alexa488 were used for postsynthetic labeling (Figure 5).

For Cy5 and ATTO647N, *N*-hydroxy-succinimidyl (NHS) esters were used. For Alexa488, the NHS-ester is only available as anisomeric mixture of 5- and 6-esters (Figure 6). After reaction with the dye, RNAs labeled with either isomer can be separated by gelelectrophoresis. However, we found it being more efficient using the tetrafluorophenyl (TFP) ester of Alexa488, which is available as the pure 5-isomer (Figure 6).

For each coupling reaction, an amount of 5 to 10 nmol RNA was dissolved in borax buffer at pH 8.7 for labeling with Cy5 and ATTO647N and pH 9.4 for reaction with Alexa488. The dyes were added dissolved in DMF, and the reaction was carried out at room temperature overnight in the dark. Afterwards, the excess of dye was removed by gel filtration. Purification of dye-labeled RNAs was carried out by PAGE (Cy5- and Alexa488-labeled strands) or by RP-HPLC (ATTO647N-labeled oligonucleotides). Successful labeling was confirmed by MALDI-TOF MS analysis. The data shown in Table 5 are representative examples of MS analyzed dye labeled oligonucleotides.

Figure 7 shows the MALDI-TOF-MS spectra of the oligonucleotide 2-L3-A prior to and after labeling with ATTO647N, as an example for the superior quality of the labeled oligonucleotides.

3.4. Hybridization of Single-Stranded RNAs for Structural Studies. To demonstrate that the amino linkers as well as the fluorescent labels do not intervene in the hybridization of single-stranded oligonucleotides, the hybrid of Alexa Fluor 488-labeled D1L3 and the complementary Cy5-labeled A1L3 was prepared (Figure 8). The oligonucleotides and their hybrid were visualized by illumination with UV-light at 254

and 365 nm. As can be seen from the gel, the two dye-labeled single-stranded oligonucleotides were fully hybridized.

In a parallel experiment, Alexa488-labeled D1L3 and Cy5-labeled A1L3 were dissolved in KH_2PO_4 - K_2HPO_4 buffer containing MgCl_2 and KCl, according to the required buffer conditions in the FRET experiments to be conducted. The mixture was mixed, incubated at 78°C for 3 minutes, and slowly cooled down to room temperature. Reaction analysis by native polyacrylamide gel electrophoresis (data not shown) confirmed quantitative hybridization. Several of the synthesized RNAs upon hybridization to their complementary strands were used in fluorescence measurements, in particular in single molecule detection (SMD) and time correlated single photon counting (TCSPC) bulk experiments [14].

3.5. Synthesis of an Activated FMN Derivative for Postsynthetic Labeling of Amino-Functionalized RNA. In previous work, we have engineered an FMN responsive hairpin aptazyme that can be switched on and off in a reversible mode [10]. As we have shown, FMN can be removed from its binding pocket by reduction of the isoalloxazine ring, which is associated with a change in geometry and subsequently with loss of binding. In the original system, FMN was externally added, requiring a 400-fold excess to saturate its binding site in the aptazyme and to obtain a functional response. Current work is focused on engineering a hairpin aptazyme with FMN being covalently attached via a suitable linker. This small molecular device is supposed to be advantageous in terms of entropic costs for FMN binding to the aptamer and iterative cycles of reversible switching. For this purpose, a suitable FMN derivative for postsynthetic labeling of amino-modified RNA was synthesized. Recent experiments with different FMN derivatives have shown that only the isoalloxazine ring system is required for interaction with the aptazyme, while the ribose unit as well as the phosphate do not seem to play a major role (T. Marschall and S. Müller, unpublished observation). Thus, we decided to attach a carboxyl group containing linker to the ribose chain of riboflavin which subsequently could be postsynthetically attached to amino-functionalized RNA as described in Figures 5, 9, and 10.

Starting from riboflavin, first a fully protected riboflavin derivative was synthesized. Protection of the 5'-OH group was achieved by chemoselective tritylation of the primary hydroxyl group with DMT-Cl in pyridine. The remaining secondary hydroxyl groups were reacted with acetic anhydride to give the fully protected derivative **22** in good yields (83–90%). The selective deprotection of the 5'-OH group under acidic conditions (2% DCA in DCM) led to the formation of a side product with the same molecular mass as the main product. This suggested the migration of one of the acetyl groups under DMT deprotection conditions. It has been shown before that under acidic conditions, the acetyl group of the C4'-OH group can move to the neighbouring C5'-OH group [24]. Separation of the two isomers by column chromatography on silica gel was not successful. Therefore, we looked for alternative procedures. Alternatively to acids, also ZnBr_2 has been suggested as superior reagent

TABLE 4: Masses of the synthetic RNAs determined by MALDI-TOF-MS (linear negative-ion mode, found $M = [M-H]^-$).

Sequence	Cal. mass	Found mass	Δm	Error (%)
L2-TEST	2677.73	2675.99	0.77	0.027
2-L3-A	2585.63	2583.99	0.64	0.024
8-L1-A	2526.53	2525.96	0.43	0.017
A1L1	8041.95	8040.41	0.53	0.010
A1L2	8155.03	8153.15	0.88	0.011
A1L3	8039.93	8039.23	0.30	0.003
A2L1	8041.95	8044.29	3.36	0.040
A2L3	8039.93	8041.17	2.24	0.030
D1L1	8098.93	8098.62	0.69	0.010
D1L2	8212.03	8211.46	0.43	0.005
D1L3	8096.93	8099.73	3.80	0.040

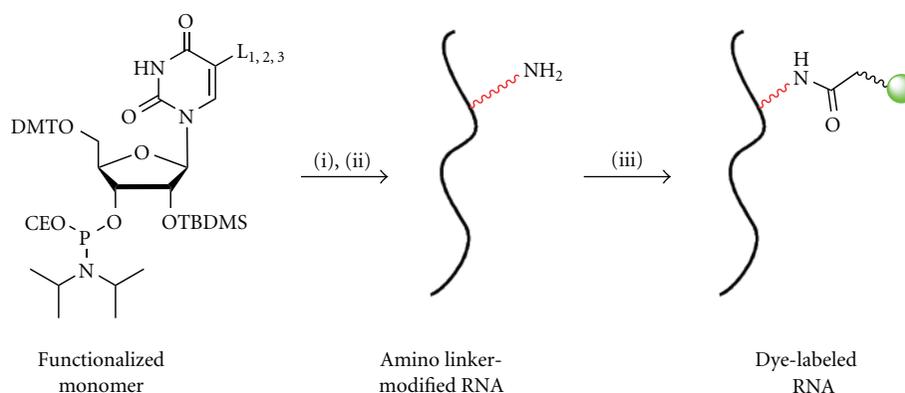


FIGURE 5: Postsynthetic labeling of amino-modified oligonucleotides: (i) Chemical synthesis of amino-modified RNA; (ii) Deprotection; (iii) Postsynthetic labeling with fluorescent dyes.

for removal of DMT groups [25, 26]. Indeed, using $ZnBr_2$ in dry nitromethane led to the formation of the desired product **23** within a few minutes, without acetyl migration. However, during purification of the detritylated riboflavin derivative by column chromatography, again migration of acetyl protecting groups was observed. Therefore, the solid crude product was just washed several times with ice cold diethyl ether, yielding compound **23** with satisfying purity.

For the synthesis of the carboxylic acid containing linker component, tetraethylene glycol was reacted with *tert*-butyl acrylate (Figure 10). The product of this Michael addition was converted to the corresponding phosphoramidite with 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite by standard procedures [27]. The coupling reaction between the riboflavin derivative **23** and the linker phosphoramidite **26** was accomplished in acetonitrile with BMT activation followed by oxidation with an aqueous solution of iodine to give the flavinmononucleotide derivative **27**. The fully protected FMN derivative was treated with trifluoroacetic acid in dichloromethane to eliminate the acid labile *tert*-butyl protecting group (**28**). After the purification of the free acid form of the FMN derivative by RP-HPLC, cleavage of the β -cyanoethyl and the acetyl groups was achieved by stirring in 30% ammonia in methanol. Evaporation of the ammonia solution and subsequent treatment with DOWEX (H^+ -form)

yielded the desired product **29** that upon activation as NHS ester was used for postsynthetic RNA functionalization.

The FMN sensitive hairpin aptazyme was chemically synthesized as a three-stranded or two-stranded system using the phosphoramidite approach. In order to facilitate coupling of the FMN derivative with the oligonucleotides, 2'-aminouridine or 5-propargylaminouridine (Figure 11) were introduced in predefined positions of the RNA chain, providing the aliphatic amine for selective attachment of the FMN derivative (Table 3). Suitable positions for FMN attachment have been defined on the basis of the structure model of the FMN aptamer [28] with the help of PyMOL.

In order to couple the FMN derivative to the amino-modified RNA, an NHS-ester was generated *in situ* by reacting the unprotected FMN derivative **29** with 1.2 eq. TSTU and 2 eq. DIPEA for one hour at room temperature in DMF (Figure 12). The formation of the activated ester was monitored by MALDI mass spectrometry. After all of the free carboxylic acid had been converted to its NHS-form, the FMN derivative was coupled to the amino group bearing oligonucleotide. Successful coupling was confirmed by polyacrylamide gel electrophoresis (Figure 12). The functional characterization of the resulting FMN-RNA constructs in ribozyme activity assays is currently in progress.

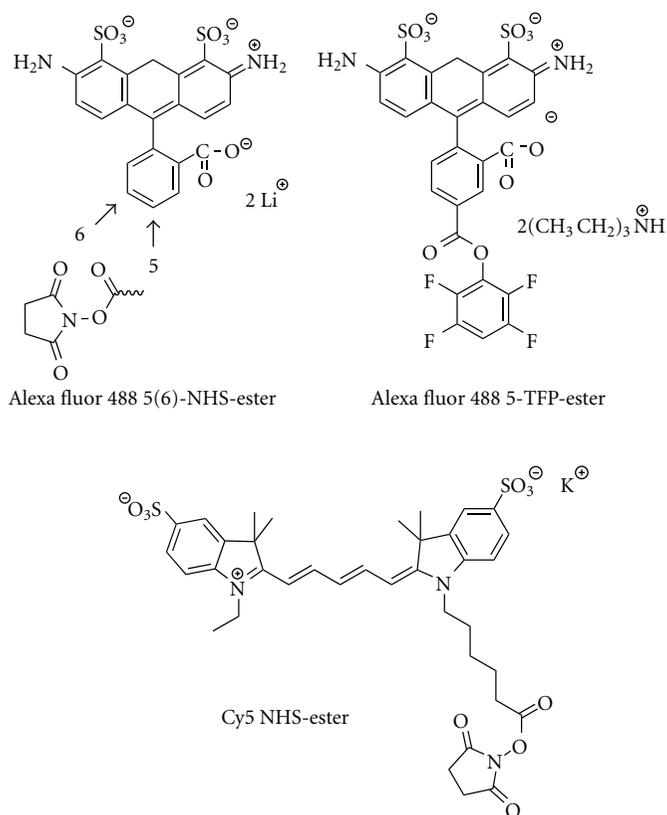


FIGURE 6: Activated derivatives of fluorescent dyes for postsynthetic labeling of RNA.

TABLE 5: Masses of dye-labeled oligonucleotides determined by MALDI-TOF-MS (linear negative-ion mode, found $M = [M+H]^-$).

Labeled sequence	Cal. mass	Found mass	Δm	Error (%)
2-L3-A-ATTO 647N	3213.63	3212.44	0.19	0.006
A1L2-ATTO 647N	8783.03	8782.28	0.25	0.003
A2L2-ATTO 647N	8783.03	8776.95	5.08	0.057
A2L2-Cy5	8793.25	8791.53	0.72	0.008

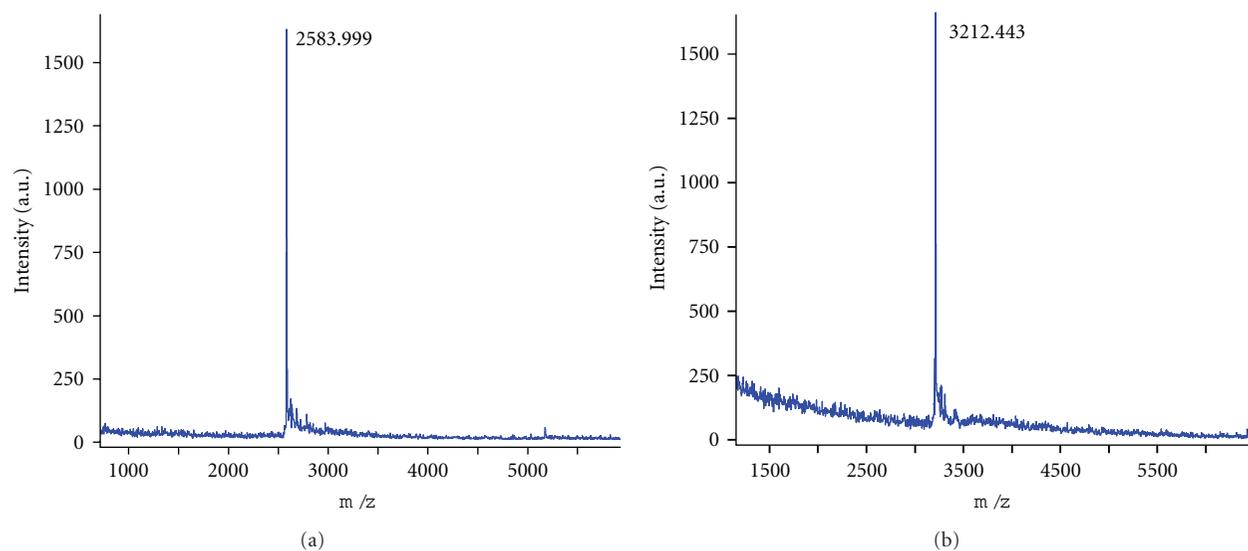


FIGURE 7: MALDI-TOF mass spectra of the amino-modified oligonucleotide 2-L3-A (a) and its corresponding conjugate with ATTO647N (b). For the sequence of 2-L3-A, compare Table 1.

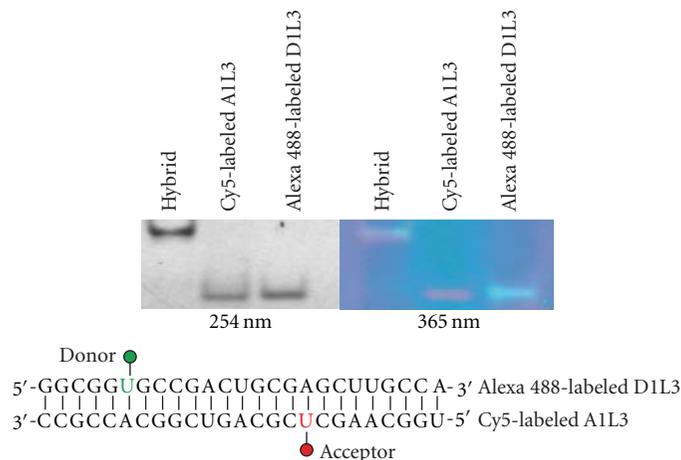


FIGURE 8: Hybridization of Alexa488-labeled D1L3 and Cy5-labeled A1L3: 3 μ M Alexa488-labeled D1L3: 3 μ M Cy5-labeled A1L3, 10 mM MgCl₂, 50 mM Tris buffer, pH 7.4, total volume 100 μ L, 78°C 3 min, slowly cooling to rt. The hybrid was analyzed by 15% native PAA-gel: (left) illuminated at 254 nm, (right) illuminated at 365 nm.

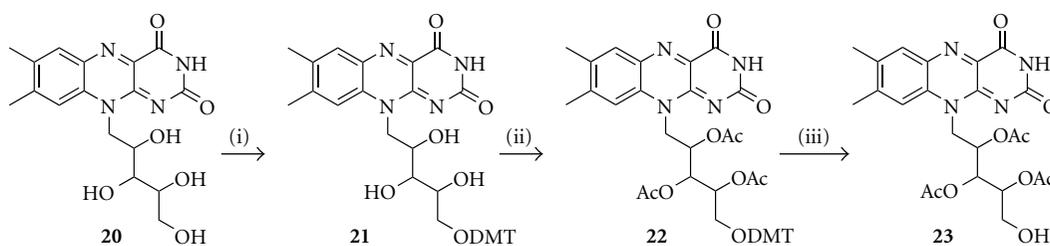


FIGURE 9: Synthesis of protected riboflavine derivatives: (i) DMT-Cl, DMAB, pyridine, rt, overnight (53%); (ii) acetic anhydride, pyridine, 0°C \rightarrow rt, overnight (90%); (iii) ZnBr₂ in nitromethane, 2 min (83%).

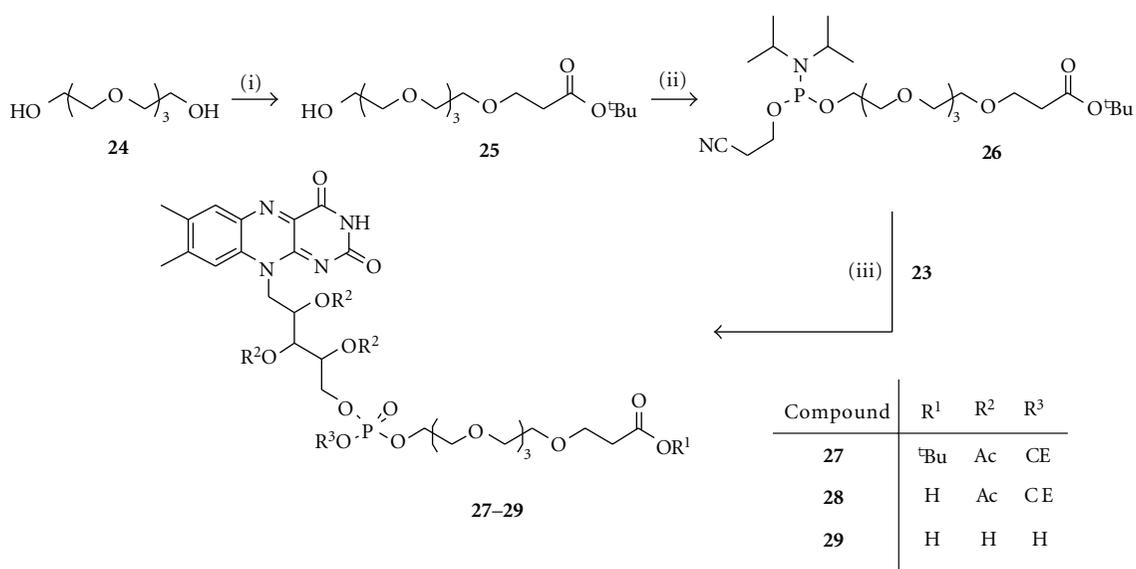


FIGURE 10: Synthesis of the fully protected linker unit: (i) Na, *tert*-butyl acrylate, THF, rt, 1 d (34%); (ii) 2-Cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 1.5 h (quant.); (iii) (a) BMT, MeCN, rt, 1 h; (b) 0.2 M I₂ in THF/pyridine/H₂O (2 : 1 : 1) (32%).

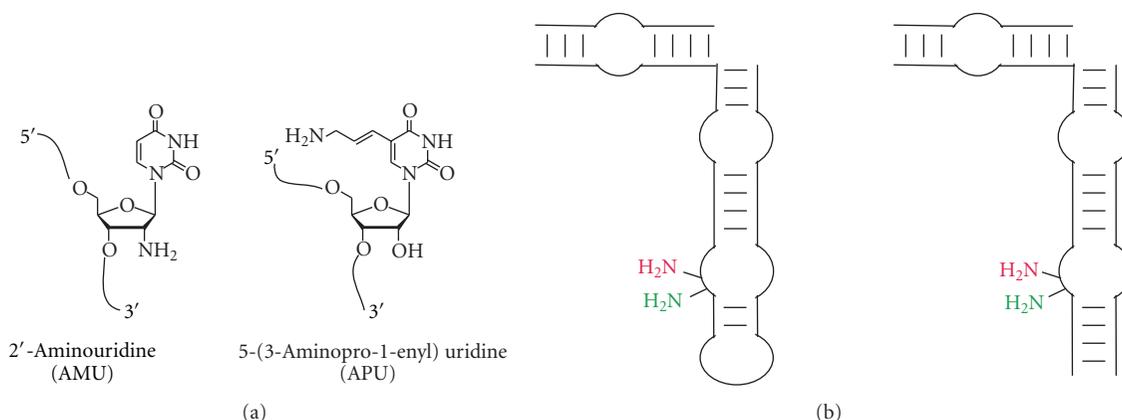


FIGURE 11: (a) Structure of the two nucleotide analogs used for the incorporation of an aliphatic amino group into the aptazyme RNA. (b) Two-stranded (left) and three-stranded (right) aptazyme. The modifications AMU and APU are located at position U12 (red) and U14 (green), respectively.

3.6. Synthesis of Specifically Spin-Labeled RNAs. Within a collaborative project on structural analysis of the tetracycline dependent riboswitch [9] and of the FMN responsive aptazyme described above [10] by electron paramagnetic resonance (EPR) spectroscopy, we have synthetically prepared RNAs carrying a suitable spin label at a predefined site. The basic strategy of site-directed spin-labeling originally has been developed with proteins and involves modification of the SH group of an appropriate cysteine with a selective paramagnetic nitroxide reagent. Applying the same technique to RNA, labeling initially was achieved by incorporating a guanosine monophosphorothioate at the 5'-end followed by spin-labeling at the sulfur [29]. The strategy was extended towards site-specific introduction of a phosphorothioate group at RNA backbone locations or of modified bases such as 4-thiouridine and subsequent attachment of a suitably activated nitroxide spin label [30–32]. More recent work employs aliphatic amino groups at the 2'-position of a specific sugar residue or attached at the heterocyclic base [33–35].

We have made use of two different strategies, incorporating either 4-thiouridine or 2'-aminouridine into RNA followed by postsynthetic reaction with the nitroxide spin label. For reaction with 4-thiouridine containing RNAs, (1-oxyl-2,2,5,5-tetramethylpyrrolin-3-methyl)methanethiosulfonate (TPM-MTS) and *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl) iodoacetamide (TMP-IAA) was used, for reaction with 2'-aminouridine containing RNA, (1-oxyl-2,2,5,5-tetramethylpyrrolin-3-carboxylate)-*N*-hydroxy-succinimidyl ester (TMC-NHS) or 4-isocyanato-2,6-tetramethylpiperidyl-*N*-oxid (TEMPO-NCO) (Figure 13). As mentioned already above, the 12 modified RNAs VW1 to VW12 shown in Table 2 were synthesized within a structural study of a tetracycline responsive RNA aptamer [9]. HPAS11 and HPAS16 were synthesized and spin labeled for an EPR study of an artificial flavin mononucleotide responsive riboswitch [10].

For RNAs labeled via 4-thiouridine, reaction with the TPM-MTS spin label was favored, because initial EPR measurements showed that the TMP-IAA label has suboptimal

properties, presumably because it is too mobile due to the rather flexible CH₂-CO-NH-CH₂-spacer (Figure 13, lower part). Therefore, the majority of 4-thiouridine containing RNAs was reacted with TPM-MTS for introduction of the less mobile nitroxide spin label. The yields shown in Table 2 were reproduced in up to five labeling reactions for each individual RNA. Thus, the variation of labeling efficiency observed for individual RNAs suggests that the structure of the RNA aptamer significantly influences the labeling reaction, making spin labeling of individual uridines dependent on the specific location in the aptamer sequence.

A second point that may account for the observed low labeling efficiency is the rather unstable disulfide bond attaching the spin label at the uracil residue. Therefore, as an alternative, we have also studied conditions for spin labeling via the RNA sugar-phosphate backbone (Figure 14). To this end, RNAs containing 2'-aminouridine at a specific position (HPAS11 to HPAS16) were synthesized and labeled using either TEMPO-NCO or TMC-NHS. It has been described in the literature that the steric hindrance of the 2'-NH₂-group hampers reaction with NHS esters [36]. Therefore, we have extended the reaction time from 15 minutes to 4 hours and used a 500-fold excess of the NHS-reagent to overcome the problem of low reaction yields. On the contrary to NHS esters, isocyanates were shown to virtually quantitatively react with 2'-amino sugars [33]. However, being less selective, side reactions at the nucleobases were observed when a high excess of the isocyanate reagent was used [37]. The authors have optimized reaction conditions such that a 2'-amino group was selectively modified without concurrent side reaction. In that, they observed an increase of the yield of the modified oligonucleotide upon lowering the temperature, which presumably is the result of repression of competing isocyanate hydrolysis [37]. In parallel rows of labeling reactions under varying conditions, we have also looked at reaction conditions to deliver the spin-labeled RNA with high yield and high quality. Best results were obtained with an only 15-fold excess of the isocyanate reagent over RNA, and a reaction temperature of -8°C . Under these conditions, about 80% labeling efficiency were reproducibly

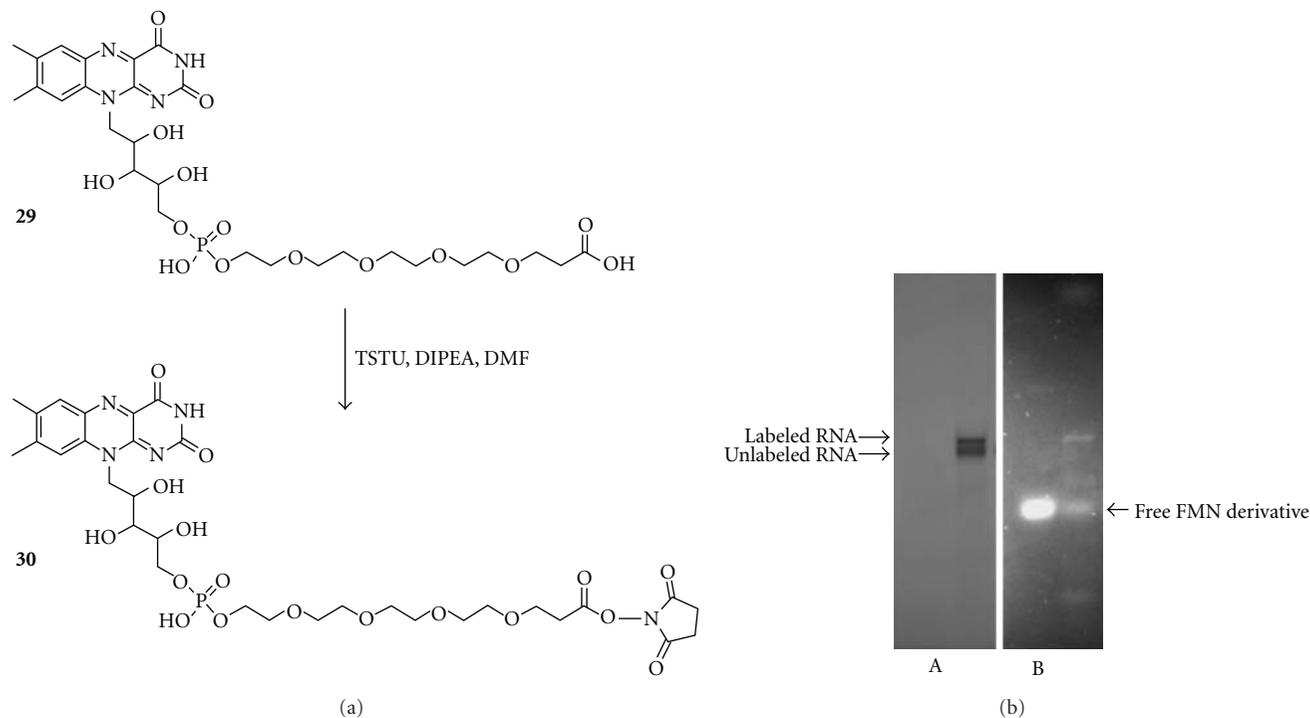


FIGURE 12: (a) Formation of the *N*-hydroxysuccinimidyl (NHS) ester. (b) PAGE analysis of the coupling reaction between amino-modified RNA HPAT1-BHI (for sequence see Table 3) and the FMN derivative. Visualization is at UV 254 nm (A) and 365 nm (B).

obtained for both labeling reagents with all RNA samples (Table 2). In order to avoid side reactions, we have added the spin label reagent only once in 15-fold excess even though it was originally described that repeated additions of the reagent give nearly quantitative results [33]. EPR spectra of RNAs showed that the spin labels conjugated via the 2'-amino group are rather immobile (H.-J. Steinhoff, University Osnabrück, oral information), making labeling via this position most favorable. It was somewhat surprising that mobility of TEMPO seemed to be even more limited than that of the TPC- label, because the 6-membered piperidyl ring is conformationally more flexible than the 5-membered pyrrolidine ring. Obviously, the nature of the spacer unit (urea versus amide bond) also has a strong influence. Overall, in the specific context of RNA sequences shown in Table 2, 2'-amino sugar labeling is advantageous over 4-thiouridine labeling in terms of efficiency and immobility.

4. Conclusion

As demonstrated above, the chemical synthesis of RNA offers the possibility of site-specific introduction of modified nucleosides. This is an important advantage over enzymatic procedures, making RNA synthesis an essential tool for a large number of studies in different fields. We have succeeded in the synthesis of a number of monomer building blocks that upon incorporation in RNA allow for postsynthetic attachment of spectroscopic reporter groups or other desired functionalities. The amino-modified monomers presented in this paper are a valuable addition to the RNA chemist's tool

box. Due to the different length and flexibility of the linkers that carry the aliphatic amino groups for postsynthetic functionalization, molecular entities that become attached to the RNA have differing degrees of motional freedom. This is of particular interest for fluorescence spectroscopic studies, because systematic errors may result from fluorescence dyes that are sterically hindered, or on the contrary are too mobile, limiting the accuracy of the measurement. Therefore, using our amino linker building blocks of the U series (Figure 2), an extensive analysis of the influence of the linker lengths and rigidity on the accuracy of FRET measurements was carried out [14]. The results will pave the way for more sophisticated studies of RNA structure dynamics. Likewise, EPR spectroscopic analysis of RNA will profit from the new building blocks, since spin labels attached to the RNA of interest, are subjected to the same conditions and restrictions as the fluorescent dyes in fluorescence spectroscopy.

Lastly, we have synthesized an FMN-RNA conjugate to be used for functional studies of an FMN responsive aptazyme. This example shows that there is much room for RNA engineering and construction of sophisticated RNA devices. Organic chemistry offers a large number of possibilities to prepare suitably activated modifiers that can be postsynthetically attached to amino-modified RNA, or to RNA carrying other functionalities such as for example thiol groups, or alkynes and azides as used for labeling by Click chemistry [38]. Thus, in combination with the development of synthetic routes to appropriately activated modifiers, RNAs can be conjugated with a large variety of tags, making the nucleic acid not only visible in spectroscopic

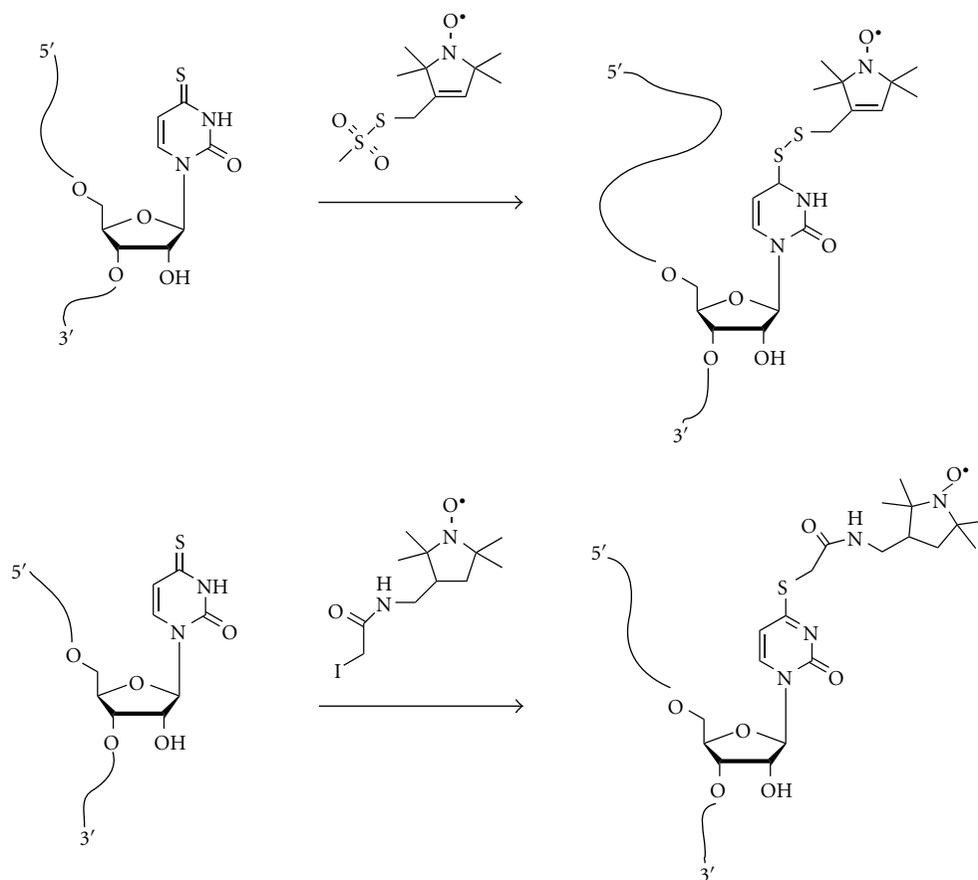


FIGURE 13: Strategy for postsynthetic introduction of spin labels to 4-thiouridine containing RNA.

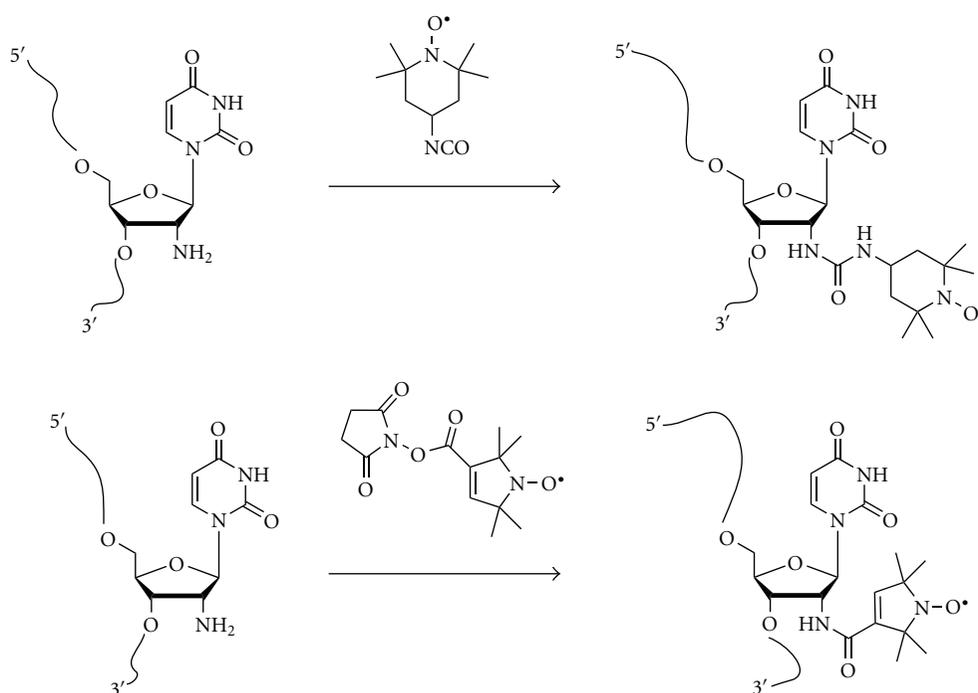


FIGURE 14: Labeling of the 2'-amino group of a model RNA.

experiments, but also adding new functionality as demonstrated with the RNA-FMN conjugate described herein.

Abbreviations

CE:	Cyanoethyl
DCM:	Dichloromethane
DIPEA:	Diisopropylethylamine
DMA:	Dimethylamine
DMAP:	Dimethylaminopyridin
DMF:	Dimethylformamide
DMT:	Dimethoxytrityl
MeCN:	Acetonitril
TBDMS:	<i>tert</i> -butyldimethylsilyl
TEA:	Triethylamine
BMT:	1-(Benzylmercapto)-1 <i>H</i> -tetrazole
TEMPO-NCO:	4-isocyanato-2,6-tetramethylpiperidyl- <i>N</i> -oxyl
TFA:	Trifluoroacetic acid
TMC-NHS:	(1-oxyl-2,2,5,5-tetramethylpyrrolin-3-carboxylat)- <i>N</i> -hydroxysuccinimidyl ester
TMP-IAA:	<i>N</i> -(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide
TPM-MTS:	(1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)methanethiosulfonate.

References

- [1] A. Mescalchin and T. Restle, "Oligomeric nucleic acids as antivirals," *Molecules*, vol. 16, no. 2, pp. 1271–1296, 2011.
- [2] I. Drude, V. Dombos, S. Vauléon, and S. Müller, "Drugs made of RNA: development and application of engineered RNAs for gene therapy," *Mini-Reviews in Medicinal Chemistry*, vol. 7, no. 9, pp. 912–931, 2007.
- [3] J. Liu, Z. Cao, and Y. Lu, "Functional nucleic acid sensors," *Chemical Reviews*, vol. 109, no. 5, pp. 1948–1998, 2009.
- [4] F. Meng, W. E. Hennink, and Z. Zhong, "Reduction-sensitive polymers and bioconjugates for biomedical applications," *Biomaterials*, vol. 30, no. 12, pp. 2180–2198, 2009.
- [5] S. Müller, J. Wolf, and S. A. Ivanov, "Current strategies for the synthesis of RNA," *Current Organic Synthesis*, vol. 1, no. 3, pp. 293–307, 2004.
- [6] M. H. Caruthers, "A brief review of DNA and RNA chemical synthesis," *Biochemical Society Transactions*, vol. 39, no. 2, pp. 575–580, 2011.
- [7] F. Narberhaus, T. Waldminghaus, and S. Chowdhury, "RNA thermometers," *FEMS Microbiology Reviews*, vol. 30, no. 1, pp. 3–16, 2006.
- [8] T. J. Wilson, M. Nahas, T. Ha, and D. M. J. Lilley, "Folding and catalysis of the hairpin ribozyme," *Biochemical Society Transactions*, vol. 33, no. 3, pp. 461–465, 2005.
- [9] D. Wunnicke, D. Strohbach, J. E. Weigand et al., "Ligand-induced conformational capture of a synthetic tetracycline riboswitch revealed by pulse EPR," *RNA*, vol. 17, no. 1, pp. 182–188, 2011.
- [10] D. Strohbach, N. Novak, and S. Müller, "Redox-active riboswitching: allosteric regulation of ribozyme activity by ligand-shape control," *Angewandte Chemie—International Edition*, vol. 45, no. 13, pp. 2127–2129, 2006.
- [11] K. A. Cruickshank and D. L. Stockwell, "Oligonucleotide labelling: a concise synthesis of a modified thymidine phosphoramidite," *Tetrahedron Letters*, vol. 29, no. 41, pp. 5221–5224, 1988.
- [12] S. Dey and T. L. Sheppard, "Ketone-DNA: a versatile postsynthetic DNA decoration platform," *Organic Letters*, vol. 3, no. 25, pp. 3983–3986, 2001.
- [13] C. M. McKeen, L. J. Brown, J. T. G. Nicol, J. M. Mellor, and T. Brown, "Synthesis of fluorophore and quencher monomers for use in Scorpion primers and nucleic acid structural probes," *Organic and Biomolecular Chemistry*, vol. 1, no. 13, pp. 2267–2275, 2003.
- [14] S. Sindbert, S. Kalinin, H. Nguyen et al., "Accurate distance determination of nucleic acids via FRET: implications of dye linker length and rigidity," *Journal of the American Chemical Society*, vol. 133, no. 8, pp. 2463–2480, 2011.
- [15] C. Schmidt, R. Welz, and S. Müller, "RNA double cleavage by a hairpin-derived twin ribozyme," *Nucleic Acids Research*, vol. 28, no. 4, pp. 886–894, 2000.
- [16] R. Welz and S. Müller, "5-(Benzylmercapto)-1*H*-tetrazole as activator for 2'-*O*-TBDMS phosphoramidite building blocks in RNA synthesis," *Tetrahedron Letters*, vol. 43, no. 5, pp. 795–797, 2002.
- [17] I. Drude, A. Strahl, D. Galla, O. Müller, and S. Müller, "Design of hairpin ribozyme variants with improved activity for poorly processed substrates," *FEBS Journal*, vol. 278, no. 4, pp. 622–633, 2011.
- [18] R. Welz, K. Bossmann, C. Klug, C. Schmidt, H.-J. Fritz, and S. Müller, "Site-directed alteration of RNA sequence mediated by an engineered twin ribozyme," *Angewandte Chemie—International Edition*, vol. 42, no. 21, pp. 2424–2427, 2003.
- [19] D. Gasparutto, T. Livache, H. Bazin et al., "Chemical synthesis of a biologically active natural tRNA with its minor bases," *Nucleic Acids Research*, vol. 20, no. 19, pp. 5159–5166, 1992.
- [20] F. Wincott, A. DiRenzo, C. Shaffer et al., "Synthesis, deprotection, analysis and purification of RNA and ribozymes," *Nucleic Acids Research*, vol. 23, no. 14, pp. 2677–2684, 1995.
- [21] W. A. Herrmann, C. Brossmer, K. Öfele et al., "Palladacycles as structurally defined catalysts for the heck olefination of chloro- and bromoarenes," *Angewandte Chemie—International Edition*, vol. 34, no. 17, pp. 1844–1848, 1995.
- [22] K. Shah, H. Y. Wu, and T. M. Rana, "Synthesis of uridine phosphoramidite analogs: reagents for site-specific incorporation of photoreactive sites into RNA sequences," *Bioconjugate Chemistry*, vol. 5, no. 6, pp. 508–512, 1994.
- [23] G. H. Hakmelahi, Z. A. Proba, and K. K. Ogilvie, "New catalysts and procedures for the dimethoxytritylation and selective silylation of ribonucleosides," *The Canadian Journal of Chemistry*, vol. 60, pp. 1106–1113, 1982.
- [24] C. Frier, J.-L. Décout, and M. Fontecave, "Method for preparing new flavin derivatives: synthesis of flavin-thymine nucleotides and flavin-oligonucleotide adducts," *Journal of Organic Chemistry*, vol. 62, no. 11, pp. 3520–3528, 1997.
- [25] V. Kohli, H. Blöcker, and H. Köster, "The triphenylmethyl (trityl) group and its uses in nucleotide chemistry," *Tetrahedron Letters*, vol. 21, no. 28, pp. 2683–2686, 1980.
- [26] M. D. Matteucci and M. H. Caruthers, "The use of zinc bromide for removal of dimethoxytrityl ethers from deoxynucleosides," *Tetrahedron Letters*, vol. 21, no. 34, pp. 3243–3246, 1980.
- [27] T. Atkinson and M. Smith, "Solid phase synthesis of oligodeoxyribonucleotides by the phosphite-triester method," in *Oligonucleotide Synthesis: A Practical Approach*, M. J. Gait, Ed., IRL Press Oxford, 1984.

- [28] P. Fan, A. K. Suri, R. Fiala, D. Live, and D. J. Patel, "Molecular recognition in the FMN-RNA aptamer complex," *Journal of Molecular Biology*, vol. 258, no. 3, pp. 480–500, 1996.
- [29] J. C. Macosko, M. S. Pio, I. Tinoco Jr., and Y.-K. Shin, "A novel 5'-displacement spin-labeling technique for electron paramagnetic resonance spectroscopy of RNA," *RNA*, vol. 5, no. 9, pp. 1158–1166, 1999.
- [30] P. Z. Qin, K. Hideg, J. Feigon, and W. L. Hubbell, "Monitoring RNA base structure and dynamics using site-directed spin labeling," *Biochemistry*, vol. 42, no. 22, pp. 6772–6783, 2003.
- [31] P. Z. Qin, S. E. Butcher, J. Feigon, and W. L. Hubbell, "Quantitative analysis of the isolated GAAA tetraloop/receptor interaction in solution: a site-directed spin labeling study," *Biochemistry*, vol. 40, no. 23, pp. 6929–6936, 2001.
- [32] A. Ramos and G. Varani, "A new method to detect long-range protein-RNA contacts: NMR detection of electron-proton relaxation induced by nitroxide spin-labeled RNA," *Journal of the American Chemical Society*, vol. 120, no. 42, pp. 10992–10993, 1998.
- [33] T. E. Edwards, T. M. Okonogi, B. H. Robinson, and S. T. Sigurdsson, "Site-specific incorporation of nitroxide spin-labels into internal sites of the TAR RNA: structure-dependent dynamics of RNA by EPR spectroscopy," *Journal of the American Chemical Society*, vol. 123, no. 7, pp. 1527–1528, 2001.
- [34] N. Piton, Y. Mu, G. Stock, T. F. Prisner, O. Schiemann, and J. W. Engels, "Base-specific spin-labeling of RNA for structure determination," *Nucleic Acids Research*, vol. 35, no. 9, pp. 3128–3143, 2007.
- [35] T. E. Edwards and S. T. Sigurdsson, "Site-specific incorporation of nitroxide spin-labels into 2'-positions of nucleic acids," *Nature Protocols*, vol. 2, no. 8, pp. 1954–1962, 2007.
- [36] S. B. Cohen and T. R. Cech, "Dynamics of thermal motions within a large catalytic RNA investigated by cross-linking with thiol-disulfide interchange," *Journal of the American Chemical Society*, vol. 119, no. 27, pp. 6259–6268, 1997.
- [37] S. T. Sigurdsson and F. Eckstein, "Site specific labelling of sugar residues in oligoribonucleotides: reactions of aliphatic isocyanates with 2' amino groups," *Nucleic Acids Research*, vol. 24, no. 16, pp. 3129–3133, 1996.
- [38] A. H. El-Sagheer and T. Brown, "Click chemistry with DNA," *Chemical Society Reviews*, vol. 39, no. 4, pp. 1388–1405, 2010.

Research Article

Commercially Supplied Amine-Modified siRNAs May Require Ultrafiltration prior to Conjugation with Amine-Reactive Compounds

Shannen Lau, Bim Graham, Ben J. Boyd, Colin W. Pouton, and Paul J. White

Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, VIC 3052, Australia

Correspondence should be addressed to Paul J. White, paul.white@monash.edu

Received 10 November 2010; Accepted 8 February 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Shannen Lau et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Conjugation of siRNA to macromolecules such as serum albumin has multiple potential benefits, including enhanced extravasation via albumin-mediated transcytosis across endothelial cells and reduced renal clearance. In attempting to conjugate siRNA to albumin, we used commercially sourced amine-modified siRNA and reacted it with the heterobifunctional linker succinimidyl 4-[*N*-maleimidomethyl]cyclohexane-1-carboxylate (SMCC) to introduce a maleimide group suitable for conjugation to the thiol group of the surface-exposed cysteine residue (Cys 34) within albumin. We found the conjugation of the SMCC-treated siRNA to bovine serum albumin (BSA) to be very inefficient and investigated the cause of the low yield of conjugate. Ultrafiltration with phosphate-buffered saline prior to activation with SMCC dramatically increased the yield of siRNA-albumin conjugate (~15-fold). Communication with the commercial supplier revealed that ammonium acetate buffer was used in a desalting step as part of the siRNA purification process prior to supply, likely resulting in ammonium counterions to the siRNA polyanion, which would interfere with conjugation by consuming the SMCC. After ultrafiltration, a greatly reduced amount of SMCC could be used to affect conjugation, without significant reduction in yield. These data indicate that amine-modified siRNA sourced commercially may require ultrafiltration or dialysis prior to use in conjugation reactions.

1. Introduction

A series of obstacles impede the successful delivery of short interfering RNA (siRNA) to the cytosol of target cells [1, 2]. Strategies involving conjugation of siRNA to a wide variety of molecules have been employed to overcome these obstacles. Typically, these approaches employ amine- [3] or thiol-modified [4, 5] siRNA at 3' or 5' ends of one strand of the siRNA duplex (see [6] for review of conjugation chemistry). Conjugations of siRNA to polyethylene glycol [5], quantum dots containing targeting moieties [4], cell penetrating peptides such as penetratin [7], and a range of lipophilic molecules, including cholesterol [8], have been demonstrated without a loss of silencing ability.

Recently, we embarked on a project aiming to deliver siRNA to the heart using siRNA conjugated to the serum protein albumin. The strategy employed involved the use

of the common heterobifunctional linker succinimidyl 4-[*N*-maleimidomethyl]cyclohexane-1-carboxylate (SMCC). Using this approach, the coupling of SMCC to siRNA containing an aminoethyl pendant at the 3' end of the sense strand results in an "activated" siRNA bearing a thiol-reactive maleimide group. The "activated" siRNA is then covalently linked to serum albumin via reaction with the thiol-containing side chain of cysteine residue 34, located on the surface of serum albumin molecules (see Figure 1). Previous studies have shown that a number of thiol-reactive drugs rapidly and selectively bind to endogenous albumin within a few minutes of administration, due to that fact that free thiol groups are not found on the majority of circulating serum proteins except for albumin [9]. The chemistry used to conjugate the siRNA to albumin is not novel; others have conjugated nanoparticles to siRNA using a similar cross-linker, *N*-succinimidyl-3-(2-pyridyldithio)propionate (SPDP) [10].

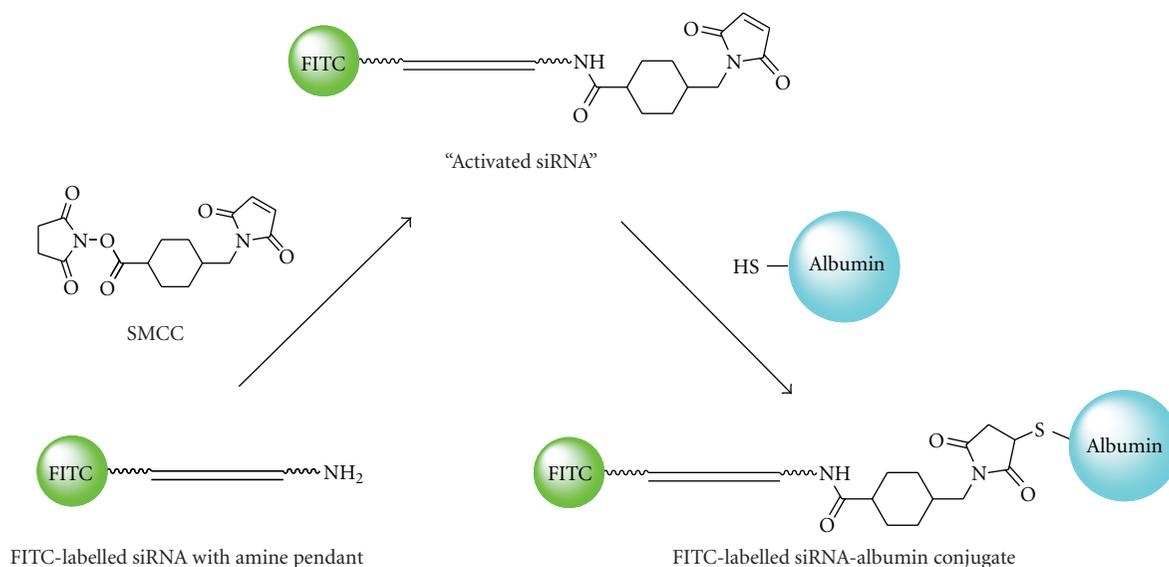


FIGURE 1: Conjugation chemistry. Formation of an amide bond between the heterobifunctional cross-linker SMCC and commercially sourced amine-modified siRNA produces maleimide-functionalized siRNA (activated siRNA). The activated siRNA reacts with cysteine-34 of BSA to produce an siRNA-BSA conjugate.

These workers employed a 40:1 SPDP:siRNA ratio to activate the siRNA and then gel filtration to remove the unreacted SPDP, prior to conjugation. We found significant batch-to-batch variability in the efficiency of the siRNA conjugation to bovine serum albumin (BSA) using similar conditions and therefore sought to understand the factors limiting the conjugation. Here, we report a successful strategy to restore conjugation efficiency, by repeated buffer exchange prior to exposing the amine-modified siRNA to SMCC.

2. Experimental Methods

2.1. Materials. Samples of IGF-IR siRNA used were sourced from Dharmacon (CO, USA) and were modified with an amine for conjugation, with 2'-O-methyl modifications for *in vivo* use, and with a fluorescein label for imaging purposes. Each sense strand contained four 2'-O-methyl modifications, whilst the antisense strand contained two 2'-O-methyl modified nucleotides. The sense strand contained a 3' amine modification (N6) and a fluorescein label (Fl) at its 5' end. The sequence used was as follows, where "m" refers to 2'-O-methyl modification:

sense strand: 5'-Fl(*)GCmCmCAUmGUGUGAGAm-AGACC(*)dT(*)dT(*)N6-3',

antisense strand: 3'-dT(*)dT(*)CGGGUACACACm-UCUUCUmGG-5'.

(*) denotes phosphorothioate linkages.

The siRNA was dissolved in freshly prepared phosphate-buffered saline (PBS, pH 7.4) or 250 mM borate buffer (pH 8 or 8.5) to a concentration of 0.2 mM and stored at -20°C prior to use. BSA used was bovine serum albumin (BSA, Grade V Sigma USA). 4-[N-maleimidomethyl]cyclohexane-1-carboxylate (SMCC) was purchased from Pierce Chemical

CO (CA, USA) and dissolved in DMSO (9.57 mM) just prior to use.

2.2. Reaction Conditions for Conjugation without Ultrafiltration Step. siRNA-BSA covalent conjugates were formed by reacting the 3'-amine-modified, fluorescein-labeled siRNA with the heterobifunctional linker reagent SMCC to form an irreversible amide linkage (see Figure 1). Aliquots of siRNA and SMCC linker were combined to give a ratio of 1 nmol siRNA (0.2 nM) to 40 nmol SMCC (9.58 nM) in a ~1:1 mixture of MilliQ water and DMSO and allowed to react for 1 h at room temperature. Excess SMCC was then removed by gel filtration chromatography using a NAP-5 column (Pharmacia, UK), with only the fluorescent flow-through collected. The quantity of siRNA collected was then determined by fluorescence using an EnVision Multilabel Plate Reader (USA) with 488 nm excitation/520 nm emission. The activated siRNA was then coupled to the surface-exposed cysteine of BSA by incubating the siRNA with 2 molar equivalents of BSA for time periods of up to 24 h (final concentrations of siRNA and BSA in the reaction mixture were 30.3 and 60.6 μM, resp.).

The components present in the reaction mixture were examined on a 10% SDS-PAGE, together with BSA and siRNA standards of known molecular weight (66.5 and 14 kDa, resp.). The gels were loaded with 1 μg protein per well and run for 90 min at a constant voltage of 100 V in SDS running buffer. siRNA was visualized under UV light or using Typhoon gel scanner for detection of fluorescein emission at 520 nm, and BSA was visualised following Coomassie Brilliant Blue staining. Quantitation of the relative amounts of conjugate and unreacted siRNA was performed by densitometric analysis of siRNA and BSA bands recorded on gel scans.

2.3. Reaction Conditions for Conjugation with Ultrafiltration Step. Prior to conjugation, the reconstituted IGF-IR siRNA was subjected to four rounds of ultrafiltration with PBS buffer using an Amicon Ultra-4 centrifugal 3 kDa filter unit spun at 4000 rpm for 50 min at 4°C after each addition of fresh buffer. The concentrated siRNA was removed from the filter unit and quantified on an EnVision Multilabel Plate Reader (USA). Conjugation was then performed as described above.

3. Results and Discussion

As documentation supplied with the commercially supplied amine-modified siRNA did not contain any information to suggest the contrary, it was presumed that the siRNA would be suitable for conjugation to amine-reactive compounds after simply reconstituting in an appropriate buffer. However, initial attempts to couple the siRNA to BSA via the use of SMCC, a heterobifunctional cross-linker routinely employed to link amine- to thiol-bearing molecules, met with very limited success. Following the literature precedent, siRNA was reacted (activated) with a 40-fold molar excess of SMCC prior to exposure to BSA, as described in Section 2, and at the completion of the reaction samples were analysed via PAGE. Figure 2(a) shows a representative PAGE gel, used to evaluate conjugation efficiency; scanning under 488 nm excitation reveals siRNA standard at 14 kDa in lane 3, with lanes 4 and 5 showing a significant band corresponding to unreacted siRNA band and a faint band at the size expected for the siRNA-BSA conjugate, approximately 80 kDa. Densitometric analysis indicated that only 4-5% of siRNA was conjugated to BSA and that there was no difference in yield between a one-hour and three-hour activation time. Whilst there was some batch-to-batch variability, the majority of conjugation reactions performed using several separate batches of amine-modified siRNA gave similarly poor results, with less than 20% of siRNA in conjugate form. Figure 2(b) shows the same PAGE gel that is shown in Figure 2(a), in this case after Coomassie (protein) staining to reveal BSA. As expected, the BSA standard ran at 66 kDa. There was no clear band apparent at 80 kDa, likely indicating that the small amount of siRNA-BSA conjugate evident in Figure 2(a) was below the limit of detection using Coomassie staining.

In troubleshooting the inefficient conjugation reaction, a range of experimental parameters were varied in an attempt to drive the reaction to completion. The reaction was performed overnight, at pH values between 7 and 8.5, and using a range of siRNA, SMCC, and BSA concentrations and molar ratios. None of these strategies were successful (data not shown). Finally, we hypothesized that there may have been low-molecular-weight species in the siRNA as supplied that interfered with the conjugation reaction and, therefore, that extensive ultrafiltration or dialysis prior to activation with SMCC to remove these species might improve the outcome of the conjugation reaction. Communication with the manufacturer revealed that the siRNA was desalted using ammonium acetate; an unwise choice for amine-modified

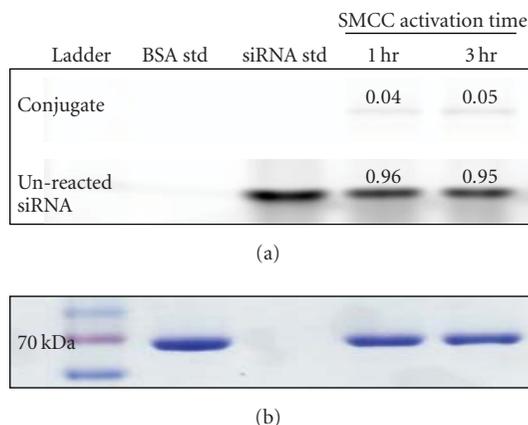


FIGURE 2: Inefficient conjugation using amine-modified siRNA without prior ultrafiltration to remove ammonium counterions. SDS-polyacrylamide gel showing outcome of siRNA-BSA conjugation reaction using 1- and 3-hour SMCC activation times. Amine-modified siRNA was activated with SMCC in a molar ratio of 40:1 for 1 or 3 h, followed by conjugation with BSA over 24 h. (a) Faint bands at the predicted conjugate size were observed under excitation with 488 nm light, with the majority of siRNA unreacted. (b) Only a single unconjugated albumin band was revealed by Coomassie Brilliant Blue staining in each case. siRNA std: siRNA standard; BSA std: BSA standard.

siRNA whose likely purpose is conjugation via the amine functional group, since ammonium ions would become counterions to the polyanionic siRNA. In solution, there is an equilibrium between ammonium ions and ammonia. Ammonia can act as a nucleophile, similar to the primary amine in the modified siRNA, and hence is also able to react with the succinimidyl ester group in the SMCC linker.

Figure 3(a) (i) shows the reaction products before (lane 4) and after (lane 5) four rounds of ultrafiltration with PBS. Ultrafiltration prior to SMCC activation resulted in a 15-fold increase in the proportion of siRNA converted to the conjugate form, as evidenced by the band at 80 kDa. After Coomassie staining of this gel, a clear band at 80 kDa was evident in addition to the native BSA band, indicating the BSA component of the conjugate (Figure 3(a) (ii)). Densitometric analysis of the unreacted and conjugate BSA bands indicated that approximately one third of the BSA was involved in the reaction. This is not unexpected given that approximately 70% of albumin is in the thiol reactive (mercaptalbumin) form [11] and that a 2:1 albumin:siRNA ratio was employed. Thus, it appears that removal of the ammonium counterions from the amine-modified siRNA prevents the majority of the SMCC from being consumed by a competitive side reaction, and thus it allows the reaction of siRNA with SMCC to proceed more effectively.

Having demonstrated that the reaction efficiency was dramatically improved by ultrafiltration using PBS, we anticipated that the ratio of SMCC linker to siRNA used (40:1)

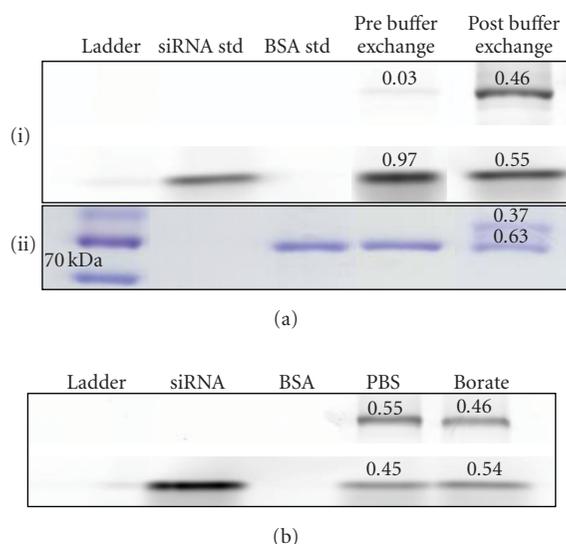


FIGURE 3: Ultrafiltration dramatically increases siRNA-albumin conjugate yield. (a) SDS-polyacrylamide gel showing outcome of siRNA-BSA conjugation reaction without prior ultrafiltration (lane 4) and subsequent to ultrafiltration of siRNA (lane 5). A 15-fold increase in conjugate band intensity was observed using ultrafiltered siRNA, compared to amine-modified siRNA simply reconstituted in buffer, when scanned at 488 nm. A visible band at the predicted conjugate size was also apparent after staining with Coomassie Brilliant Blue. (b) siRNA conjugation was efficient after buffer exchange with either PBS (pH 7.4) or borate buffer (pH 8). siRNA std : siRNA standard; BSA std : BSA standard.

was unnecessarily high. Figure 4 shows evaluation of the effect of reducing the SMCC:siRNA ratio on the outcome of the conjugation reaction. There was no reduction in the proportion of siRNA in conjugate form upon reducing the SMCC:siRNA ratio from 40:1 to 10:1, and only a modest decrease occurred when the ratio was further reduced to 2:1 (48% of siRNA in conjugate form versus 61% for 40:1 ratio, Figure 4(a)). Thus, it was possible to reduce the SMCC concentration during the reaction, with the advantage of reducing the amount of excess linker remaining after the reaction, as well as lowering the cost of synthesis.

4. Conclusions

These data clearly demonstrate that siRNA supplied commercially with terminal amine groups can require ultrafiltration or dialysis prior to exposure to linker reagents in order to remove ammonium ions remaining after the desalting step used in the purification process. Researchers planning to perform conjugation experiments should consult with the supplier as to the purification conditions, and if it is known or suspected that ammonium-containing salts were used, an ultrafiltration step, as described herein, should be performed prior to use of the siRNA.

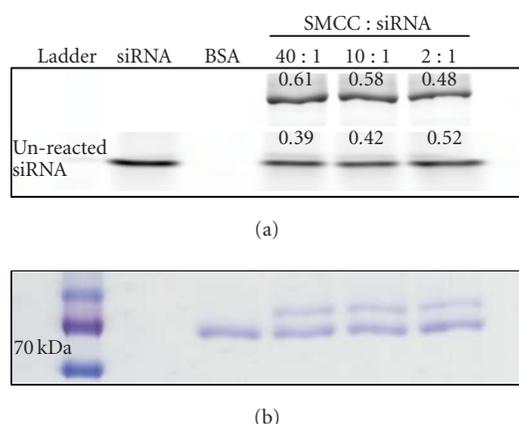


FIGURE 4: SMCC:siRNA ratio can be reduced without loss of yield after ultrafiltration of amine-modified siRNA. Conjugation of amine-modified siRNA with BSA using 40:1, 10:1, and 2:1 SMCC-to-siRNA molar ratios. Conjugates formation was observed for each of the SMCC:siRNA ratios; reducing the ratio from 40:1 to 2:1 had relatively little effect on conjugate yield. siRNA std : siRNA standard; BSA std : BSA standard.

Acknowledgment

The authors gratefully acknowledge the Heart Foundation of Australia (Grant-in-Aid G09M 4519).

References

- [1] R. Juliano, M. R. Alam, V. Dixit, and H. Kang, "Mechanisms and strategies for effective delivery of antisense and siRNA oligonucleotides," *Nucleic Acids Research*, vol. 36, no. 12, pp. 4158–4171, 2008.
- [2] P. J. White, "Barriers to successful delivery of short interfering RNA after systemic administration," *Clinical and Experimental Pharmacology and Physiology*, vol. 35, no. 11, pp. 1371–1376, 2008.
- [3] D. B. Rozema, D. L. Lewis, D. H. Wakefield et al., "Dynamic PolyConjugates for targeted in vivo delivery of siRNA to hepatocytes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 32, pp. 12982–12987, 2007.
- [4] A. M. Derfus, A. A. Chen, D. H. Min, E. Ruoslahti, and S. N. Bhatia, "Targeted quantum dot conjugates for siRNA delivery," *Bioconjugate Chemistry*, vol. 18, no. 5, pp. 1391–1396, 2007.
- [5] S. Jung, S. H. Lee, H. Mok, H. J. Chung, and T. G. Park, "Gene silencing efficiency of siRNA-PEG conjugates: effect of PEGylation site and PEG molecular weight," *Journal of Controlled Release*, vol. 144, no. 3, pp. 306–313, 2010.
- [6] K. Lu, Q. P. Duan, L. Ma, and D. X. Zhao, "Chemical strategies for the synthesis of peptide-oligonucleotide conjugates," *Bioconjugate Chemistry*, vol. 21, no. 2, pp. 187–202, 2010.
- [7] A. Muratovska and M. R. Eccles, "Conjugate for efficient delivery of short interfering RNA (siRNA) into mammalian cells," *FEBS Letters*, vol. 558, no. 1–3, pp. 63–68, 2004.
- [8] J. Soutschek, A. Akinc, B. Bramlage et al., "Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs," *Nature*, vol. 432, no. 7014, pp. 173–178, 2004.

- [9] F. Kratz, "Albumin as a drug carrier: design of prodrugs, drug conjugates and nanoparticles," *Journal of Controlled Release*, vol. 132, no. 3, pp. 171–183, 2008.
- [10] H. R. Kim, I. K. Kim, K. H. Bae, S. H. Lee, Y. Lee, and T. G. Park, "Cationic solid lipid nanoparticles reconstituted from low density lipoprotein components for delivery of siRNA," *Molecular Pharmaceutics*, vol. 5, no. 4, pp. 622–631, 2008.
- [11] M. Sogami, S. Era, S. Nagaoka et al., "High-performance liquid chromatographic studies on non-mercapt α mercapt conversion of human serum albumin. II," *Journal of Chromatography A*, vol. 332, pp. 19–27, 1985.

Review Article

tRNA Modification and Genetic Code Variations in Animal Mitochondria

Kimitsuna Watanabe^{1,2} and Shin-ichi Yokobori¹

¹ Department of Molecular Biology, School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji, Tokyo 192-0392, Japan

² Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-Ku, Tokyo 113-8657, Japan

Correspondence should be addressed to Kimitsuna Watanabe, kwatanab@ft.catv.ne.jp

Received 31 May 2011; Accepted 4 July 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 K. Watanabe and S.-i. Yokobori. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In animal mitochondria, six codons have been known as nonuniversal genetic codes, which vary in the course of animal evolution. They are UGA (termination codon in the universal genetic code changes to Trp codon in all animal mitochondria), AUA (Ile to Met in most metazoan mitochondria), AAA (Lys to Asn in echinoderm and some platyhelminth mitochondria), AGA/AGG (Arg to Ser in most invertebrate, Arg to Gly in tunicate, and Arg to termination in vertebrate mitochondria), and UAA (termination to Tyr in a planaria and a nematode mitochondria, but conclusive evidence is lacking in this case). We have elucidated that the anticodons of tRNAs deciphering these nonuniversal codons (tRNA^{Trp} for UGA, tRNA^{Met} for AUA, tRNA^{Asn} for AAA, and tRNA^{Ser} and tRNA^{Gly} for AGA/AGG) are all modified; tRNA^{Trp} has 5-carboxymethylaminomethyluridine or 5-taurinomethyluridine, tRNA^{Met} has 5-formylcytidine or 5-taurinomethyluridine, tRNA^{Ser} has 7-methylguanosine and tRNA^{Gly} has 5-taurinomethyluridine in their anticodon wobble position, and tRNA^{Asn} has pseudouridine in the anticodon second position. This review aims to clarify the structural relationship between these nonuniversal codons and the corresponding tRNA anticodons including modified nucleosides and to speculate on the possible mechanisms for explaining the evolutionary changes of these nonuniversal codons in the course of animal evolution.

1. Introduction

Up to now six codons have been known which are deciphered by the corresponding tRNAs as amino acids different from those assigned by the universal genetic code in animal mitochondria (Figure 1) [1]. UGA termination codon in the universal genetic code is deciphered to Trp in all animal mitochondria, AUA Ile to Met in most metazoan except echinoderm, planarian, cnidarian, placozoan and poriferan mitochondria, AAA Lys to Asn in echinoderm and some platyhelminth mitochondria, and AGA/AGG Arg to Ser in most invertebrate mitochondria, Gly in tunicate (urochordata) mitochondria, and termination codon in vertebrate mitochondria. UAA termination codon was assumed to be a Tyr codon in a planaria [2] and a nematode mitochondria [3], but there is neither structural information on mt tRNA^{Tyr} that decodes the UAA codon, nor information

about the mitochondrial (mt) release factor relevant to this phenomenon. Thus, this issue is no more discussed here.

The codon-amino acid correspondence was first deduced by comparison of mt DNA sequence containing the codon with amino acid sequence of the corresponding protein [4]. Since mt proteins exist in a small number (in most cases, 13) and which are encoded by a small sized mt DNA (16,500 bp in the case of human mitochondria) [5], the correspondence can be unambiguously accomplished. In the next step, in order to analyze the molecular mechanism of the codon change, the corresponding tRNA sequence was analyzed especially focused on the anticodon sequence [6]. It is an advantageous point for analysis of mt tRNAs that in metazoan mitochondria, tRNA genes are restricted to 22~24 species on the mt genomes, and that tRNAs are not imported from the cytoplasm in almost all metazoan mitochondria [7] except for a few cases such as in Cnidaria. As the results,

several modified nucleosides such as 5-carboxymethylaminomethyl(2-thio)uridine (cmnm⁵(s²)U), 5-taurinomethyl(2-thio)uridine (τ m⁵(s²)U), 5-formylcytidine (f⁵C), 7-methylguanosine (m⁷G) in the anticodon first position, and pseudouridine (Ψ) in the anticodon second position were found to be involved in the genetic code variations (Figure 2), in which τ m⁵(s²)U [8] and f⁵C [9] are novel modified nucleosides found by our group. Thus, an expanded wobble rule was established in which cmnm⁵(s²)U, τ m⁵U, and f⁵C at the anticodon first position pair with A or G in the codon third position, and m⁷G pairs with all nucleotides in the codon third position [10, 11] (Figure 3). It was speculated that unmodified G in the anticodon first position should pair with C, U, and A in the codon third position in the case of fruit fly, *Drosophila melanogaster* tRNA^{Ser}_{GCU} for decoding AGU/AGC/AGA codons [11], and Ψ in the anticodon second position strengthens the pairing interaction with A in the codon second position in the case of echinoderm tRNA^{Asn}_{G Ψ U} for decoding AAA codon [12]. In the third step, the wobble pairings as inferred from the above-expanded wobble rule were confirmed by an *in vitro* translation system of animal mitochondria, in which *in vitro* translation was performed by *E. coli* or bovine mt translation system using synthetic polyribonucleotide made of a series of nonuniversal codon as a messenger, and incorporation of a certain amino acid corresponding to the nonuniversal codon was identified [13, 14]. In the fourth step, the wobble pairings were confirmed by an *in vitro* experiment, in which natural mRNA including a specific codon which was replaced with a certain nonuniversal codon was translated *in vitro*, and the mRNA activity was detected by enzymatic activity if the mRNA encodes a certain enzyme such as dihydrofolate reductase (DHFR) (Hanada, T., Suzuki, T. and Watanabe, K., unpublished results).

In this review, we summarize mostly the results obtained in the second step, together with a few cases in the third step, and inquire into the nature of codon-anticodon interaction in the mt translation process and the relationship between nonuniversal genetic code and modified nucleoside in the tRNA anticodon deciphering the codon, during the course of animal evolution. These studies may lead to the understanding as to how the genetic code evolves and how mt tRNAs keep up with the genetic code variations by providing modification at the anticodon of tRNAs.

It is also important to consider the involvement of aminoacyl-tRNA synthetase (aaRS) in the recognition of tRNA, especially in the case when aaRS recognizes the anticodon region of tRNA as the identity determinant. Since a few cases have been known about the animal mt aaRSs, some discussions will be added in the applicable sections, about the recognition mechanisms of aaRSs toward the corresponding tRNAs involved in the genetic code variations.

2. Genetic Code Variations and the Anticodon Structure of the Corresponding tRNA

2.1. cmnm⁵(s²)U for Decoding UGA/UGG Codons. In mitochondria of all animal phyla, UGA is read as Trp instead

of termination codon in the universal genetic code [4, 6]. In the protostome (nematode and probably platyhelminth) mitochondria, the anticodon first position (wobble position) is modified to 5-carboxymethylaminomethyluridine (cmnm⁵U) or 5-carboxymethylaminomethyl-2-thiouridine (cmnm⁵s²U) [15, 16]. Therefore, the modification of U to cmnm⁵(s²)U in the anticodon first position in tRNA^{Trp}_{UCA} may restrict the base pairing only with purine nucleosides (A/G, symbolized by R) at the codon third position (Figures 3(a) and 3(b)), which would overcome the competition with mitochondrial release factor. Thus, UGA codon is read as Trp. It is known that unmodified uridine (U) at the wobble position recognizes all four nucleosides (A/G/U/C, symbolized by N) at the codon third position [17]; thus, in mitochondria all four codon boxes are mostly read by the respective single tRNAs with unmodified U at the anticodon wobble position (see Figure 1).

cmnm⁵U was first found in yeast mitochondria [18]. This modified uridine is similar in the chemical structure with cmnm⁵Um (5-carboxymethylaminomethyl-O-2'-methyluridine) in *Escherichia coli* tRNA^{Leu}₄ [19] and mnm⁵U in *E. coli* tRNA^{Arg} and is considered to fix its conformation by interresidual hydrogen bonding. The mnm⁵U possesses the same side chain at position 5 of uracil base as mnm⁵s²U in tRNA^{Glu}, which enables to take "rigid" conformation for the construction of the C3'-endo form.

In the *in vitro* translation system using MS2 RNA as a messenger, tRNA^{Leu}₄ having cmnm⁵Um and tRNA^{Leu}₅ having 2'-O-methylcytidine (Cm) at the anticodon first position, both recognized UUA and UUG codons, but not UUU and UUC codons at all [20]. It was clarified by NMR analysis that the orthodox C3'-endo-G⁻ form of both cmnm⁵Um and Cm are very stable. It is considered that the posttranscriptional modification to form cmnm⁵Um and Cm fixes their conformations very rigid, which regulates not to recognize the UUU/UUC codons [21].

In summary, in the tRNAs recognizing A- or G-ending codons by dividing 2:2 in the codon box, U and C (in the case of f⁵C, see Section 2.3) at the anticodon first position are modified by introducing side chains to position 5 of the nucleobase, 2-thiolation at position 2 of the nucleobase, or methylation at position 2 of the ribose moiety. By combining these modifications, the conformation of the nucleoside becomes more rigid and which guarantees the precise recognition toward NNA/NNG codons (Figures 3(a) and 3(b)) [22].

2.2. τ m⁵U for Decoding UGA/UGG and AGA/AGG Codons. In the tunicate (ascidian) and vertebrate mitochondria, the anticodon first letter of tRNA^{Trp} is modified to 5-taurinomethyluridine (τ m⁵U) ([8, 23]). The same position of ascidian *Halocynthia roretzi* mt tRNA^{Gly} is also occupied by τ m⁵U [23]. That τ m⁵U recognizes only A- and G-ending codons (Figures 3(a) and 3(b)) were verified in two ways. Yasukawa et al. examined the translation activity of human tRNA^{Leu}_{UAA} possessing τ m⁵U (at that time it was an unknown modified uridine (U*) [24], and later it was elucidated to be τ m⁵U [8]) in a bovine mitochondrial *in vitro* translation system using synthetic mRNAs possessing 30

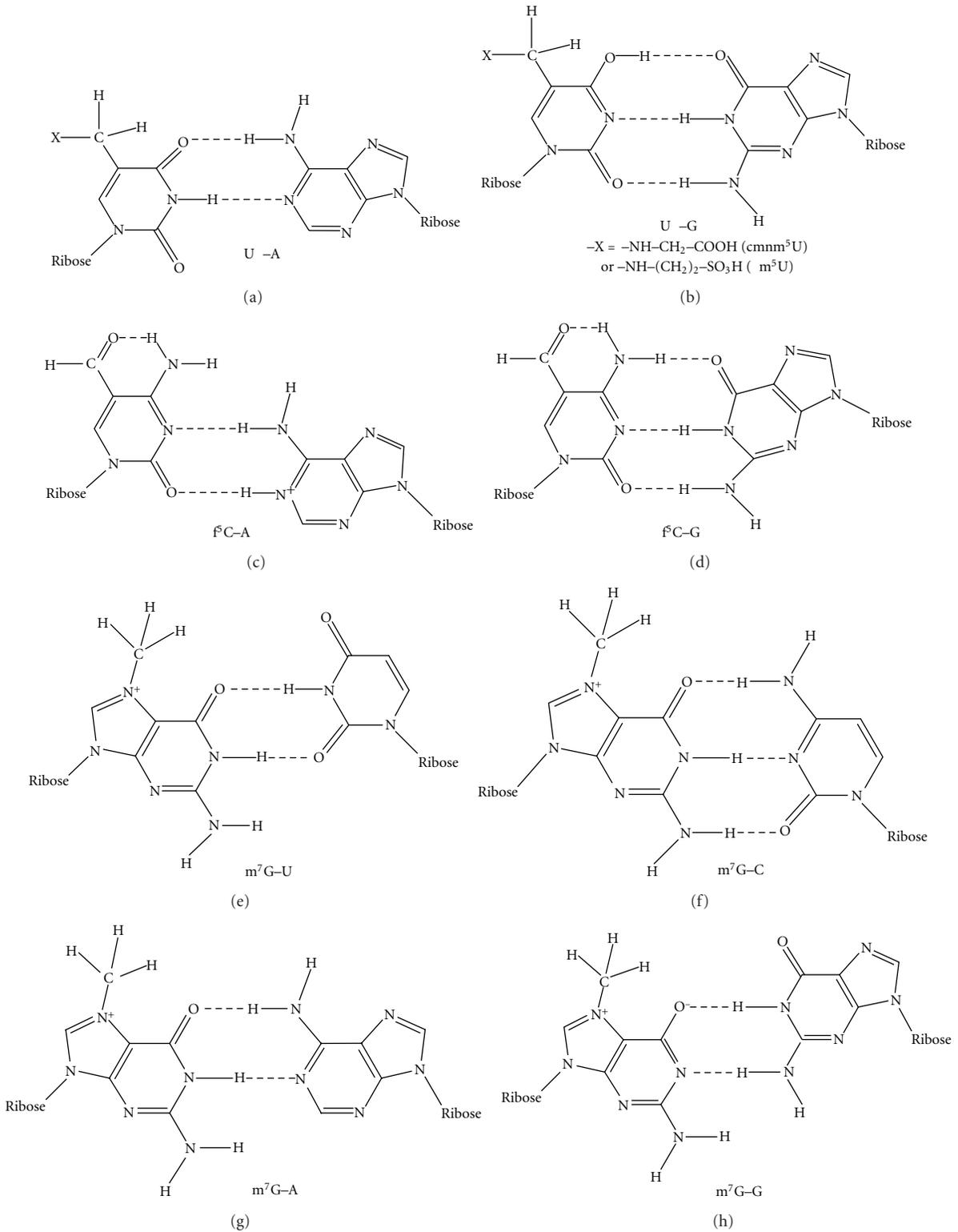


FIGURE 3: Possible scheme of base pairing between modified nucleosides at the anticodon wobble position and nucleosides at the codon third position in animal mitochondrial translation systems. These structures are only schematic drawing of base-pairing and do not show precise dimension of each nucleoside.

triplet repeats for the Leu codons UUA and UUG, as well as UUC as a negative control [24, 25]. They showed clearly that wild-type tRNA^{Leu}_{UAA} translated both poly(UUA)₃₀ and poly(UUG)₃₀ to poly(Leu) efficiently, but scarcely translated poly(UUU)₃₀ or poly(UUC)₃₀.

Kurata et al. measured the decoding activity of *E. coli* tRNA^{Leu}_{UAA} possessing τm^5U , $cmnm^5U$, or unmodified U (negative control) in the anticodon wobble position which was constructed by using molecular surgery technique [26] in *E. coli* S30 *in vitro* cell-free system with synthetic oligoribonucleotides including UUN codon as messengers [27]. They clearly demonstrated that tRNA^{Leu}_{UAA} possessing either τm^5U or $cmnm^5U$ in the anticodon wobble position could translate UUA/UUG-containing messengers efficiently but could translate neither UUU- nor UUC-containing messenger. The tRNA^{Leu}_{UAA} with unmodified U34 efficiently translated the UUA codon, but decoding of the UUG codon was approximately one third the activity of decoding UUA. However, this tRNA^{Leu}_{UAA} could not translate either UUU or UUC in consistent with the results obtained for mitochondrial translation [28]. Thus, either τm^5U - or $cmnm^5U$ -modification was proved to be essential for restricting the base-pairing to the purine-ending codons (Figures 3(a) and 3(b)).

The conformation of τm^5U has not been analyzed, but it can be speculated that it is very similar to that of $cmnm^5U$, because only their side chains at position 5 of the uridine base are different; glycine and taurine are linked to 5-methyluridine in $cmnm^5U$ and τm^5U , respectively (Figure 2). As described above, the modified uridine possessing the side chain at position 5 of uridine base which is linked through a methylene group takes very rigid conformation, which enables base pair with G as well as A in the codon third position.

Mitochondrial glycyl-tRNA synthetase (GlyRS) has scarcely been studied at the molecular level. It is suggested that a tunicate (*Ciona intestinalis*) genome encodes single GlyRS gene responsible for synthesis of both cytoplasmic and mt GlyRSs (Yokobori et al., unpublished results). Kondow et al. elucidated that a tunicate (*H. roretzi*) mt tRNA^{Gly} _{τm^5UCU} is glycylylated *in vivo* [29]. These results strongly suggest that tunicate mt GlyRS is possible to recognize tRNA^{Gly} _{τm^5UCU} possessing τm^5U34 .

2.3. f⁵C for Decoding AUA Codon. 5-formylcytidine (f⁵C) occurs at the anticodon wobble position of tRNA^{Met} of most invertebrate (fruit fly, *Drosophila melanogaster* [12], squid, *Loligo bleekeri* [30], and nematode, *Ascaris suum* [31]), and vertebrate (bovine, *Bos taurus* [9]) mitochondria, where AUA codon is read as Met instead of Ile [1, 5]. However, in echinoderm [32], some platyhelminth (such as planaria), cnidarian, placozoan, and poriferan mitochondria, AUA codon is read as Ile, like universal genetic code. In the former case, it is considered that f⁵C34 of tRNA^{Met} could restrict the base pair with A and G in the codon third letter (Figures 3(c) and 3(d)), but in the latter case, the wobble nucleoside of tRNA^{Ile} is G34, so that this tRNA would translate AUA codon as not Met, but Ile.

By constructing an *in vitro* translation system from bovine liver mitochondria, Takemoto et al. examined the decoding properties of the native mt tRNA^{Met} carrying f⁵C in the anticodon compared to a transcript that lacks the modification [14]. The native mt Met-tRNA^{Met} could recognize both AUG and AUA codons as Met, but the corresponding synthetic tRNA^{Met} lacking f⁵C (anticodon CAU) recognized only the AUG codon in both the codon-dependent ribosomal binding and *in vitro* translation assays. Furthermore, the *E. coli* elongator tRNA^{Met}_m with the anticodon ac⁴CAU (ac⁴C=4-acetylcytidine) and the bovine cytoplasmic initiator tRNA^{Met} (anticodon CAU) translated only the AUG codon for Met on mt ribosome. The codon recognition patterns of these tRNAs were the same on *E. coli* ribosomes. These results demonstrate that the f⁵C modification in mt tRNA^{Met} plays a crucial role in decoding the nonuniversal AUA codon as Met, and that the genetic code variation is compensated by a change in the tRNA anticodon, not by a change in the ribosome.

Conformation analysis of f⁵C by 500-MHz NMR showed that the nucleoside takes a very rigid C3'-endo-anti form [33]. This feature may be advantageous for the decoding properties of tRNA^{Met}, because a very rigid pyrimidine in the first position of the anticodon cannot form base pairs with U and C in the codon third position [34, 35], so that the tRNA^{Met} cannot decode the AUU and AUC Ile codons. In addition, it is to be anticipated that the stability conferred by a rigid ribose moiety in the first anticodon nucleotide will to some extent be propagated to the second and third anticodon residues. This would result in greater overall stability of the stacked anticodon bases, and thus of codon-anticodon pairings.

It is apparent that f⁵C can interact in the expected manner with G of the AUG Met codon (Figure 3(d)), reflecting the finding that the conformation of f⁵C is similar to that of cytidine [33]. In order to read the AUA Met codon, the intriguing possibility exists that f⁵C pairs with A in the AUA Met codon by protonation (Figure 3(c)). The protonation of A at N-1 in an A-C pair at pH values above the pK of the monomer has been demonstrated in oligoribonucleotide duplexes [36]. In addition, A was found to adopt a pK of 6.5 in the active site of a Pb-dependent ribozyme [37].

Human mt MetRS has been studied by Spemulli's group [38]. They found that the enzyme recognizes the anticodon region of tRNA^{Met} as an identity determinant and aminoacylates both human mt tRNA^{Met}_{CAU} transcript and bovine tRNA^{Met}_{f⁵CAU} with similar K_M (0.15~0.16 μ M) and k_{cat} (0.02 s⁻¹) values. As the nucleotide sequences of human and bovine mt tRNAs^{Met}_{CAU} are almost identical to each other with 5 nucleotide differences at positions 16, 27, 50, 56, and 60, and it has turned out that human mt tRNAs^{Met}_{CAU} also possesses f⁵C at the anticodon wobble position [39] these kinetic data clearly demonstrate that human mt MetRS recognizes the substrate tRNA^{Met}_{CAU} irrespective of the presence or absence of f⁵C34.

2.4. m⁷G and Unmodified G Which Decodes AGA or AGG Codon in Most Invertebrates. AGA and AGG codons are read as Ser in most metazoan mitochondria [32]. Matsuyama et al.

found that 7-methylguanosine (m^7G) is present at the wobble position of $tRNA^{Ser}_{GCU}$ in most invertebrate mitochondria [10]. Therefore, the anticodon m^7GCU of $tRNA^{Ser}_{GCU}$ is most likely responsible for reading all four AGN codons as Ser (Figures 3(e)–3(h)) [10]. On the other hand, the AGG codon is absent from some metazoan mitochondria, such as fruit fly *Drosophila melanogaster* mitochondria (Table 1), and the anticodon wobble position of $tRNA^{Ser}_{GCU}$ is unmodified G [12]. In this case, the unmodified GCU anticodon of $tRNA^{Ser}_{GCU}$ seems to read the three codons AGU, AGC, and AGA as Ser. Therefore, it is summarized that G34 of *Drosophila* $tRNA^{Ser}_{GCU}$ base pairs with only U, C, and A in the third letter of Ser codon, and in most invertebrate mt $tRNAs^{Ser}_{GCU}$, it is prerequisite for G34 to be modified to m^7G for recognizing G in the third letter of Ser codon.

The purine-purine base pairing is well discussed by Murphy and Ramakrishnan for I-A pair [43]. They reported the crystal structure of I-C and I-A base pairs in the context of the ribosomal decoding center, clearly showing that the I-A base pair is of an $I_{anti} - A_{anti}$ conformation, as predicted by Crick [44] although the distance between $C_1 - C_1$ of I and A residues is broader than the usual one. Owing to this observation, G-A and also m^7G -A base pairs in question would be possible to take a similar structure as that of I-A base pair (Figure 3(g)). Since m^7G can form a structure in which a proton is cleaved from HN_1 and O_6 becomes O^- (but G cannot form the structure), m^7G -G base pair can be formed in which m^7G moves to the minor groove side in the context of the ribosomal decoding center (Figure 3(h)) [45]. Such base pair stacks well on the neighboring second base pair of codon-anticodon pairing, so that the whole interaction would be stabilized [46]. This speculation must be confirmed experimentally, the simplest way of which would be to use an *in vitro* translation system.

The recognition mechanism of bovine mt SerRS toward $tRNA^{Ser}$ has been elucidated well by biochemical [47, 48] as well as X-ray crystallographic studies [49]. Both results clearly demonstrated that the anticodon region of $tRNA^{Ser}$ is not involved in the identity determinant for bovine mt SerRS. Thus, it can be concluded that the presence or absence of τm^5U in the anticodon wobble position has no influence on the recognition of SerRS toward $tRNA^{Ser}$.

2.5. Ψ in the Anticodon Second Position of $tRNA^{Asn}_{GUU}$ Responsible for Decoding AAA Codon. In echinoderm [32] and some platyhelminth mitochondria [50], not only the usual Asn codons AAU and AAC, but also the usual Lys codon AAA, are read as Asn by a single mt $tRNA^{Asn}$ with the anticodon GUU. Tomita et al. elucidated that starfish mt $tRNA^{Asn}$ possesses the anticodon G Ψ U, whose second position is modified to pseudouridine (Ψ) [42]. In contrast, mt $tRNA^{Lys}$, corresponding to another Lys codon, AAG, has the anticodon CUU. Mt $tRNAs$ possessing anticodons closely related to that of $tRNA^{Asn}$, but responsible for decoding only two codons each ($tRNA^{His}$, $tRNA^{Asp}$, and $tRNA^{Tyr}$) (see Figure 1), were found to possess unmodified U35 in all cases, suggesting the importance of Ψ 35 in $tRNA^{Asn}$ for decoding the AAU, AAC, and AAA codons. Experiments with an

E. coli in vitro translation system confirmed that $tRNA^{Asn}_{G\Psi U}$ has about two-fold higher translational efficiency than $tRNA^{Asn}_{GUU}$ [42], in which $tRNA^{Asn}_{G\Psi U}$ was constructed by chemical synthesis and ligation, and $tRNA^{Asn}_{GUU}$ was obtained by *in vitro* run-off transcription. It is exceptional that modification at the anticodon second nucleoside is involved in the codon-anticodon interaction efficiency.

There is a report that Ψ within the base-paired region at position 35 in a model for the codon-anticodon interaction in $tRNA^{Tyr}$ increases the T_m by several degrees [51]. This is also supporting evidence for the above-mentioned phenomenon.

3. Relationship between Modified Nucleosides and Genetic Code Variations in the Course of Animal Evolution

In considering the evolution of the genetic code, there are so far two main hypotheses, the “codon-capture” hypothesis based on directional mutation pressure proposed by Osawa and Jukes [52], and the “ambiguous intermediate” hypothesis proposed by Schultz and Yarus [53]. Codon capture hypothesis proposes temporary disappearance of a sense codon (or stop codon) from coding frames by conversion to another synonymous codon, followed by loss of the corresponding tRNA that translates the codon. (For a stop codon, the release factor (RF) must change simultaneously so as not to recognize the stop codon.) This change may be caused by directional mutation pressure acting on genome (AT or GC pressure), changes of RF, or genome economization, which will produce an “unassigned codon.” The codon reappears later by the conversion of another codon caused by directional mutation pressure and emergence of a tRNA (or RF) that translates (or recognizes) the reappeared codon with a different assignment. Thus, the codon is reassigned or captured. In the “ambiguous intermediate” hypothesis, it is presumed that reassignment of codons is facilitated by a translationally ambiguous intermediate where the transitional codon is read simultaneously as two different amino acids by two tRNAs, one cognate and the other near cognate. There has been little experimental evidence in support of the presence of these tRNAs that decode a single codon as two different amino acids simultaneously. Simulation study on codon reassignment [54] suggested that the pathway of codon reassignment in favor of either “codon capture” or “ambiguous codon” depends on the initial conditions such as genome size and number of codons in interest.

We adopt here the former hypothesis essentially to explain the genetic code variations of animal mitochondria, because the characteristics of the animal mitochondrial genome such as AT richness and genome economization would be suitable for the adoption of the codon-capture hypothesis. The hypothesis would nicely fit in explaining the UGA codon change in animal mitochondria. There has been so far no report of a release factor recognizing UGA codon [55, 56]. Such a release factor corresponding to eubacterial RF2 was probably lost in the animal mt system, so that

TABLE 1: Relationship between the genetic code variations and modified nucleosides in the anticodon of the corresponding tRNAs in the animal phyla.

(a) UGA codon.		
Animal phyla	UGA specificity	Anticodon
Vertebrata	Trp	$\tau m^5UCA^{(a)}$
Tunicata (Urochordata)	Trp	$\tau m^5UCA^{(b)}$
Cephalochordata	Trp	(TCA)
Echinodermata	Trp	(TCA)
Arthropoda	Trp	$U^*CA^{(c)}$
Most invertebrate phyla	Trp	$cmnm^5(s^2)UCA^{(d)}$
Platyhelminthes	Trp	(TCA)
Cnidaria, Placozoa, Porifera	Trp	(TCA)
(b) AUA codon.		
Animal phyla	AUA specificity	Anticodon
Vertebrata	Met	$f^5CAU^{(e)}$
Tunicata (Urochordata)	Met	$\tau m^5UAU^{(f)}$
Cephalochordata	Met	(CAT)
Echinodermata	Ile	(GAT)
Arthropoda	Met	$f^5CAU^{(g)}$
Most invertebrate phyla	Met	$f^5CAU^{(h,i)}$
Platyhelminthes		
Most (Echinostomida, Trematoda)	Met	(CAT)
Rhabditiophora, Planaria	Ile	(GAT)
Cnidaria	Ile	— ^(j)
Placozoa	Ile	(GAT)
Porifera	Ile	(CAT) ^(k)
(c) AAA codon.		
Animal phyla	AAA specificity	Anticodon
Vertebrata	Lys	(CTT)
Tunicata (Urochordata)	Lys	(TTT)
Cephalochordata	Lys	(TTT)
Echinodermata	Asn	$G\Psi U^{(l)}$
Arthropoda	Lys	$CUU^{(m)}/(TTT)$
Most invertebrate phyla	Lys	(CTT/TTT)
Platyhelminthes	Asn	(GTT)
Cnidaria	Lys	— ^(j)
Placozoa	Lys	(TTT)
Porifera	Lys	(TTT)
(d) AGA/AGG codons.		
Animal phyla	AGA/AGG specificity	Anticodon
Vertebrata	Term.	None
Tunicata (Urochordata)	Gly	$\tau m^5UCU^{(n)}$
Cephalochordata	Ser	(GCT)
Echinodermata	Ser	$m^7GCU^{(o)}$
Arthropoda		
Most Arthropods	Ser	(GCT/TCT)
<i>Drosophila melanogaster</i>	$Ser^{(p)}$	$GCU^{(q)}$
Most invertebrate phyla	Ser	$m^7GCU^{(r)}$
Platyhelminthes	Ser	(GCT/TCT)

(d) Continued.

Animal phyla	AGA/AGG specificity	Anticodon
Cnidaria	Arg	— ⁽ⁱ⁾
Placozoa	Arg	(TCT)
Porifera	Arg	(TCT)

In case that only the tRNA gene sequences are known, the anticodon sequences at DNA level are shown in parentheses. ^(a)*Homo sapiens* (Suzuki et al. [39]) ^(b)*Halocynthia roretzi* (Suzuki et al. [23]) ^(c)*Drosophila melanogaster* (Tomita et al. [12]). U* is an unknown modified uridine. ^(d)*Ascaris suum* (Sakurai et al. [15]) ^(e)*Bos taurus* (Moriya et al. [9]) ^(f)*Halocynthia roretzi* (Suzuki et al. [23]) ^(g)*Drosophila melanogaster* (Tomita et al. [12]) ^(h)*Loligo bleekeri* (Tomita et al. [30]) ⁽ⁱ⁾*Ascaris suum* (Watanabe et al. [31]) ^(j)Corresponding tRNA gene is not encoded in the mitochondrial genome. The tRNA decoding this codon is presumed to be imported from cytoplasm [40]. ^(k)Poriferan mt genomes encode three tRNA genes with anticodon CAT at DNA level. One of these tRNA is thought to be tRNA^{Ile}_{LAU} gene (L is the modified C (lysidine) found at the anticodon first position of most bacterial tRNA^{Ile} decoding AUA codon) [41]. ^(l)*Asterias amurensis* (Tomita et al. [42]) ^(m)*Drosophila melanogaster* (Tomita et al. [12]) ⁽ⁿ⁾*Halocynthia roretzi* (Kondow et al. [29], Suzuki et al. [23]) ^(o)*Asterias amurensis* (Matsuyama et al. [10]) ^(p)No AGG codon appears in the *D. melanogaster* mitochondrial genome. ^(q)*Drosophila melanogaster* (Tomita et al. [12]) ^(r)*Loligo bleekeri* (Tomita et al. [11]).

the UGA codon became unassigned. When the anticodon wobble position of tRNA^{Trp} changed from C to modified U (U*: cmnm⁵(s²)U or τm^5U), and a part of the Trp UGG codon in reading frames was changed to UGA by AT pressure, the UGA codon was captured by tRNA^{Trp}_{U*CA} and read as Trp (Figure 4(a)). Why lower animals use cmnm⁵(s²)U for U*, whereas higher animals use τm^5U instead of cmnm⁵(s²)U (Table 1(a)), and from which stage of animals conversion of cmnm⁵(s²)U to τm^5U occurs remain to be clarified.

To explain the evolutionary change of other nonuniversal codons in metazoan mitochondria, in addition to AT pressure, the genome economization effect (especially, that tRNA genes are restricted to 22~24 species, and that tRNAs are not imported from the cytoplasm in almost all metazoan mitochondria [7] except for a few cases in Cnidaria and so on) should be taken into consideration. At the same time, we should include the idea of competition between two different tRNAs toward a certain codon, to explain the role of modified nucleoside in the anticodon region of tRNA. Namely, any codon can be read by the corresponding tRNA, but if a competitor tRNA or a release factor arises that has stronger affinity toward the codon than does the original tRNA, the codon will then be read by the competitor.

The AUA codon is read as Ile in cnidarian, placozoan, poriferan, some platyhelminths, and echinoderm mitochondria, but it is read as Met in most metazoan mitochondria (Table 1(b)). ^fC in most metazoan mitochondria [8, 9, 12, 31] and τm^5U in tunicate mitochondria [23] found at the wobble position of tRNA^{Met} may be important for understanding the codon reassignment from AUA-Ile to AUA-Met. Since the interaction between tRNA^{Ile}_{GAU} and AUA on the ribosome might be more unstable than that between the tRNA^{Ile}_{GAU} and AUY codons, when tRNA^{Met} acquired the capacity to decode AUA by 5'-formylation of C or 5'-taurinometylation of U at the wobble position, tRNA^{Met}_{^fCAU} or tRNA^{Met} _{τm^5U AU} may have prevailed over tRNA^{Ile}_{GAU} in the interaction with the AUA codon. Thus, the reassignment of Ile to Met could have easily occurred (Figure 4(b)). In echinoderm mitochondria, ^fC or τm^5U modification was lost in tRNA^{Met}, so that AUA is read as Ile by tRNA^{Ile}_{GAU} because the competitor tRNA was lost. Why most metazoan tRNAs^{Met} possess ^fC modification, but only

tunicates tRNA^{Met} possess τm^5U modification is also to be solved.

The AAA codon is read as Lys in most metazoan mitochondria, but only in some platyhelminth and echinoderm mitochondria is read as Asn (Table 1(c)). In this case, pseudouridylation at the second position of the anticodon of tRNA^{Asn} [42] was critical for the codon reassignment (Figure 4(c)). The interaction between tRNA^{Asn}_{GΨU} and AAA may have prevailed over that between tRNA^{Lys}_{UUU} or tRNA^{Lys}_{CUU} (in various metazoan mitochondria such as those of echinoderms and fruit fly *Drosophila melanogaster*) and AAA. It is interesting observation that echinoderm and a part of platyhelminth mitochondria use AUA as a universal Ile codon and AAA as a nonuniversal Asn codon, while all the other animal mitochondria use AUA as a nonuniversal Met codon, and AAA as a universal Lys codon. This should be clarified in the relationship between genetic code change and animal evolution.

AGR codons are used as Ser codons in most invertebrate mitochondria, Gly codons in tunicate mitochondria and termination codons in vertebrate mitochondria (Table 1(d)). The scenario for this codon change would be as follows (Figure 4(d)). In the evolutionary process from Cnidaria, Placozoa, and Porifera to higher invertebrates, tRNA^{Arg}_{U*CU} to decode AGR codons as Arg disappeared from the mt genome, because of the reduction in genome size (in Cnidaria, most of tRNA genes are absent from the mt genome and most tRNAs are thought to be imported from the cytoplasm [40], but, in most higher animal mitochondria, import of tRNA has not been reported [7]). This has created a situation in which AGR codons cannot be translated. Since the AGR-Arg sites in the mitochondrial genomes of Protista such as *Trypanosoma* are mostly replaced with CGN-Arg codons and partially by a few other codons throughout metazoan mitochondria [57], AGR codons were converted mainly to CGN upon the deletion of tRNA^{Arg}_{U*CU}, so that AGR codons became unassigned (first step). Once the anticodon wobble position (G34) of tRNA^{Ser}_{GCU} was modified to m⁷G, all four AGN codons became ready to be decoded as Ser [10, 11]. AGR codons pairing with this tRNA^{Ser}_{m⁷GCU} then appeared in reading frames because of the mutation of AGY-Ser codons or other codons to AGR codons, and they were captured by Ser (second step). In ancestors of tunicates and

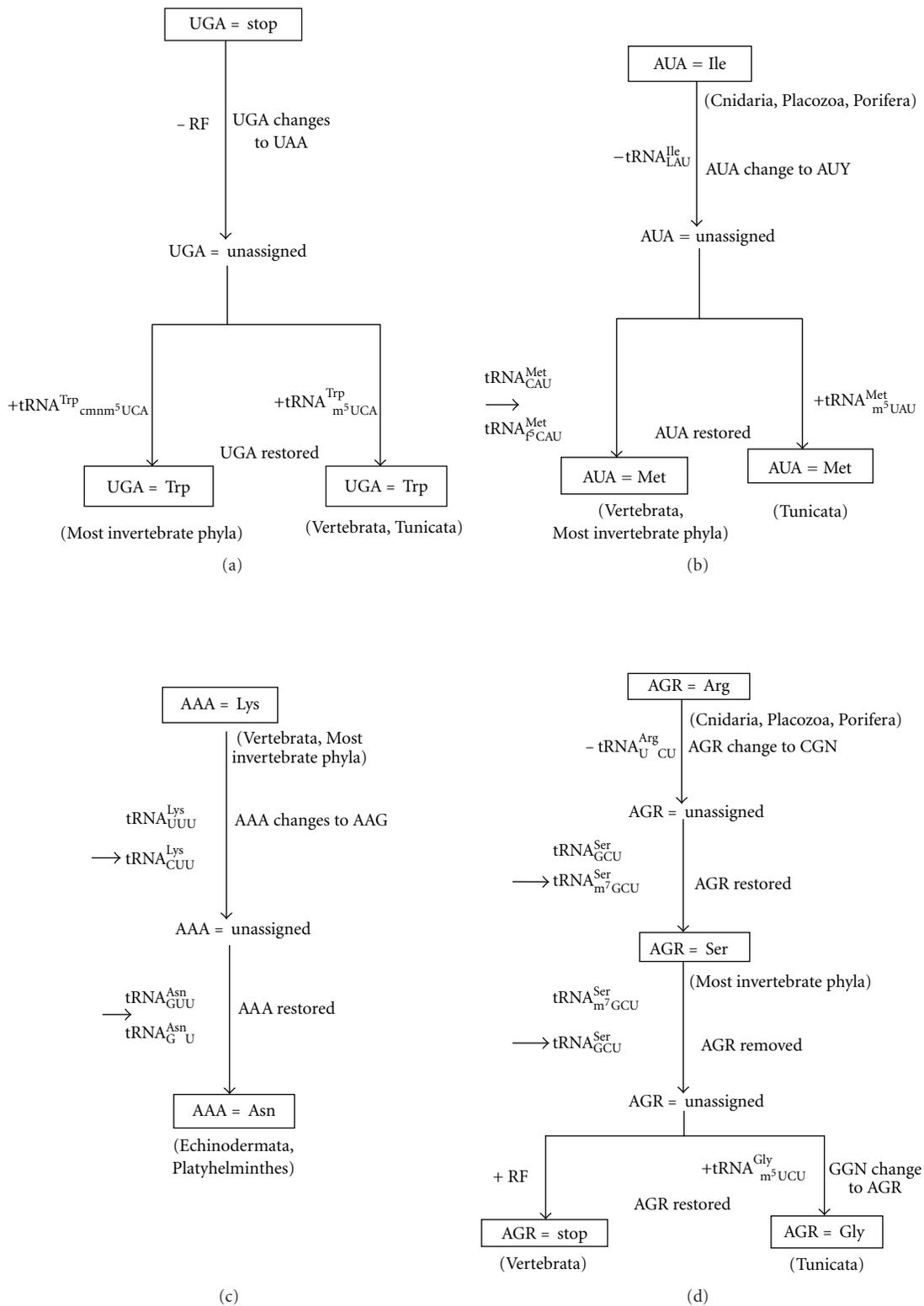


FIGURE 4: Schematic drawings showing the possible evolutionary changes of UGA (a), AUA (b), AAA (c), and AGR (d) codons in animal mitochondrial genomes. * In order to exclude too complicated situations, simple formulas were adopted for the process of the possible evolutionary changes of AGR codons in this figure. For details, see Text.

vertebrates, demethylation of m^7G of mt tRNA^{Ser} _{m^7GCU} may have occurred, so the resulting mt tRNA^{Ser} _{GCU} no longer reads the AGG codon. Strong selective constraints resulting from the lost translation of AGG caused the AGG codon to change mainly to AGY or other codons, so that the AGG codon became unassigned (third step). After tunicates were separated from vertebrate ancestors, a fourth event may have occurred; the mt tRNA^{Gly} _{UCC} gene was duplicated in mitochondrial DNA (mtDNA) (in fact, the ascidian mt genome has two tRNA^{Gly} genes [58]), and the anticodon of one species of tRNA^{Gly} _{UCC} was converted from TCC to TCT because of AT pressure, resulting in tRNA^{Gly} _{UCU} , which might have occurred in tunicate mt genomes. Then, the anticodon wobble position of tRNA^{Gly} _{UCU} must have been modified to τm^5U (tRNA^{Gly} _{τm^5UCU} ; ([8, 23]) so as to decode AGR codons. At the same time, AT pressure caused GGN codons to change to AGR codons. Since the interaction of tRNA^{Gly} _{τm^5UCU} with AGR is stronger than that of tRNA^{Ser} _{GCU} with AGR, AGR codons are captured by tRNA^{Gly} _{τm^5UCU} , resulting in the translation of AGR codons as Gly [59]. The low GC content of the cytochrome oxidase subunit I region in ascidian (tunicate) mitochondria in comparison to that of vertebrates is consistent with the speculation that AT pressure occurred on the tunicate mt genome. In the ancestors of vertebrate mitochondria, AGR codons may have appeared in the reading frames by the deletion of U from the UAG termination codon [57], concomitantly with a functional change in the vertebrate RF so as to recognize AGR codons (a possible candidate for an mt RF capable of recognizing AGR codons was reported [56]). The RF prevails over tRNA^{Ser} _{GCU} in decoding the AGA codon and, thus, changes AGR codons to termination codons (Figure 4(d)) [57]. Why m^7G - and τm^5U -modifications have become to be used in most metazoan and tunicate mitochondria, respectively, for decoding AGR codons are intriguing problems, which should be pursued in future.

4. Conclusion

The concept of the universal genetic code was forced to change since the discovery of nonuniversal codon in human mitochondria [4, 5], and it was known that not only mitochondria but also usual cellular systems (nuclear genomes) contain the nonuniversal genetic codes [60]. Nowadays, it is widely accepted that the genetic code is not universal but is changeable depending on the lineage of organisms [61]. Especially, it should be mentioned that animal mt genomes contain several variations in the genetic code, and the nonuniversal genetic codes are decoded by tRNAs which possess modified nucleosides in their anticodon first or second positions.

In animal mt genetic code, variations occur only in the 2 codon boxes where one family box is divided 2:2 (which may be easily understandable), and only purine-ending codon(s) are the target for codon reassignment (Figure 1). The questions may arise as to why no codon changes has occurred in two codon boxes of His(CAY)-Gln(CAR) and/or Asp(GAY)-Glu(GAR) and why pyrimidine-ending codon(s)

(NNY) do not become the target for codon reassignment. For the former question, such cases may be found in future [61] and for the latter question, it may be because it is difficult to provide any modified nucleoside in the tRNA anticodon wobble position which can decode purine-ending codons (NNR) and one or both of pyrimidine-ending codons (NNY).

The way of codon-anticodon interaction in the decoding process of the nonuniversal genetic code follows the base-pairing rule based on the physiochemical property of nucleosides. It is the future problems to be answered why various modified nucleosides such as $cmm^5(s^2)U$, τm^5U , f^5C , m^7G , and Ψ have been selected as the key molecules for decoding the nonuniversal genetic code and how these modified nucleosides have been distributed among various animal mitochondria.

The genetic code consists of interaction between codon and anticodon of tRNA, and inspections of genetic code changes in animal mitochondria in which considerable number of changes occur in various animal lineages will lead to the elucidation of the basic principle of origin and evolution of the genetic code. It is expected that these studies will shed light on a way to approach the origin of life.

Acknowledgments

The authors thank Professors Hiroyuki Hori and Kazuyuki Takai, Faculty of Engineering, Ehime University, Japan, for their kind instruction and discussion on the base-pairing property of m^7G . This work was supported by grants in aid for scientific research (C) from Japan Society for the Promotion of Science (JSPS).

References

- [1] K. Watanabe, "Unique features of animal mitochondrial translation systems: the non-universal genetic code, unusual features of the translational apparatus and their relevance to human mitochondrial diseases," *Proceedings of the Japan Academy B*, vol. 86, no. 1, pp. 11–39, 2010.
- [2] Y. Bessho, T. Ohama, and S. Osawa, "Planarian mitochondria II. The unique genetic code as deduced from cytochrome c oxidase subunit I gene sequence," *Journal of Molecular Evolution*, vol. 34, no. 4, pp. 331–335, 1992.
- [3] J. E. M. Jacob, B. Vanholme, T. Van Leeuwen, and G. Gheysen, "A unique genetic code change in the mitochondrial genome of the parasitic nematode *Radopholus similis*," *BMC Research Notes*, vol. 2, article 192, 2009.
- [4] B. G. Barrell, A. T. Bankier, and J. Drouin, "A different genetic code in human mitochondria," *Nature*, vol. 282, pp. 189–194, 1979.
- [5] S. Anderson, A. T. Bankier, B. G. Barrell et al., "Sequence and organization of the human mitochondrial genome," *Nature*, vol. 290, no. 5806, pp. 457–465, 1981.
- [6] K. Watanabe and S. Osawa, "tRNA sequences and variations in the genetic code," in *tRNA: Structure, Biosynthesis and Function*, D. Söll and U. L. RajBhandary, Eds., pp. 215–250, ASM Press, Washington, DC, USA, 1995.
- [7] B. A. Roe, J. F. H. Wong, E. Y. Chen, and P. A. Armstrong, "Sequence analysis of mammalian mitochondrial tRNAs," in

- Recombinant DNA: Proceedings of the 3rd Cleveland Symposium on Macromolecules*, A. G. Walton, Ed., pp. 167–176, Elsevier, Cleveland, Ohio, USA, June 1981.
- [8] T. Suzuki, T. Suzuki, T. Wada, K. Saigo, and K. Watanabe, "Taurine as a constituent of mitochondrial tRNAs: new insights into the functions of taurine and human mitochondrial diseases," *The EMBO Journal*, vol. 21, no. 23, pp. 6581–6589, 2002.
- [9] J. Moriya, T. Yokogawa, K. Wakita et al., "A novel modified nucleoside found at the first position of the anticodon of methionine tRNA from bovine liver mitochondria," *Biochemistry*, vol. 33, no. 8, pp. 2234–2239, 1994.
- [10] S. Matsuyama, T. Ueda, P. F. Crain, J. A. McCloskey, and K. Watanabe, "A novel wobble rule found in starfish mitochondria. Presence of 7-methylguanosine at the anticodon wobble position expands decoding capability of tRNA," *Journal of Biological Chemistry*, vol. 273, no. 6, pp. 3363–3368, 1998.
- [11] K. Tomita, T. Ueda, and K. Watanabe, "7-Methylguanosine at the anticodon wobble position of squid mitochondrial tRNA^{Ser}GCU: molecular basis for assignment of AGA/AGG codons as serine in invertebrate mitochondria," *Biophysica Acta*, vol. 1399, no. 1, pp. 78–82, 1998.
- [12] K. Tomita, T. Ueda, S. Ishiwa, P. F. Crain, J. A. McCloskey, and K. Watanabe, "Codon reading patterns in *Drosophila melanogaster* mitochondria based on their tRNA sequences: a unique wobble rule in animal mitochondria," *Nucleic Acids Research*, vol. 27, no. 21, pp. 4291–4297, 1999.
- [13] T. Hanada, T. Suzuki, T. Yokogawa, C. Takemoto-Hori, M. Sprinzl, and K. Watanabe, "Translation ability of mitochondrial tRNAs^{Ser} with unusual secondary structures in an *in vitro* translation system of bovine mitochondria," *Genes to Cells*, vol. 6, no. 12, pp. 1019–1030, 2001.
- [14] C. Takemoto, L. L. Spemulli, L. A. Benkowski, T. Ueda, T. Yokogawa, and K. Watanabe, "Unconventional decoding of the AUA codon as methionine by mitochondrial tRNA^{Met} with the anticodon ^fCAU as revealed with a mitochondrial *in vitro* translation system," *Nucleic Acids Research*, vol. 37, no. 5, pp. 1616–1627, 2009.
- [15] M. Sakurai, T. Ohtsuki, T. Suzuki, and K. Watanabe, "Unusual usage of wobble modifications in mitochondrial tRNAs of the nematode *Ascaris suum*," *FEBS Letters*, vol. 579, no. 13, pp. 2767–2772, 2005.
- [16] M. Sakurai, T. Ohtsuki, and K. Watanabe, "Modification at position 9 with 1-methyladenosine is crucial for structure and function of nematode mitochondrial tRNAs lacking the entire T-arm," *Nucleic Acids Research*, vol. 33, no. 5, pp. 1653–1661, 2005.
- [17] Y. Andachi, F. Yamao, A. Muto, and S. Osawa, "Codon recognition patterns as deduced from sequences of the complete set of transfer RNA species in *Mycoplasma capricolum*: resemblance to mitochondria," *Journal of Molecular Biology*, vol. 209, no. 1, pp. 37–54, 1989.
- [18] R. P. Martin, A.-P. Sibley, C. W. Gehrke et al., "5-[[Carboxymethyl]amino]methyl]uridine is found in the anticodon of yeast mitochondrial tRNAs recognizing two-codon families ending in a purine," *Biochemistry*, vol. 29, no. 4, pp. 956–959, 1990.
- [19] M. Sekine, L. S. Peshakova, T. Hata, S. Yokoyama, and T. Miyazawa, "Novel method for regioselective 2'-O-methylation and its application to the synthesis of 2'-O-methyl-5-[[carboxymethyl]- amino]methyl]uridine," *Journal of Organic Chemistry*, vol. 52, no. 22, pp. 5060–5061, 1987.
- [20] K. Takai, N. Horie, Z. Yamaizumi, S. Nishimura, T. Miyazawa, and S. Yokoyama, "Recognition of UUN codons by two leucine tRNA species from *Escherichia coli*," *FEBS Letters*, vol. 344, no. 1, pp. 31–34, 1994.
- [21] N. Horie, Z. Yamaizumi, Y. Kuchino et al., "Modified nucleosides in the first positions of the anticodons of tRNA^{Leu}₄ and tRNA^{Leu}₅ from *Escherichia coli*," *Biochemistry*, vol. 38, no. 1, pp. 207–217, 1999.
- [22] S. Yokoyama and S. Nishimura, "Modified nucleosides and codon recognition," in *tRNA: Structure, Biosynthesis and Function*, D. Söll and U. L. RajBhandary, Eds., pp. 207–223, ASM Press, Washington, DC, USA, 1995.
- [23] T. Suzuki, K. Miyauchi, T. Suzuki et al., "Taurine-containing uridine modifications in tRNA anticodons are required to decipher non-universal genetic codes in ascidian mitochondria," *The Journal of Biological Chemistry*, vol. 286, no. 41, pp. 35494–35498, 2011.
- [24] T. Yasukawa, T. Suzuki, S. Ohta, and K. Watanabe, "Wobble modification defect suppresses translational activity of tRNAs with MERRF and MELAS mutations," *Mitochondrion*, vol. 2, no. 1–2, pp. 129–141, 2002.
- [25] T. Yasukawa, T. Suzuki, N. Ishii, S. Ohta, and K. Watanabe, "Wobble modification defect in tRNA disturbs codon-anticodon interaction in a mitochondrial disease," *The EMBO Journal*, vol. 20, no. 17, pp. 4794–4802, 2001.
- [26] T. Suzuki, T. Ueda, and K. Watanabe, "The "polysemous" codon—a codon with multiple amino acid assignment caused by dual specificity of tRNA identity," *The EMBO Journal*, vol. 16, no. 5, pp. 1122–1134, 1997.
- [27] S. Kurata, A. Weixlbaumer, T. Ohtsuki et al., "Modified uridines with C5-methylene substituents at the first position of the tRNA anticodon stabilize U·G wobble pairing during decoding," *Journal of Biological Chemistry*, vol. 283, no. 27, pp. 18801–18811, 2008.
- [28] Y. Kirino, T. Yasukawa, S. Ohta et al., "Codon-specific translational defect caused by a wobble modification deficiency in mutant tRNA from a human mitochondrial disease," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 42, pp. 15070–15075, 2004.
- [29] A. Kondow, T. Suzuki, S. Yokobori, T. Ueda, and K. Watanabe, "An extra tRNA^{Gly}(U*CU) found in ascidian mitochondria responsible for decoding non-universal codons AGA/AGG as glycine," *Nucleic Acids Research*, vol. 27, no. 12, pp. 2554–2559, 1999.
- [30] K. Tomita, T. Ueda, and K. Watanabe, "5-formylcytidine (^fC) found at the wobble position of the anticodon of squid mitochondrial tRNA^{Met}CAU," *Nucleic Acids Symposium Series*, no. 37, pp. 197–198, 1997.
- [31] Y. Watanabe, H. Tsurui, T. Ueda et al., "Primary and higher order structures of nematode (*Ascaris suum*) mitochondrial tRNAs lacking either the T or D stem," *Journal of Biological Chemistry*, vol. 269, no. 36, pp. 22902–22906, 1994.
- [32] H. Himeno, H. Masaki, T. Kawai et al., "Unusual genetic codes and a novel gene structure for tRNA^{Ser}_{AGY} in starfish mitochondrial DNA," *Gene*, vol. 56, no. 2–3, pp. 219–230, 1987.
- [33] G. Kawai, T. Yokogawa, K. Nishikawa et al., "Conformational properties of a novel modified nucleoside, 5-formylcytidine, found at the first position of the anticodon of bovine mitochondrial tRNA^{Met}," *Nucleosides and Nucleotides*, vol. 13, no. 5, pp. 1189–1199, 1994.
- [34] S. Yokoyama, T. Watanabe, K. Murao et al., "Molecular mechanism of codon recognition by tRNA species with modified uridine in the first position of the anticodon," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 82, no. 15, pp. 4905–4909, 1985.

- [35] G. Kawai, T. Hashizume, M. Yasuda, T. Miyazawa, J. A. McCloskey, and S. Yokoyama, "Conformational rigidity of N⁴-acetyl-2'-O-methylcytidine found in tRNA of extremely thermophilic archaeobacteria (Archaea)," *Nucleosides and Nucleotides*, vol. 11, no. 2–4, pp. 759–771, 1992.
- [36] S. Wang, H. Gao, B. L. Gaffney, and R. A. Jones, "Nitrogen-15-labeled oligodeoxynucleotides. 3. Protonation of the adenine N1 in the A·C and A·G mispairs of the duplexes {d[CG(¹⁵N1)AGAATTC¹⁵CG]}₂ and {d[CGGGAATTC(¹⁵N1)ACG]}₂," *Journal of the American Chemical Society*, vol. 113, no. 14, pp. 5486–5489, 1991.
- [37] M. D. Canny, F. M. Jucker, E. Kellogg, A. Khvorova, S. D. Jayasena, and A. Pardi, "Fast cleavage kinetics of a natural hammerhead ribozyme," *Journal of the American Chemical Society*, vol. 126, no. 35, pp. 10848–10849, 2004.
- [38] A. C. Spencer, A. Heck, N. Takeuchi, K. Watanabe, and L. L. Spremulli, "Characterization of the human mitochondrial methionyl-tRNA synthetase," *Biochemistry*, vol. 43, no. 30, pp. 9743–9754, 2004.
- [39] T. Suzuki, A. Nagano, and T. Suzuki, "Human mitochondrial diseases caused by lack of taurine modification in mitochondrial tRNAs," *WIREs RNA*, vol. 2, no. 3, pp. 376–386, 2011.
- [40] C. T. Beagley, R. Okimoto, and D. R. Wolstenholme, "The mitochondrial genome of the sea anemone *Metridium senile* (Cnidaria): introns, a paucity of tRNA genes, and a near-standard genetic code," *Genetics*, vol. 148, no. 3, pp. 1091–1108, 1998.
- [41] D. V. Lavrov, L. Forget, M. Kelly, and B. F. Lang, "Mitochondrial genomes of two demosponges provide insights into an early stage of animal evolution," *Molecular Biology and Evolution*, vol. 22, no. 5, pp. 1231–1239, 2005.
- [42] K. Tomita, T. Ueda, and K. Watanabe, "The presence of pseudouridine in the anticodon alters the genetic code: a possible mechanism for assignment of the AAA lysine codon as asparagine in echinoderm mitochondria," *Nucleic Acids Research*, vol. 27, no. 7, pp. 1683–1689, 1999.
- [43] F. V. Murphy IV and V. Ramakrishnan, "Structure of a purine-purine wobble base pair in the decoding center of the ribosome," *Nature Structural & Molecular Biology*, vol. 11, no. 12, pp. 1251–1252, 2004.
- [44] F. H. C. Crick, "Codon-anticodon pairing: the wobble hypothesis," *Journal of Molecular Biology*, vol. 19, no. 2, pp. 548–555, 1966.
- [45] K. Takai, "Classification of the possible pairs between the first anticodon and the third codon positions based on a simple model assuming two geometries with which the pairing effectively potentiates the decoding complex," *Journal of Theoretical Biology*, vol. 242, pp. 564–580, 2006.
- [46] H. Mizuno and M. Sundaralingam, "Stacking of Crick Wobble pair and Watson-Crick pair: stability rules of G-U pairs at ends of helical stems in tRNAs and the relation to codon-anticodon Wobble interaction," *Nucleic Acids Research*, vol. 5, no. 11, pp. 4451–4461, 1978.
- [47] T. Yokogawa, N. Shimada, N. Takeuchi et al., "Characterization and tRNA recognition of mammalian mitochondrial seryl-tRNA synthetase," *Journal of Biological Chemistry*, vol. 275, no. 26, pp. 19913–19920, 2000.
- [48] N. Shimada, T. Suzuki, and K. Watanabe, "Dual mode recognition of two isoacceptor tRNAs by mammalian mitochondrial seryl-tRNA synthetase," *Journal of Biological Chemistry*, vol. 276, no. 50, pp. 46770–46778, 2001.
- [49] S. Chimnarok, M. G. Jeppesen, T. Suzuki, J. Nyborg, and K. Watanabe, "Dual-mode recognition of noncanonical tRNAs^{Ser} by seryl-tRNA synthetase in mammalian mitochondria," *The EMBO Journal*, vol. 24, no. 19, pp. 3369–3379, 2005.
- [50] M. J. Telford, E. A. Herniou, R. B. Russell, and D. T. J. Littlewood, "Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 21, pp. 11359–11364, 2000.
- [51] D. R. Davis, "Biophysical and conformational properties of modified nucleosides in RNA (nuclear magnetic resonance studies)," in *Modification and Editing of RNA*, H. Grosjean and R. Benne, Eds., pp. 103–112, ASM Press, Washington, DC, USA, 1998.
- [52] S. Osawa and T. H. Jukes, "Codon reassignment (codon capture) in evolution," *Journal of Molecular Evolution*, vol. 28, no. 4, pp. 271–278, 1989.
- [53] D. W. Schultz and M. Yarus, "Transfer RNA mutation and the malleability of the genetic code," *Journal of Molecular Biology*, vol. 235, no. 5, pp. 1377–1380, 1994.
- [54] S. Sengupta, X. Yang, and P. G. Higgs, "The mechanisms of codon reassignments in mitochondrial genetic codes," *Journal of Molecular Evolution*, vol. 64, no. 6, pp. 662–688, 2007.
- [55] C. C. Lee, K. M. Timms, C. N. A. Trotman, and W. P. Tate, "Isolation of a rat mitochondrial release factor. Accommodation of the changed genetic code for termination," *Journal of Biological Chemistry*, vol. 262, no. 8, pp. 3548–3552, 1987.
- [56] H. R. Soleimanpour-Lichaei, I. Kuhl, M. Gaisne et al., "mtRF1a is a human mitochondrial translation release factor decoding the major termination codons UAA and UAG," *Molecular Cell*, vol. 27, no. 5, pp. 745–757, 2007.
- [57] S. Osawa, T. Ohama, T. H. Jukes, and K. Watanabe, "Evolution of the mitochondrial genetic code I. Origin of AGR serine and stop codons in metazoan mitochondria," *Journal of Molecular Evolution*, vol. 29, no. 3, pp. 202–207, 1989.
- [58] S. Yokobori, T. Ueda, G. Feldmaier-Fuchs et al., "Complete DNA sequence of the mitochondrial genome of the ascidian *Halocynthia roretzi* (Chordata, Urochordata)," *Genetics*, vol. 153, no. 4, pp. 1851–1862, 1999.
- [59] S. Yokobori, T. Ueda, and K. Watanabe, "Codons AGA and AGG are read as glycine in ascidian mitochondria," *Journal of Molecular Evolution*, vol. 36, no. 1, pp. 1–8, 1993.
- [60] F. Yamao, A. Muto, Y. Kawaguchi et al., "UGA is read as tryptophan tRNAs in *Mycoplasma capricolum*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 82, no. 8, pp. 2306–2309, 1985.
- [61] S. Osawa, *Evolution of the Genetic Code*, Oxford University Press, New York, NY, USA, 1995.

Research Article

Control of Aptamer Function Using Radiofrequency Magnetic Field

Kenichi Taira,¹ Koichi Abe,¹ Takayuki Ishibasi,^{2,3} Katsuaki Sato,^{2,4} and Kazunori Ikebukuro¹

¹Department of Biotechnology, Graduate School of Engineering, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan

²Department of Applied Physics, Graduate School of Engineering, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan

³Department of Material Science and Technology, Nagaoka University of Technology 1603-1, Kamitomioka, Nagaoka, Niigata 940-2188, Japan

⁴PRESTO Program, Japan Science and Technology Agency, 5 Sambancho, Chiyodaku, Tokyo 102-0075, Japan

Correspondence should be addressed to Kazunori Ikebukuro, ikebu@cc.tuat.ac.jp

Received 22 May 2011; Accepted 28 June 2011

Academic Editor: Daisuke Miyoshi

Copyright © 2011 Kenichi Taira et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Remote control of aptamer function has allowed us to control protein function in space and time. Here, we propose a novel control system for aptamer function by radiofrequency magnetic field- (RFMF-) induced local heating of a gold nanoparticle conjugated with an aptamer. In this study, we used a 31-mer thrombin-binding aptamer (TBA), which can inhibit thrombin activity, as a model aptamer. We evaluated the RFMF control of the inhibitory activity of a gold nanoparticle-conjugated TBA. To evaluate the effect of RFMF on enzymatic activity, we utilized a complementary DNA strand that maintains the broken structure during the activity assay. We observed a decrease in the inhibitory activity of TBA after RFMF irradiation. It indicates that RFMF is capable of controlling the TBA structure. Because RFMF allows noninvasive control of aptamer function, this strategy is expected to be novel way of controlling aptamer drug activity.

1. Introduction

Aptamers, single-stranded DNA or RNA molecules, recognize various target molecules on forming a particular three-dimensional structure [1, 2]. Each aptamer has a different topology and three-dimensional structure that recognizes a target [3]. Although most proteins lose their function by denaturing their structure irreversibly, aptamers can form correct structures reversibly even after denaturation of their structure. Therefore, aptamer function can be controlled reversibly by controlling the structures.

Many researchers have reported the remote control of functional nucleic acids by UV light irradiation. Caged compounds, which can be controlled by light, have been utilized for the remote control of aptamers and siRNA [4, 5]. However, because only a small amount of UV light penetrates deep tissues and also damages DNA structure,

the application of UV light would be limited unless specific devices are used that generate two-photon absorption [6].

Hamad-Schfferil et al. have reported remote electronic control of the DNA structure by inductive coupling of a radiofrequency magnetic field to a metal nanocrystal covalently linked to DNA [7]. Gold nanoparticles were heated by irradiation of a radiofrequency magnetic field (RFMF), leading to denaturation of a DNA and leaves the surrounding molecules relatively unaffected.

In this study, we propose a novel system for controlling aptamer function by RFMF irradiation (Figure 1). Local heating of a gold nanoparticle by RFMF irradiation would destabilize the aptamer structure and result in a decrease in aptamer function. However, the aptamer can easily refold to its native structure after denaturation of its structure. To evaluate their function easily, we introduced a complementary DNA strand that would not affect aptamer

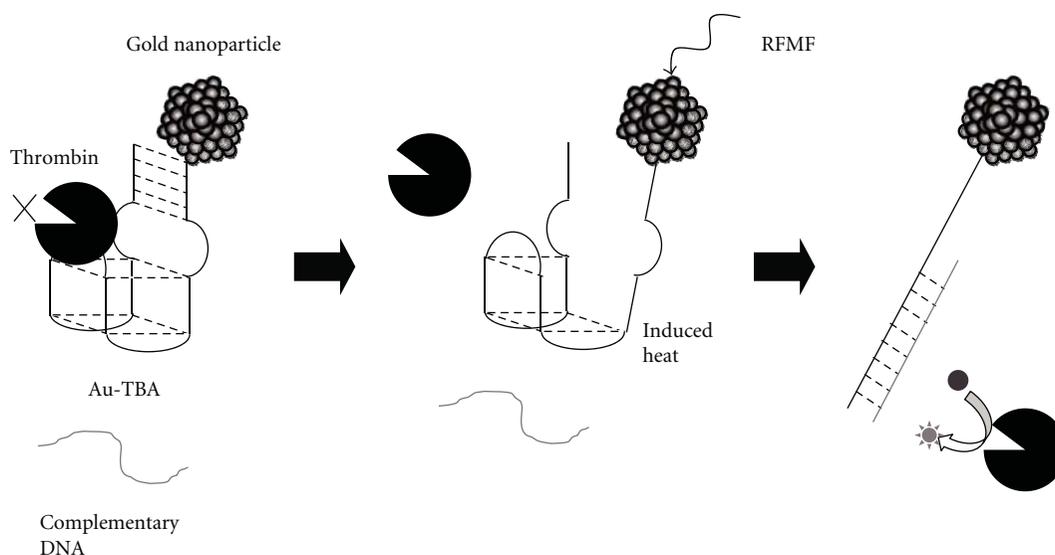


FIGURE 1: The scheme for the controlling of the function of gold nanoparticle-conjugated thrombin-inhibiting aptamer by RFMF.

function in the absence of RFMF irradiation. We expect that the complementary DNA strand hybridizes the aptamer on RFMF irradiation and remains in the hybridized state even after RFMF irradiation. Because the complementary DNA would help repress the refolding of the aptamer to its native structure, we can easily evaluate the aptamer function activity after RFMF irradiation.

In this study, we selected thrombin-binding aptamer (TBA) as a model aptamer. TBA forms a G-quadruplex structure to inhibit thrombin activity [8]. We selected a 31-mer TBA, because it has higher inhibitory activity than the more commonly used 15-mer TBA [9, 10]. In addition, because the 31-mer TBA forms the G-quadruplex structure attached to the stem region, we expect that conjugation of the gold nanoparticle has little effect on thrombin inhibitory activity of TBA. We prepared a 31-mer TBA conjugated with a gold nanoparticle (Au-TBA) and examined the control of its function by RFMF irradiation.

2. Materials and Methods

2.1. Linkage of Gold Nanocrystals to TBA. A monomaleimide-modified gold nanoparticle (diameter, 1.4 nm) was purchased from Nanoprobe (NY, USA). The DNA was synthesized by Operon Biotechnologies (Tokyo, Japan). The 5' end of the 31-mer TBA (5'-CACTGGTAGGTTGGTGTGGTTGGGGCCAGTG-3') was thiolated. A gold nanoparticle-conjugated 31-mer TBA was prepared according to the procedure published in the literature [7]. After gold nanoparticles were prepared with 10% isopropanol and diluted with water, a thiolated 31-mer thrombin-inhibiting aptamer (3 nmol) was coupled with an excess of monomaleimide-gold nanoparticles (30 nmol) at 4°C for 24 h. A gold nanoparticle-conjugated TBA was purified by Superdex 75

HiLoad 16/60 (GE Healthcare, UK Ltd, Buckinghamshire, England) in 5 mM NaH₂PO₄, 150 mM NaCl (pH 6.5).

Successful conjugation of gold nanoparticles with TBA (Au-TBA) was verified by native polyacrylamide gel (11%) electrophoresis (PAGE) in TBE buffer (89 mM Tris, 89 mM H₃BO₃, 2 mM EDTA) at room temperature. The band size of Au-TBA was compared with the fluorescein-modified 31-mer TBA. Fluorescein-modified 31-mer TBA was detected by a Typhoon 8600 (GE Healthcare, UK Ltd, Buckinghamshire, England) and Au-TBA was visualized by LI Silver staining (Nanoprobe, NY, USA) for 5 min to confirm Au-TBA purification.

2.2. Au-TBA Characterization. In TBS buffer (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 5 mM KCl), Au-TBA (final concentration, 100 nM) and 13.5-nM thrombin (Wako Chemicals, Osaka, Japan) were mixed at room temperature for 5 min. Then, 2 mg/mL fibrinogen (Wako Chemicals, Osaka, Japan) was mixed, and the clotting time was measured using Amelung KC-4A Micro Coagulation Analyzer (RWTUV, Essen, Germany) at 37°C.

2.3. Evaluation of Complementary DNA to Control Au-TBA Inhibitory Activity. Each complementary DNA (Table 1) (100 nM) was mixed with Au-TBA (100 nM) and thrombin (13.5 nM). Mixtures were incubated at room temperature for 5 min and then mixed with fibrinogen (2 mg/mL). The clotting time was measured using Amelung KC-4A Micro Coagulation Analyzer at 37°C.

2.4. Control of the Inhibitory Activity of Au-TBA by RFMF. Au-TBA (100 nM) and 20-mer complementary DNA (100 nM) and thrombin (13.5 nM) were incubated for 30 s at room temperature in 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 5 mM KCl and were irradiated by RFMF (1 GHz,

TABLE 1

Name	Sequence	Tm value (°C)*
TBA	5'-CACTGGTAGGTTGGTGTGGTTGGGGCCAGTG-3'	64
14-mer	5'-CACTGGCCCCAACC-3'	52.2
15-mer	5'-CACTGGCCCCAACCA-3'	54.2
16-mer	5'-CACTGGCCCCAACCAC-3'	56.4
17-mer	5'-CACTGGCCCCAACCACA-3'	58
20-mer	5'-CACTGGCCCCAACCACACCA-3'	63.3
23-mer	5'-CACTGGCCCCAACCACACCAACC-3'	66.7

*Tm (melting temperature) values of complementary DNA were predicted using mfold program. Tm value of TBA was measured using UV melting assay.

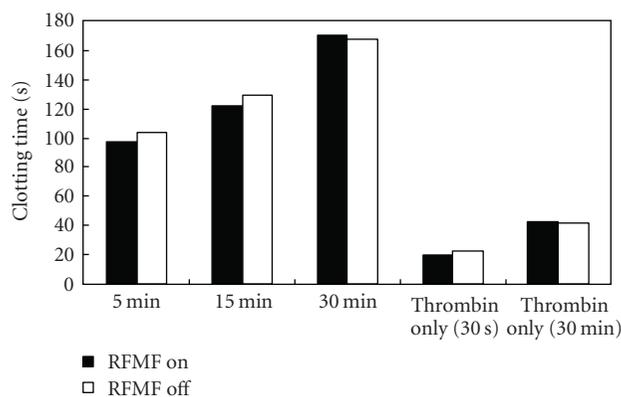


FIGURE 2: Evaluation of the effect of RFMF on thrombin activity. Mixtures of Au-TBA and thrombin were irradiated for 5 min, 15 min, or 30 min. As control, the mixtures were incubated for the same time at room temperature. Fibrinogen was added, and then, the clotting time was measured. White bars show clotting times without RFMF irradiation. Black bars show clotting times with RFMF irradiation.

1.5 W) for 5 min by using an HP 8648C RF signal generator (Hewlett-Packard Company, Calif, USA). The diameter of a coil was 1 cm and its height was 5 cm with 35 turns. To optimize RFMF power, we connected a tuning horn to an amplifier for matching. Then, fibrinogen (2 mg/mL) was mixed, and the clotting time was measured using Amelung KC-4A Micro Coagulation Analyzer at 37°C.

2.5. Confirmation of Complementary DNA and Au-TBA Hybridization. Au-TBA (100 nM), 20-mer complementary DNA (100 nM), and thrombin (13.5 nM) were incubated for 30 s at room temperature in 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 5 mM KCl, and some of these samples were irradiated by RFMF (1 GHz, 1.5 W) for 5 min by using the HP 8648C RF signal generator. After the RFMF irradiation, we carried out native polyacrylamide gel electrophoresis (PAGE) using 11% polyacrylamide gel in TBE buffer (89 mM Tris, 89 mM H₃BO₃, 2 mM EDTA) at room temperature. The fluorescence intensities of complementary DNA were analyzed by a Typhoon 8600.

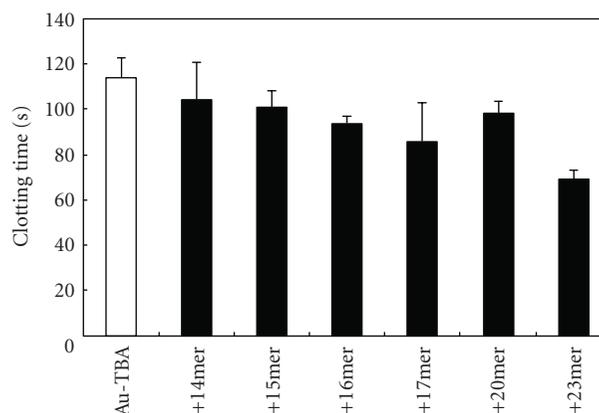


FIGURE 3: Evaluation of the effect of complementary DNA on TBA inhibitory activity. 14–17-mer, 20-mer, and 23-mer complementary DNA (100 nM) were incubated with Au-TBA (100 nM). The white bar shows the clotting time of thrombin and Au-TBA. The gray bar shows the clotting time of thrombin, Au-TBA, and each complementary DNA strand. Each clotting time represents the average of the clotting times in two experiments, and error bars are given.

3. Results and Discussion

First, we designed Au-TBA that has RFMF-controllable inhibitory activity. We selected a 31-mer TBA as a model aptamer. The 31-mer TBA forms a G-quadruplex-type structure that is 15-mer TBA attached to 3 base loops and a 5-bp stem. The internal loop region of the G-quadruplex recognizes the thrombin [11]. The stem region would stabilize the G-quadruplex structure and thus increase inhibitory activity. We then selected the 5' terminal of the aptamer for conjugation with gold nanoparticles and prepared Au-TBA using a monomaleimide gold nanoparticle and 5' thiol-modified TBA. We expected that the gold nanoparticle would not have an effect on the TBA inhibitory activity; further, heating of the gold nanoparticles by RFMF irradiation destabilizes the stem structure of the aptamer, resulting in a decrease in its inhibitory activity. In fact, the prepared Au-TBA shows inhibitory activity similar to that shown by unmodified TBA. Conjugation of the 5' terminal of TBA with gold nanoparticles did not significantly affect TBA function (data not shown).

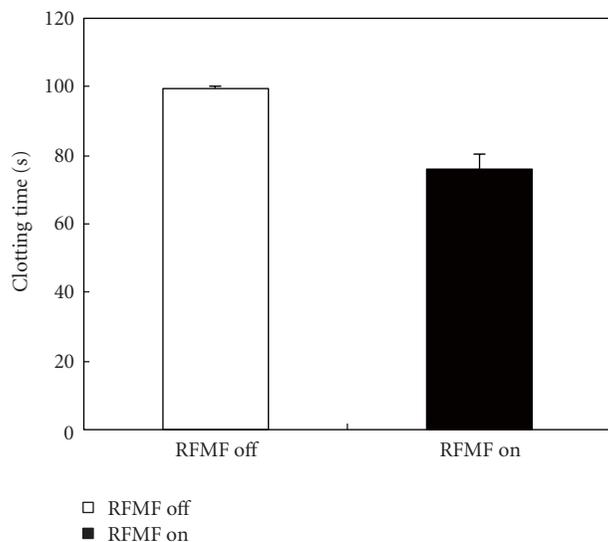


FIGURE 4: The control of the function of Au-aptamer by RFMF. The white bar termed “RFMF off” shows the clotting time of the mixture after 5 min of incubation at room temperature. The black bar termed “RFMF on” shows the clotting time of the mixture of thrombin, fibrinogen, TBA, and 20-mer complementary DNA after 5 min of irradiation using RFMF. Each clotting time represents the average of the clotting times in two experiments, and error bars are given.

Because the heat generated from nanoparticles by RFMF irradiation depends on nanoparticle size and magnetic field frequency, larger gold nanoparticles possibly affect not only the structure of the aptamer, but also thrombin structure. To control aptamer function, we have to select a suitable nanoparticle size and a magnetic field frequency that do not produce an unexpected effect on enzyme activity. Then, we selected gold nanoparticles with a diameter of 1.4 nm, which is the same as that of nanoparticles used in the study of Hamad-Schfferil et al. [7]. They mentioned that the effect of heat induced by this nanoparticle, because 1 GHz RFMF irradiation was limited to proximity. To evaluate the effect of RFMF irradiation on thrombin activity, RFMF was used to irradiate a mixture of Au-TBA and thrombin, and thrombin activity was measured after RFMF irradiation. DNA can easily refold to its correct structure, but it is difficult for proteins to renature correctly. A similar level of inhibition was observed regardless of whether the mixture was irradiated with RFMF (Figure 2). This indicates that heating induced by RFMF irradiation did not affect thrombin activity.

However, it is difficult to evaluate thrombin activity upon RFMF irradiation, because a complicated instrument is required for the measurement of thrombin activity with RFMF irradiation. Then, we used a complementary DNA strand of TBA. We expected RFMF would enhance hybridization between Au-TBA and a complementary DNA strand. The complementary DNA strand would repress refolding of the aptamer to its functional structure after RFMF irradiation, after which one can easily evaluate Au-TBA inhibitory activity. We selected an optimal complementary

DNA strand that insignificantly affects the Au-TBA structure without RFMF irradiation.

First, we investigated the optimal length of the complementary DNA strand for the evaluation of this system. Because the complementary DNA strand should have a T_m value lower than that of TBA (64°C) [12], we designed 14–17-mer, 20-mer, and 23-mer DNA strands complementary to the 3' terminal region of TBA; the 14–20-mer was designed such that its T_m value was lower than that predicted for TBA (Table 1) [13]. When we incubated a 14–17-mer or 20-mer complementary DNA with Au-TBA, they did not significantly affect the inhibitory activity of TBA without RFMF irradiation (Figure 3). On the other hand, the 23-mer complementary DNA decreased the inhibitory activity of TBA. We then selected a 20-mer complementary DNA strand to evaluate the effect of RFMF on the inhibitory activity of TBA.

We evaluated the controlling of TBA inhibitory activity by RFMF irradiation. We mixed thrombin, Au-TBA, and 100-nM 20-mer complementary DNA and incubated these at room temperature for 30 s; this mixture was subsequently subjected to RFMF irradiation for 5 min (Figure 4). Then, we added fibrinogen and measured the clotting time. The thrombin solution containing Au-TBA and 20-mer complementary DNA showed a shorter clotting time after RFMF irradiation than after incubation for 5 min at room temperature without RFMF irradiation (Figure 4). We observed a decrease of about 20% in the inhibitory activity of TBA by RFMF irradiation.

In order to confirm the successful hybridization of complementary DNA and Au-TBA, we performed native PAGE using fluorescein-modified complementary DNA (FcDNA). Because the gold nanoparticles quench the fluorescence of some fluorescent molecules and the TBA was labeled with a gold nanoparticle, it was difficult to obtain a distinct fluorescence band indicating the formation of complex of Au-TBA and 20-mer complementary DNA. Therefore, we quantified the hybridization reaction by comparing the fluorescence intensity of free FcDNA. We observed a decrease in the fluorescence intensity when Au-TBA was mixed with FcDNA (Figure 5(a), lane 2) rather than when FcDNA was used alone (Figure 5(a), lane 1) after RFMF irradiation. Although this decrease was also observed without RFMF (Figure 5(a), lane 3), we observed a weaker fluorescence intensity for free FcDNA subjected to RFMF irradiation than FcDNA that was not subjected to the same. The ratio of decrease in fluorescence intensity was approximately 20% (Figure 5(b)), and it shows good agreement with inhibitory activity results. These results indicate that RFMF irradiation destabilizes the Au-TBA structure and results in acceleration of hybridization between a complementary DNA strand and Au-TBA. We can say that we were able to control the inhibitory activity of TBA with complementary DNA and RFMF irradiation.

RFMF irradiation might induce heating of Au nanoparticle resulting in destabilization of Au-TBA structure. During RFMF irradiation, hybridization of complementary DNA with Au-TBA might compete with refolding to native structure of Au-TBA. Since RFMF irradiation would be limited

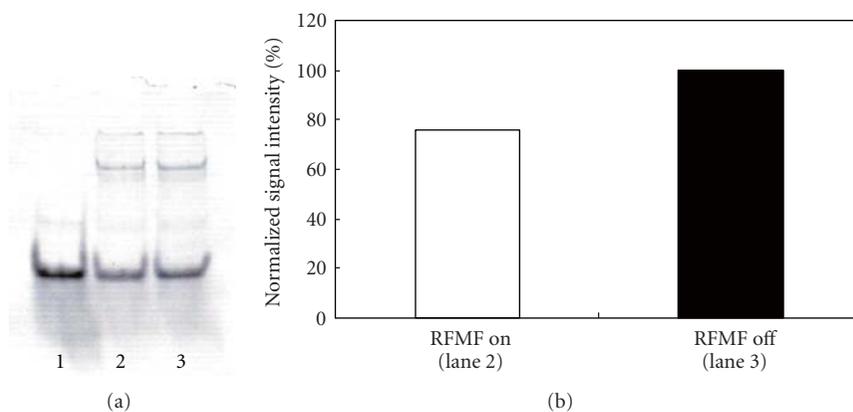


FIGURE 5: The confirmation of hybridization between FcDNA and Au-TBA by native PAGE. (a) Lane 1 shows a fluorescence image of 20-mer FcDNA. Lane 2 shows a fluorescence image of FcDNA with thrombin and Au-TBA after RFMF irradiation. Lane 3 shows a fluorescence image of FcDNA with thrombin and Au-TBA after incubation for 5 min at room temperature. (b) Signal intensity of FcDNA of Lane 2 and 3 normalized to Lane 2.

to local area around Au nanoparticle, induced heating of Au nanoparticle would not have so much effect on the solution temperature. Therefore, free Au-TBA might refold to native structure, whereas duplex might keep their structures after RFMF irradiation. Additionally, complementary DNA strand concentration would affect hybridization efficiency. Therefore, the control of structure of Au-TBA would be improved by adjusting of RFMF irradiation time and complementary DNA strand concentration.

In order to apply this system in therapeutic medicine, we should be able to control aptamer function by RFMF without the need for a complementary DNA strand. Because RFMF irradiation would destabilize the aptamer structure, the function of the aptamer is possibly repressed by RFMF irradiation in the absence of a complementary DNA strand. It may not be easy to determine the exact effect of RFMF irradiation on the aptamer function without a complementary DNA strand, but we can control it by adjusting the RFMF frequency and size of the gold nanoparticle, because the extent of instability of the structure of the gold nanoparticle-conjugated aptamer would depend on the energy generated from the gold nanoparticle after RFMF irradiation.

Control of aptamer function by RFMF irradiation has the following 3 advantages over that by UV light irradiation: in vivo penetration, no DNA damage, and limited effect on conjugated DNA. Although there are many reports describing the remote control of aptamers by UV light, deep tissue penetration is difficult and limits their application in vitro or requires two-photon absorption, which can be achieved only by a specific device. On the other hand, because RFMF can penetrate deep tissues, its application is not limited to an in vitro setting and can also be used in vivo. In addition, because the effect of heating induced by gold nanoparticles irradiated by RFMF is limited to the structure of the aptamer conjugated with gold nanoparticles, the heating of gold nanoparticles breaks the tertiary structure of the aptamer but does not modify its chemical structure. In addition, the heating of an Au nanoparticle can be limited to a local region by adjusting the frequency of RFMF and size

of the gold nanoparticle [7]. Therefore, RFMF irradiation prevents the unexpected effect caused by DNA damage and excessive heating of gold nanoparticles.

4. Conclusions

When therapeutic medicine produces an unexpected effect, it becomes necessary to control therapeutic medicine. Although antidotes based on complementary DNA have been developed for the use of therapeutic aptamers [14], remote control in therapeutic medicine is an attractive strategy. RFMF is widely used for magnetic resonance imaging and photothermal therapy [15, 16]. Therefore, heating induced by gold nanoparticles irradiated by RFMF is an attractive option for controlling aptamer function remotely in vivo. In this study, we demonstrated the control of thrombin inhibitory activity of TBA by RFMF irradiation and the use of using complementary DNA. Although we should develop a control system that does not require complementary DNA, we expect that the controlling of aptamer function by RFMF would be a useful tool in the future.

Authors' Contributions

Both K. Taira and K. Abe equally contributed to this paper.

Acknowledgment

This work is partly supported by the Grant-in-Aid for the 21st Century COE "Future Nano-Materials" from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan.

References

- [1] A. D. Ellington and J. W. Szostak, "In vitro selection of RNA molecules that bind specific ligands," *Nature*, vol. 346, no. 6287, pp. 818–822, 1990.

- [2] C. Tuerk and L. Gold, "Systemic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase," *Science*, vol. 249, no. 4968, pp. 505–510, 1990.
- [3] T. Hermann and D. J. Patel, "Adaptive recognition by nucleic acid aptamers," *Science*, vol. 287, no. 5454, pp. 820–825, 2000.
- [4] V. Mikat and A. Heckel, "Light-dependent RNA interference with nucleobase-caged siRNAs," *RNA*, vol. 13, no. 12, pp. 2341–2347, 2007.
- [5] M. C.R. Buff, F. Schäfer, B. Wulffen et al., "Dependence of aptamer activity on opposed terminal extensions: improvement of light-regulation efficiency," *Nucleic Acids Research*, vol. 38, no. 6, pp. 2111–2118, 2009.
- [6] K. König, "Multiphoton microscopy in life sciences," *Journal of Microscopy*, vol. 200, no. 2, pp. 83–104, 2000.
- [7] K. Hamad-Schifferli, J. J. Schwartz, A. T. Santos, S. Zhang, and J. M. Jacobson, "Remote electronic control of DNA hybridization through inductive coupling to an attached metal nanocrystal antenna," *Nature*, vol. 415, no. 6868, pp. 152–155, 2002.
- [8] L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas, and J. J. Toole, "Selection of single-stranded DNA molecules that bind and inhibit human thrombin," *Nature*, vol. 355, no. 6360, pp. 564–566, 1992.
- [9] K. Ikebukuro, Y. Okumura, K. Sumikura, and I. Karube, "A novel method of screening thrombin-inhibiting DNA aptamers using an evolution-mimicking algorithm," *Nucleic Acids Research*, vol. 33, no. 12, p. e108, 2005.
- [10] A. B. Dobrovolsky, E. V. Titaeva, S. G. Khaspekova, V. A. Spiridonova, A. M. Kopylov, and A. V. Mazurov, "Inhibition of thrombin activity with DNA-aptamers," *Bulletin of Experimental Biology and Medicine*, vol. 148, no. 1, pp. 33–36, 2009.
- [11] K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler, and A. Tulinsky, "The structure of α -thrombin inhibited by a 15-mer single-stranded DNA aptamer," *Journal of Biological Chemistry*, vol. 268, no. 24, pp. 17651–17654, 1993.
- [12] K. Abe, D. Ogasawara, W. Yoshida, K. Sode, and K. Ikebukuro, "Aptameric sensors based on structural change for diagnosis," *Faraday Discussions*, vol. 149, pp. 93–105, 2011.
- [13] N. R. Markham and M. Zuker, "DINAMelt web server for nucleic acid melting prediction," *Nucleic Acids Research*, vol. 33, no. 2, pp. W577–W581, 2005.
- [14] C. P. Rusconi, J. D. Roberts, G. A. Pitoc et al., "Antidote-mediated control of an anticoagulant aptamer in vivo," *Nature Biotechnology*, vol. 22, no. 11, pp. 1423–1428, 2004.
- [15] A. Ito, Y. Kuga, H. Honda et al., "Magnetite nanoparticle-loaded anti-HER2 immunoliposomes for combination of antibody therapy with hyperthermia," *Cancer Letters*, vol. 212, no. 2, pp. 167–175, 2004.
- [16] J. Cardinal, J. R. Klune, E. Chory et al., "Noninvasive radiofrequency ablation of cancer targeted by gold nanoparticles," *Surgery*, vol. 144, no. 2, pp. 125–132, 2008.

Review Article

Cell-Specific Aptamers as Emerging Therapeutics

Cindy Meyer, Ulrich Hahn, and Andrea Rentmeister

Chemistry Department, MIN Faculty, Institute for Biochemistry and Molecular Biology, Hamburg University, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany

Correspondence should be addressed to Andrea Rentmeister, rentmeister@chemie.uni-hamburg.de

Received 16 May 2011; Accepted 23 June 2011

Academic Editor: Souvik Maiti

Copyright © 2011 Cindy Meyer et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aptamers are short nucleic acids that bind to defined targets with high affinity and specificity. The first aptamers have been selected about two decades ago by an *in vitro* process named SELEX (systematic evolution of ligands by exponential enrichment). Since then, numerous aptamers with specificities for a variety of targets from small molecules to proteins or even whole cells have been selected. Their applications range from biosensing and diagnostics to therapy and target-oriented drug delivery. More recently, selections using complex targets such as live cells have become feasible. This paper summarizes progress in cell-SELEX techniques and highlights recent developments, particularly in the field of medically relevant aptamers with a focus on therapeutic and drug-delivery applications.

1. Introduction

Aptamers are short nucleic acids (typically 12–80 nucleotides long) capable of specific and tight binding to their target molecules. The term aptamer is derived from the Latin word *aptus* (fitting) and the Greek word *meros* (part). Aptamers are selected by a process called SELEX (systematic evolution of ligands by exponential enrichment), which was established independently by Ellington and Szostak [1], Tuerk and Gold [2], and Robertson and Joyce [3] in 1990. A typical SELEX experiment starts with a library of up to 10^{15} random oligonucleotides, which can be DNA, RNA, or modified RNA (e.g., 2'-OMe or 2'-F). Some members of this enormous library are anticipated to bind a desired target. The key step of the SELEX procedure is to efficiently separate those few from the nonbinding species. Selected nucleic acids are then amplified and used for further selection rounds. A successful SELEX experiment will usually result in a collection of aptamers, which can subsequently be cloned and tested individually for their binding properties.

The possible aptamer targets show a great diversity ranging from small molecules, like organic dyes [4], amino acids [5] or antibiotics [6], peptides [7], proteins [8], and viruses [9] to whole cells [10]. The dissociation constants (K_d values) of aptamer-target complexes are comparable to those of antibodies and can reach the picomolar range. In addition,

aptamers exhibit the following interesting features, which set them apart from antibodies: they are selected entirely *in vitro*, their synthesis has been automated, and they can easily be chemically modified [11]. Furthermore, they can be stored and shipped without problems, because the stability of DNA aptamers, in particular, is almost infinite. Importantly, they are not immunogenic. Due to their binding properties, aptamers have become useful tools for diagnostic and therapeutic applications [12, 13].

The first representative of an aptamer drug is Macugen, a highly modified RNA aptamer for treatment of age-related macular degeneration (AMD), which has been released in 2004 [14].

In this paper, we are focusing on recent developments in the selection of aptamers for complex targets and their application in diagnostics and therapy. We are emphasizing the potential of cell-based aptamer selections (cell-SELEX) as a means to enrich binders to live cells and highlight the potential of recent technical advances.

2. Aptamers for Intra- and Extracellular Targets and Their Application

Alternatively: for each class of targets, aptamers show outstanding characteristics. For small molecules, aptamers were

found that discriminate between closely related targets, such as caffeine and theophylline [15], differing by a single methyl group, or adenosine triphosphate and adenosine monophosphate (differing by their 5'-phosphorylation state) [16]. These aptamers have been successfully used as diagnostic tools and have been recently reviewed [17].

Nucleic acid-binding proteins, like Tat [18] and Rev [19] of HIV-1, were found to be excellent protein targets yielding high-affinity aptamers. However, numerous examples proved that nucleic acid binding is not a required target property for the selection of tightly and specifically binding aptamers. Several examples will be mentioned in this paper. There are many ways of classifying aptamer targets into different groups. Here, we want to distinguish between intracellular and extracellular targets.

Intracellular targets are difficult to address in living cells, because cytosolic delivery of aptamers, so-called intramers, in sufficient quantities is not facile. The problem becomes even harder for the delivery of aptamers *in vivo*. Most of the applied methods rely on protocols established for siRNA-transfection which have been recently summarized in excellent reviews [20, 21]. Briefly, lipofection-based methods are technically simple but do not always yield sufficient amounts of cytosolic aptamer to cause measurable effects. There are few examples for the successful delivery of aptamers into cells. An aptamer specific for the N-terminal part of cytohesin-2 was successfully transferred into cells via lipofection and led to downregulation of the serum-mediated MAPK activation [22]. Another example that should be mentioned here is an EGFRvIII-binding aptamer that intracellularly binds to the unglycosylated form of the receptor, thereby disrupting its posttranslational modifications [23].

The main alternative to lipofection is the intracellular production of the intramers by generating appropriate aptamer overproducing cell lines. This process has been realized by the application of different expression vectors [24, 25] and by adeno- [26] or vaccinia-virus-based expression systems [27]. However, the generation of stably transfected cell lines is time consuming and of limited use for therapeutic applications. These difficulties may explain why there is a big gap between the number of aptamers available for intracellular targets and the number of successfully used intramers. Consequently, there is a concern whether intramers are suitable therapeutic agents *in vivo*. A smart way to circumvent the aptamer delivery problem is to find cell-permeable inhibitory small molecules interfering with the aptamer-protein interaction by aptamer displacement screening [28].

Apart from aptamers for intracellular molecules, a variety of aptamers has been selected that bind extracellular proteins. The latter aptamers have the big advantage that they can be administered intravenously or subcutaneously. We can distinguish between shedded and cell-surface bound targets. Only the latter are within the scope of this paper.

In the case of extracellular targets, aptamer stability becomes a major issue. Therefore, stabilized RNA (e.g., by 2'-O-methyl or 2'-fluoro modifications [29–32]) or DNA is preferable when extracellular proteins are targeted. Even

then, degradation and clearance is inevitable, and repeated administration will be required until treatment is complete.

3. Cell-Specific Aptamers

Aptamers for extracellular targets have high potential for diagnostic and therapeutic applications. For diagnostics, their ability to differentiate among different cell types, namely, between tumor cells and normal cells, is highly interesting. More specialized aptamers can even be applied to distinguish between different types of cancers. In therapy, aptamers for extracellular targets can be used directly as effectors (activators or inhibitors) or indirectly as vehicles for drug delivery.

Surfaces of live cells and parasites represent very complex and convoluted targets. Although it is clear that the cell surface contains many recognition elements, for example, tumor markers in the case of cancer cells, its exact composition is unknown. Subtle differences in cell surface compositions can be very important and may be associated with the differentiation of cells and the onset of malignancies. Analyzing these changes and elucidating the exact composition of different cell types can yield important insights into the molecular details associated with these transformations. However, analysis of the molecular composition (e.g., by protein sequencing and mass spectrometric identification) is not sufficient to identify binders capable of targeting specific cell types.

Cell-SELEX is an evolutionary approach, and thus allows the selection of aptamers even without prior knowledge of specific targets [33–35]. In theory, cell-SELEX may generate aptamers for multiple targets in parallel. Another benefit of the cell-SELEX process is that it generates aptamers for accessible cell-surface markers. Aptamers can, therefore, speed up the discovery of new biomarkers that are per se targetable.

Of course, several cell-surface markers that can be used as target molecules in the traditional SELEX process have already been identified. Using purified proteins as target molecules certainly raises the odds of obtaining aptamers for this exact target molecule. However, cell-surface proteins tend to be membrane bound and often cannot be easily purified. In many cases, recombinant production and purification of extracellular domains is possible, but there is a concern whether the domain will still adopt the native conformation. In the worst case scenario, one could select aptamers that will only bind to the purified protein or protein domain and do not recognize the native form on living cells. This risk can be eliminated using cell-SELEX. Here, all the cell-surface molecules are presented in their natural surrounding and native conformation and, moreover, contain possible posttranslational modifications.

4. Recent Advances in Cell-SELEX

4.1. Different Ways of Obtaining Cell-Specific Aptamers. There has been a long-standing interest in the selection of aptamers for complex targets. Already in 1998, Gold and

coworkers used human red blood cell membranes as a model system to determine if SELEX could be used for a complex mixture of potential targets [36]. Aptamers for multiple targets were generated simultaneously during the selection process, and the binding affinities of these aptamers for their targets turned out to be comparable to those found in similar experiments with pure targets [36]. Using a secondary selection scheme (deconvolution SELEX), the authors rapidly isolated the aptamers for targets of special interest within the mixture.

In many cases, binding to a specific cell-type can be achieved by selecting aptamers for an abundantly presented protein on the cell-surface using traditional protein SELEX. Subsequently, the selected aptamers have to be tested for their ability to bind the native protein presented on the cell surface. One example is the *in vitro* selection of RNA aptamers using recombinant L-selectin as target molecule [37]. The resulting aptamers bound L-selectin with about 10^5 times higher affinity than the conventional oligosaccharide ligand sialyl Lewis X [37]. The nuclease stabilized aptamers showed calcium-dependent binding to native L-selectin on peripheral blood lymphocytes and blocked L-selectin-dependent interactions with the natural ligands on high endothelial venules. In a different approach, also DNA aptamers were selected with purified L-selectin. These even showed effects on lymphocyte trafficking *in vivo* [38].

Chen et al. selected RNA aptamers for the extracellular domain of the human epidermal growth factor receptor-3 (HER3) produced in insect cells and demonstrated that the aptamers interfere with receptor activation in MCF7 cells [39]. Another example for this strategy is the selection of 2'-F-modified aptamers using murine OX40 human IgG Fc fusion protein, a member of the tumor necrosis factor receptor superfamily [40].

Although the aforementioned examples were successful, it would be preferable to use the native form of the cell surface protein directly. However, cell surfaces are very complex targets. If aptamers for a specific target molecule within the cell surface are desired, cell-SELEX bears the risk to select aptamers for molecules different from the desired target. Therefore, some laboratories have combined selections using purified target proteins with cell-SELEX. Examples include aptamers for E-selectin and tenascin-C: E-Selectin is a cell adhesion molecule presented only on endothelial cells activated by cytokines. Mann et al. obtained thio-aptamers binding to endothelial cells presenting E-selectin [41]. Tenascin-C is an extracellular matrix protein that is overproduced during tissue remodeling processes, including tumor growth. Hicke et al. performed protein SELEX, cell-SELEX, and a crossover of both and selected aptamers with K_d values in the low nanomolar range [42].

Although both, the traditional SELEX and a combined approach with cell-SELEX, have been successfully applied for known targets, they do not exploit the full range of possibilities that cell-SELEX has to offer. If live cells are used, there is no need for prior information of the cell surface composition—a true asset. Figure 1 shows a schematic cell-SELEX process including the most common separation methods, which are centrifugation of cells in suspension,

washing of adherent cells, and fluorescence activated cell sorting (FACS).

As complex target for a cell-SELEX, Homann and Göringer used live *Trypanosoma brucei*, a sleeping sickness causing protozoan parasite [43]. Nonbinding nucleic acids were removed by repeated cycles of washing and centrifugation. Finally, they identified RNA aptamers specific for a 42-kDa protein located in the parasite's flagellar pocket. Bruno and Kiel selected DNA-aptamers for different spores such as *Anthrax* or *Bacillus thuringiensis* [44].

Blank et al. used endothelial cells as targets and selected DNA aptamers that could be used as histological markers of microvessels in brain tumors [45]. Partitioning of bound and unbound ssDNA molecules was done by centrifugation. The progress of the selection as well as the binding properties of individual aptamers were analyzed by FACS.

Gold and coworkers performed cell-SELEX using the tumor cell line glioblastoma U251 as target. They selected DNA aptamers for tenascin-C by simply incubating the nucleic acids with cell monolayers in serum-free medium and repeated washing steps [46].

This technique was also used by Tan and coworkers who published multiple aptamers that bound selectively to certain cancer cell lines [47]. To do so, they performed a counter selection with a control cell line before the actual incubation step to remove cell matrix binders. Table 1 gives an overview of cell lines used for live cell-SELEX.

4.2. Technical Challenges and Advances in Cell-SELEX. Until recently, most technical advances of the SELEX procedure focused on improving the individual steps of the process, in particular the actual selection step, in order to make the generation of aptamers as fast and efficient as possible. The main points for optimization are automation, separation, counter selection, and the partitioning of dead cells. Automation of the SELEX process was introduced by Ellington and coworkers [48]. As a consequence, selection times were significantly reduced. Since then, aptamers are also commercially available.

Further optimization is required to avoid matrix-binding species within an *in vitro* selection. In the classical protocol for SELEX, the target protein is immobilized on a solid support such as agarose, sepharose, or magnetic beads. The latter have become particularly popular due to low matrix binding. Alternative separation protocols include filter binding and innovative approaches such as capillary electrophoresis [49]. As mentioned above, the classical approaches have been used successfully to select aptamers for cell-surface proteins that recognize specific live cells.

Target immobilization however, is not compatible with live cells. Most cell-SELEX protocols so far separate binding from nonbinding nucleic acids by washing (adhesive cells) or centrifugation (cells growing in suspension). For obvious reasons it is not easy to enrich exclusively the desired target-binding species in complex target SELEX approaches, and therefore, counter selections have to be applied in most cases (Table 1).

Cerchia et al. performed even two of those counter selections, one with suspended parental Pc12 cells and a second

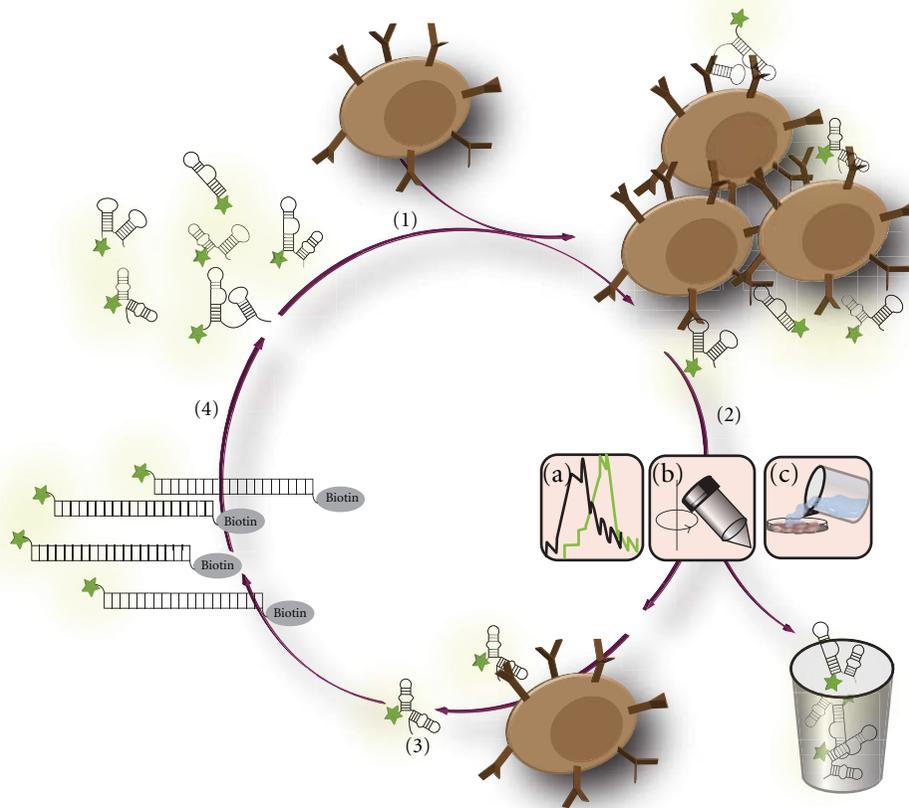


FIGURE 1: Cell-SELEX. (1) Target cells are incubated with a complex library consisting of up to 10^{15} individual single stranded DNA molecules (ssDNA). (2) The separation of cell-binding nucleic acids from nonbinding species can be performed by different methods: FACS (2a), centrifugation (2b) or washing (rinsing) steps (2c), respectively. (3) Cell-binding DNA molecules—aptamers—are eluted and (4) subsequently amplified by PCR using two specific primers. The 5'-end of the reverse primer carries a biotin. Additionally, a fluorescently labeled forward primer is used for PCR, if FACS is used for separation. After strand displacement the enriched DNA library is used for further selection rounds. This cell-SELEX procedure ends up with aptamers that bind specifically to the surface of target cells.

TABLE 1: Cell-lines used for cell-SELEX and for corresponding counter selection. The column “When?” indicates whether counter selection was performed before (B) or after (A) the selection, or not at all (—).

Selection cell line	Abbreviation	Counter selection	Abbreviation	When?	Lit.
T-cell line, human ALL	CCRF-CEM	B-cell line, human Burkitt's lymphoma	Ramos	A	[54]
Carcinoma; small-cell lung cancer	NCI-H69	Carcinoma; large-cell lung cancer	NCI-H661	A	[55]
Mouse liver hepatoma cell line	BNL 1ME A.7R.1 (MEAR)	Normal mouse liver cell line	BNL CL.2	A	[56]
Acute myeloid leukemia	HL60	Acute promyelocytic leukemia	NB4	A	[57]
Human lung carcinoma cells infected with vaccinia virus	A549	Human lung carcinoma cells without infection	A549	A	[58]
Pheochromocytoma of the rat adrenal medulla expressing human RET ^{C634Y} mutant	Pc12/MEN2A	Pheochromocytoma of the rat adrenal medulla parental cells followed by cells expressing human RET ^{M918T} mutant	Pc12 and Pc12/MEN2B	B	[50]
B-cell line, human Burkitt lymphoma	Burkitt lymphoma	None, but FACS to sort out dead cells		—	[10]
Human glioblastoma U251	U251	None		—	[46]
Endothelial cell line YPEN-1	YPEN-1	N9 microglial cells		B	[45]

one with adherent Pc12/MEN2B cells before incubating their 2'-F-modified RNA library with Pc12/MEN2A cells, the actual selection target presenting the human RET^{C634Y} mutant receptor [50].

Most people who have actively selected aptamers are familiar with the problems of matrix binders. The logic behind this problem is apparent and has been called the first law of directed evolution: "You get what you select for" [51]. Since in every cycle, multiple selection pressures are involved, it is not surprising that artifacts may arise. It can be anticipated that this problem becomes more serious when using complex surfaces of cells or spores as selection target.

To make things even worse, there is an additional problem for cell-SELEX with live cells: in every cell preparation also dead cells are present. Dead cells show strong, sequence-independent affinity for nucleic acids [52]. They cannot be removed by traditional methods like washing or centrifugation and thus lead at least to an inefficient selection process or even to failure of the experiment [53]. Sample preparation is, therefore, very important. These problems may be the reason why cell-SELEX has not yet become a widely used method for the generation of cell-specific aptamers.

The above issues were addressed in a recent publication by the laboratories of Mayer and Famulok [10]. They used FACS to remove the fraction of dead cells and to select aptamers binding to live cells. Thus, they obtained aptamers specific for B-cells that were able to differentiate between different cell subpopulations.

5. Cell-Specific Aptamers for Affinity Purification and Diagnostics

There are numerous applications for aptamers, including aptamers as sensors, as inhibitors, as diagnostics, as therapeutics, and as delivery agents. There are excellent reviews dealing with these topics, for example, [17]. Here, we want to focus on the application of cell-specific aptamers for affinity purification, diagnostics, and most importantly therapy.

The first and obvious application of cell-specific aptamers is their use for affinity purification. Affinity purification has been described with several aptamers selected for single targets [59, 60]. More interesting purification tools are aptamers with affinity for complex targets like cells. The specific target molecule may be identified afterwards using mass spectrometry.

Using this approach Blank et al. took advantage of an aptamer that binds to tumor microvessels for affinity purification and identified p187 as the target protein [45]. The above-mentioned aptamer selected for human glioblastoma U251 cells was also used to identify the target protein tenascin-C [46]. An inherent advantage of this aptamer-based target purification-identification approach is that the identified targets are per se addressable.

Tan and coworkers used several of their cell-specific aptamers for extraction of specific cells from complex mixtures. They reported aptamers for small-cell lung cancer (SCLC) cells and used them in conjugates with nanoparticles to isolate and enrich SCLC cells from mixed cell media [55].

The same laboratory developed a method for the rapid collection and detection of leukemia cells based on an aptamer for CCRF-CEM acute leukemia cells and two kinds of nanoparticles (magnetic and fluorescent). This dual approach allowed target cell extraction from complex mixtures including whole blood samples as well as sensitive cell detection [61].

Aptamers have also a long tradition in classical diagnostic applications, in analogy to antibodies. For the creation of sensors, it is ideal to use a ligand-induced conformational change to create a signal. This has been realized by differential dye binding, fluorescence quenching, FRET, or coupling to different types of nanoparticles. Also, modular aptamer sensors consisting of a recognition domain and a signaling aptamer have been constructed (for a review, see [62]). Recently, a sensor based on a cell-specific aptamer was realized by a gold nanoparticle-aptamer conjugate on a lateral flow device that could be used for the detection of Ramos cells [63]. This approach represents first steps towards the detection of circulating cancer cells.

6. Cell-Specific Aptamers for Therapy

Aptamers are not only a new and promising alternative to antibodies in diagnostics. They may also be used in therapy and medicine. For therapeutic applications, it is of particular interest to selectively deliver drug molecules to disease-related cells or tissues of interest thereby minimizing the exposure of these possibly harmful agents to surrounding healthy tissues. This target-controlled transfer can be achieved by conjugating desired drugs to carriers that specifically interact with components (e.g., proteins) only presented on the surface of pathologically relevant cells. Cell-surface proteins are targeted by more than 60% of all currently available drugs—underscoring their therapeutic importance [64]. Receptors play the most prominent role among these cell-surface proteins. They receive a signal, transduce it *via* the cellular membrane, and forward it inside the target cell in many cases to finally trigger a cellular response, such as gene expression.

As mentioned before, cell-SELEX can yield aptamers that are able to interact with specific disease-related proteins on cell surfaces. These aptamers are potential therapeutic compounds because they may cause inhibitory as well as stimulatory effects. Furthermore, most cell-surface proteins undergo recycling processes, such as ligand-induced internalization. This opens new strategies for aptamer-based drugs. Aptamers binding to receptors that are subject to internalization can be infiltrated into target cells, and thus serve as vehicles for desired cargo molecules.

In the following sections, we will give an overview of the current state of cell-specific aptamers. We have classified the aptamers into three categories: aptamers eliciting (i) agonistic or (ii) antagonistic effects on the target protein as well as (iii) aptamers as drug delivering agents.

6.1. Aptamers as Blocking Agents. Aptamers might be of therapeutic relevance due to their inhibitory effects on disease-related cell-surface components. One promising approach

would be the application of aptamers acting on deregulated processes in inflamed tissues or pathological vasculature during angiogenesis.

In 2010, Mann et al. reported phosphorothioate-modified aptamers binding to E-selectin on endothelial cells and above all to the inflamed vasculature of human carcinomas presenting E-selectin in high amounts on their cell surfaces [41]. Selectins comprise a family of calcium-dependent cell adhesion glycoproteins which can be subdivided into 3 groups: E-selectin that can be found in endothelial cells, L-selectin on the surface of leukocytes, and P-selectin on platelets and endothelial cells. All family members play crucial roles during normal and abnormal inflammation processes. Due to structural similarities, they are all able to interact with specific carbohydrates (sialyl Lewis X and sialyl Lewis A) on the surface of leukocytes thereby mediating the contact of the leukocytes to the endothelial wall [70]. E-selectin levels are increased on the endothelial surface of blood vessels within numerous inflammatory diseases, for instance, rheumatoid arthritis and several types of cancer [71]. Therefore, targeting E-selectin in inflamed tissues is a therapeutically highly relevant strategy to prevent cell transmigration processes along endothelial tissues. Mann et al. performed a two-step SELEX procedure (see above). In the first 10 SELEX rounds, purified recombinant E-selectin served as target. Subsequently, the 14 obtained aptamers were screened for binding to E-selectin presented on endothelial cells. Only one of the aptamers named ESTA-1 exhibited binding to cell surface bound E-selectin. Furthermore, this aptamer inhibited the adhesion of sialyl Lewis X positive HL-60 cells to E-selectin producing endothelial cells (Figure 2). After intravenous administration, a distinct interaction of the E-selectin-specific aptamers with tumor vasculature could be highlighted in a breast cancer xenograft model [41].

Within the same month Gutsaeva et al. reported on the application of the P-selectin binding RNA aptamer ARC5690 that contained 2'-fluoro pyrimidine and 2'-methoxy purine building blocks for increased nuclease stability. The aptamer effectively blocked the adhesion of circulating sickle red blood cells (RBCs) and leukocytes to endothelial cells in sickle cell disease (SCD) after injection into SCD mice [65]. Administration of the aptamer led to a decreased mortality rate of the treated mice opening the possibility to use these anti-P-selectin aptamers as potential anti-cell adhesion compounds in sickle cell disease.

In 1996, O'Connell et al. succeeded in selecting 2'-amino-modified RNA aptamers binding to L-selectin-Ig chimera (LS-Rg) [37]. The aptamers revealed a calcium-dependent binding to native L-selectin on peripheral blood mononuclear cells (PBMCs). Moreover, the aptamers specifically inhibited the interaction of LS-Rg and of human PBMCs with lymph node high endothelial venules that carry the natural ligands of L-selectin. Nevertheless, there was one significant drawback: the aptamers possessed their highest affinity at 4°C and 22°C. At 37°C, the affinity for L-selectin was remarkably lower. Thus, these aptamers did not seem to be particularly suitable for *in vivo* applications [38].

Consequently, Hicke et al. reported on the selection of DNA aptamers that did bind to L-selectin with nanomolar affinity [38]. These aptamers inhibited the L-selectin interaction with sialyl Lewis X. Additionally, the aptamers bound specifically to human L-selectin on lymphocytes and neutrophil cells. The DNAs inhibited lymphocyte rolling on activated endothelial cells in an *in vitro* flow system. *In vivo*, the aptamers prevented the homing of human lymphocytes to lymph nodes in severe combined immunodeficiency (SCID) mice.

Aside from adhesion molecules, also cytokines and their respective receptors are involved in a variety of inflammatory diseases like osteoarthritis (OA), the most common degenerative joint disease, caused by cartilage loss in a joint [72]. The cytokine interleukin-17 in conjunction with one of its receptors IL-17R is involved in the pathogenesis of OA. IL-17R is a glycoprotein that displays a wide tissue distribution and is a central mediator in inflammatory responses [66]. Following IL-17R activation multifunctional cytokines like IL-1 and IL-6 are produced and enforce inflammatory processes. Chen et al. reported on a novel cell-based selection method leading to DNA aptamers specific for IL-17R. The authors used whole IL-17R presenting NIH3T3 cells as targets and unmodified NIH3T3 cells as negative control for counterselection. Aptamer RA10-6 was able to interfere with the interaction of IL-17 and IL-17R *in vitro*. After intra-articular injections in OA mice the aptamers impeded synovial thickening and inflammation in combination with celecoxib, a small nonsteroidal anti-inflammatory drug (NSAID) often used in the treatment of osteoarthritis. Furthermore, its administration led to a dose-dependent reduction of IL-6 levels. In the future, it may be conceivable to additionally administer RA10-6 to NSAIDs in standard therapies for OA to circumvent high NSAID doses and associated side effects.

Aptamers binding to receptor tyrosine kinases (RTKs) also reveal a high therapeutic potential. RTKs are cell surface receptors that bind their ligands like growth factors or hormones with high affinity. Besides their role in the regulation of common cellular processes, such as cell growth and differentiation, RTKs are involved in the development and progression of many cancer types [73]. One example is the human epidermal growth factor receptor-3 (HER3), a member of the type I RTKs. HER3 is catalytically deficient and self-associates in the absence of its ligand heregulin [74]. After binding of heregulin to HER3, heterodimerization with other RTKs, especially HER2, occurs. Overproduction of both HER2 and HER3 can be found in several types of cancer [75]. Chen et al. were able to select RNA aptamers binding HER3 in its oligomeric state [39]. One aptamer, A30, was further investigated. It did not compete with heregulin for binding to HER3 but inhibited heregulin-dependent tyrosine kinase activity of HER2 and growth of MCF7 cells. Although the inhibition mechanism could not be clarified in detail, this study demonstrates the potential of aptamers as anticancer drugs.

Another RTK used as target for the enrichment of aptamers was the human RET receptor tyrosine kinase [50] (selection see above). One of the obtained aptamers

TABLE 2: Aptamers with inhibitory function on target molecules and corresponding selection strategies.

Target	Aptamer	Selection method	Lit.
E-selectin	Thiophosphate-mod. DNA	Two-step selection using recombinant ES-Ig chimera and testing on cells as target	[41]
P-selectin	2'-amino- and 2'-methoxy-mod. RNA	Traditional SELEX using recombinant mouse P-selectin as target	[65]
L-selectin	2'-amino-mod. RNA and DNA	Traditional SELEX using recombinant L-selectin-Ig chimera as target	[37, 38]
IL-17R	DNA	Cell-SELEX using IL-17R presenting NIH3T3 cells as target, unmodified NIH3T3 cells for counter selection	[66]
HER3	RNA	Traditional SELEX using the extracellular part of HER3 as target	[39]
RET	2'-F-mod. RNA	Cell-SELEX using RET ^{C634Y} receptor presenting Pc12/MEN2A cells as target, Pc12 and Pc12/MEN2B cells for counterselection	[50]
NgR	RNA	Traditional SELEX using the extracellular part of NgR-Fc fusion protein as target	[67]
CTLA-4	2'-F-mod. RNA	Traditional SELEX using CTLA-4-Fc fusion protein as target	[68]
gp120	2'-F-mod. RNA	Traditional SELEX using purified recombinant gp120 as target	[69]

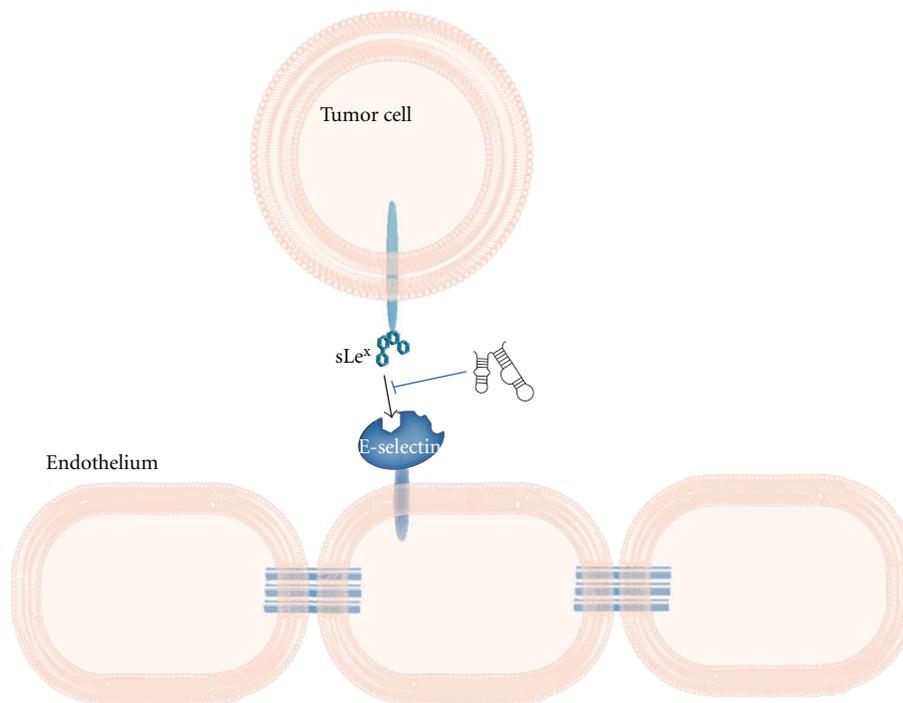


FIGURE 2: Aptamers as blocking agents. Aptamers, in this case composed of phosphorothioates, selectively bind to E-selectin that is presented on the surface of human endothelial cells. The aptamers serve as antagonist for the interaction to sialyl Lewis X on HL-60 cells, a human promyelocytic leukemia cell line, thereby inhibiting the adhesion of tumor cells by more than 75%.

bound to different cell types presenting RET on their cell surface. The aptamer interfered with ligand-induced receptor dimerization or dimerization caused by activating mutations within the receptor, and thereby inhibiting intracellular RET signaling pathways.

In another study, aptamers were used to block the Nogo-66 receptor (NgR) on axons after spinal cord injury [67]. Axon regeneration is usually prevented by three inhibitors from the myelin sheath, which are ligands of NgR [76]. The NgR-specific aptamers blocked the myelin-derived inhibitors

TABLE 3: Aptamers for cell-specific delivery of drug molecules.

Target molecule	Aptamer	Drug molecules linked or complexed with the aptamer	Lit.
PSMA	2'-F-modified RNA	(1) siRNAs/shRNAs (2) Protein toxin gelonin (3) Dox (4) QDs, SPION or gold NPs with Dox (5) Pt(IV)-encapsulated PLGA-PEG NPs	(1) [88–92] (2) [93] (3) [94] (4) [95–99] (5) [100]
gp120	2'-F-modified RNA	siRNAs	[101, 102]
Nucleolin	DNA	PDT agent TMPyP4	[103]
PTK7	DNA	Dox	[104]
TfR	DNA	L-iduronidase	[105]
MUC1	DNA	PDT agent chlorin <i>e</i> ₆	[106]
EGFR	RNA	Gold NPs	[107]
IgM heavy chain	DNA	Micelle-like nanostructures	[108]

by competing for binding to NgR [67] and fostered neurite outgrowth of primary cultured neuronal cells. This result highlights the therapeutic potential of aptamers in the treatment of myelin-associated inhibition of neuronal regeneration.

Santulli-Marotto et al. developed a promising aptamer-based strategy to inhibit the activation of an immune receptor. The authors selected aptamers specific for the T-cell receptor CTLA-4 [68]. CTLA-4 (cytotoxic T lymphocyte antigen-4) is a high affinity receptor for the B7 family members of immune regulatory ligands presented on activated T-cells. This receptor exists in a dimeric form and is stimulated by two molecules of the ligand protein B7-1 [77]. CTLA-4 inhibits antitumor immunity. It transmits an inhibitory signal to T cells, which reduces their ability to respond to stimuli by enhancing the signaling threshold. As a result higher amounts of signal molecules would be required for further stimulation. Inhibitors of CTLA-4 (e.g., antibodies) are potential agents to enhance T-cell anti-tumor immunity [78]. CTLA-4-specific high affinity aptamers enforced tumor immunity in mice [68]. Tetramers of these CTLA-4-binding aptamers showed even stronger inhibitory effects on CTLA-4 function *in vitro* and *in vivo*. These constructs seem to be useful anti-tumor agents and regulators of the immune system.

Similarly 2'-F-modified aptamers could be selected with the potential to lower HIV-1 (human immunodeficiency virus 1) infectivity by blocking gp120, a cell surface receptor, both, on HIV-1 and on HIV-1 infected immune cells [69, 79, 80]. Further developments and applications will be discussed later.

6.2. Aptamers as Stimulating Agents. In this section we will focus on aptamers causing activating effects by binding to therapeutically relevant cell-surface proteins.

Recently, McNamara et al. introduced agonistic aptamers specific for the costimulatory receptor 4-1BB, a member of the tumor necrosis factor (TNF) receptor family [81]. This receptor is presented on activated CD8⁺ T-cells and pro-

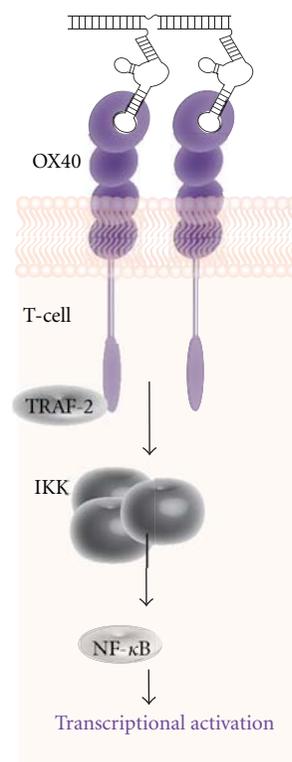


FIGURE 3: Aptamers as stimulating agents. Dimerized OX40-specific RNA aptamers show activating effects on the OX40 receptor *in vitro* and *in vivo*. The OX40 receptor on the surface of activated T cells undergoes receptor oligomerization and provides a costimulatory signal for T cell activation. Downstream factors (e.g., TRAF-2, IKK and NF-κB) of the signaling cascade are activated leading to an enhanced proliferation of T cells and an increased anti-tumor immune response.

longs the survival and expansion of CD8⁺ T lymphocytes [82]. Since CD8⁺ T-cells play an important role in tumor immunity, enhancing 4-1BBs costimulatory activity could support anti-tumor immune responses. Indeed, agonistic

monoclonal antibodies against 4-1BB that had been systemically administered to mice ameliorated tumor immunity and tumor rejection [83, 84]. Aptamers for 4-1BB represent an attractive alternative because antibodies can be immunogenic in some cases and might cause problems in clinical trials. The aptamers specifically bound to 4-1BB presented on the surface of activated T cells [81]. Stimulation of receptors of the TNF receptor family, including 4-1BB, often involves oligomerization. Notably, the oligomerization could be induced by dimeric and multimeric forms of the aptamers and triggered the activation of 4-1BB. Thus, the aptamer-based di- and multimers acquired the ability to costimulate T cells and induce tumor rejection in mice. As mentioned above, there are aptamers that block CTLA-4 functions on activated T-cells *in vivo* and enforce tumor immunity in mice [68]. Combinations of aptamers for CTLA-4 and 4-1BB may even be better antitumor agents than each aptamer by itself.

Dollins et al. extended the approach of inducing receptor multimerization by aptamers to trigger agonistic effects. They selected aptamers for the target molecule OX40 [40]. OX40, another member of the TNF receptor family, is presented on the surface of activated T cells. Binding of its ligand OX40L (OX40 ligand) causes receptor oligomerization. This activation leads to an enhanced proliferation rate of T cells, to an increased cytokine production, and to an intensified anti-tumor immune response. After selection, Dollins et al. converted the monomeric aptamers into receptor activating dimers using a molecular scaffold (Figure 3). This aptamer complex mimicked the OX40 ligand and induced OX40 multimerization and function both *in vitro* and *in vivo* in a tumor immunotherapy approach. Systemic administration of the agonistic aptamer complex showed a remarkable effect in the treatment of highly metastatic and aggressive types of melanoma [40]. Both the 4-1BB and the OX40 aptamers impressively demonstrated how cell-surface receptor binding aptamers could be turned into receptor activating complexes by di- or multimerization [40, 81]. These nucleic acid-based agonists could become a new safe and powerful class of therapeutically relevant effectors.

A related approach combined two DNA aptamers for different tumor associated targets [85]. Boltz et al. described the generation of chimeric DNA aptamers that simultaneously bound two surface receptors of immune cells and mediated tumor lysis.

Specifically, one of the aptamers bound to the Fcγ receptor IIIA also known as CD16α. CD16α belongs to the immunoglobulin superfamily and is mostly presented on natural killer cells but also on monocytes and certain tissue macrophages, respectively. CD16α plays an important role during the antibody-dependent cellular cytotoxicity (ADCC). The recruitment of CD16α-positive NK-cells to specific tumor cells could increase the antitumor-effective function during ADCC. The second aptamer bound to the hepatocyte growth factor receptor (HGF-R) also known as c-Met. The c-Met receptor belongs to the RTK family. As mentioned above, RTKs are important regulators in cellular processes and thus key players in the development and progression of many cancer types [86]. The c-Met protein is presented on epithelial cells, important for embryonic

development and overproduced in many tumors [85, 86]. Aptamers recognizing both, the CD16α and c-Met could potentially recruit NK-cells to c-Met-positive tumor cells, inducing ADCC and, therefore, acting as anti-tumor agents. Boltz et al. generated high-affinity aptamers for CD16α and for c-Met and connected them by an oligonucleotide linker. The resulting bi-specific aptamers mediated cellular cytotoxicity, induced killing of NK cells and subsequently specifically lysed tumor cells in ADCC assays. In the future, the authors will expand these findings to *in vivo* experiments using xenograft mouse models. Taken together the described bi-specific aptamers illustrate the ability of nucleic acid-based therapeutics to stimulate the lysis of tumor cells, a property that opens up the possibility to serve as reagents in cancer therapy.

6.3. Aptamers for Delivery of Toxic Payloads. Aptamers can also serve as modules that selectively recognize and bind to defined cell types or tissues. By appending drug molecules, the aptamers can be used to deliver cargo molecules to or into specific cells or tissues of interest.

6.3.1. Prostate Specific Membrane Antigen (PSMA). One of the best-studied aptamers binds to the prostate specific membrane antigen (PSMA), a well-known tumor marker for prostate cancer [87].

In 2002, Lupold et al. reported the selection of 2'-F-modified RNA aptamers binding to the extracellular portion of PSMA. Aptamer A10 and its truncated version A10-3 were able to bind PSMA presenting human prostate cancer cells with high affinity and served as a prostate specific tumor cell marker [8]. The detection of specific endocytosis of PSMA into cells via clathrin-coated pits suggested that PSMA-specific aptamers might be potential drug delivery vehicles [109]. Therefore, aptamer-drug chimeras were constructed. The idea was that the aptamers would specifically recognize and bind to prostate cancer cells and then become endocytosed together with PSMA—resulting in drug delivery to the target cells.

Several groups created chimeras consisting of small interfering RNAs (siRNAs) and the PSMA-specific aptamer A10 [88–91]. McNamara et al. used a completely RNA-based approach by directly connecting the PSMA aptamer A10 at its 3'-end with therapeutic siRNAs specific for *plk1* and *bcl-2*, two survival genes that are overexpressed in many human tumors [89] (Figure 4). The chimeras were supposed to bind PSMA on target cells via the aptamer and to silence the target mRNAs via the siRNA portions. Administration of the chimeras to PSMA presenting cells indeed resulted in endocytosis, gene silencing, reduced cell proliferation, and finally cell death. Importantly, the chimeras did not affect cells that did not present PSMA. The next step was the intratumoral application of the aptamer-siRNA chimeras into athymic mice that carried tumors derived from PSMA-positive human prostate cancer cells. In these mice, the chimeras indeed reduced tumor growth. The same experiment was performed with mice bearing tumors from PSMA-negative human prostate cancer cells. These tumors were

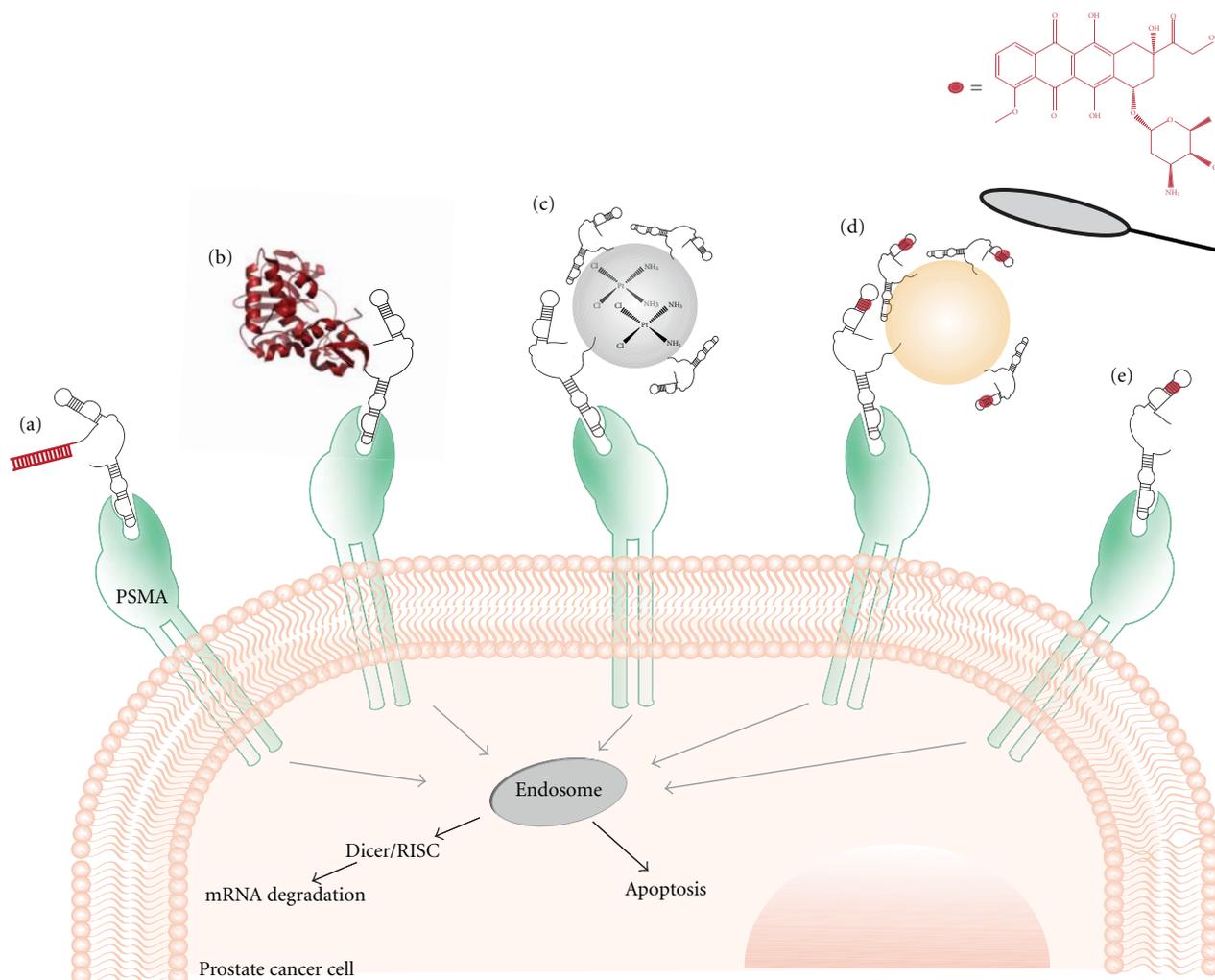


FIGURE 4: Aptamers as drug delivery agents. Prostate specific membrane antigen (PSMA) is internalized into cells via endocytosis. PSMA-binding RNA aptamers serve as drug delivery vehicles for small interfering RNAs (siRNAs, (a)). The siRNAs escape from the endosome, are processed by Dicer and RISC, and lead to the degradation of target mRNAs. Furthermore, the aptamers can escort cytotoxic molecules into prostate cancer cells like the ribosomal toxin gelonin (b), cisplatin encapsulated in organic PLGA-PEG-NPs (c), nanoparticles (NPs, gold, iron oxide and quantum dots; (d)), or doxorubicin as a physical conjugate (e). For references see text.

not affected by the chimeras. This strongly suggests that the cellular uptake of the RNA ligands was highly efficient and specific for PSMA. The aptamer-siRNA chimeras were further optimized by the same group [90]. The authors truncated the aptamer portion and enhanced the silencing activity and specificity by incorporating modifications within the siRNA molecules leading to a more efficient processing effectivity of the RNAi machinery. The optimized chimeras were systemically administered into athymic mice bearing tumors from PSMA-positive human prostate cancer cells leading to distinct tumor regression or anti-tumor activity, respectively.

Chu et al. used another conjugation strategy between a PSMA-specific aptamer and two different siRNAs specific for lamin A/C or GAPDH (Glyceraldehyde 3-phosphate dehydrogenase) mRNAs [91]. Both the aptamer and the siRNAs were chemically modified with biotin and bound to

tetrameric streptavidin in a 2 : 2 : 1 ratio. The resulting RNA-protein complexes were cell type-specifically internalized and caused down regulation of target genes and cell-death of PSMA-presenting cancer cells. Wullner et al. created multivalent aptamer-siRNA constructs to further increase potency and anti-tumor capacity of the chimeras [88]. The authors generated two different bivalent chimeras consisting of the anti-PSMA aptamer and siRNA specific for the *EEF2* (eukaryotic elongation factor 2) mRNA. In one case, the siRNA itself served as linker between two aptamer molecules, whereas in the other case, siRNAs were attached to the 3'-end of each aptamer. The increased valency of the aptamer-siRNA chimeras enhanced cytotoxicity with respect to the monovalent counterparts.

Very recently, Ni et al. used aptamer-shRNA chimeras to inhibit DNA repair pathways in prostate cancer cells and subsequently to enhance ionizing radiation (IR) therapy

[92]. The target mRNA of the shRNA was the catalytic subunit of DNAPK (DNA-activated protein kinase, catalytic polypeptide). DNAPK is a serine/threonine protein kinase that is necessary for the double-strand break repair in DNA [110]. Therefore, downregulation of DNAPK by aptamer-shRNA chimeras should selectively sensitize PSMA-positive cells to IR. Cell-specific delivery of the aptamer-shRNA chimeras to PSMA-positive LNCaP cells and to tumor models led to reduced levels of DNAPK mRNA [92]. A combination of the application of aptamer-shRNA chimeras and IR for the treatment of LNCaP tumor models resulted in an extended reduction of tumor volume due to enhanced radiosensitivity. Finally, the authors could also demonstrate high efficacy of the chimeras in human prostate tissue.

Apart from aptamer chimeras with siRNA, aptamers can be fused to toxic proteins to deliver them into target cells. This approach was realized by covalently linking a PSMA-specific aptamer [93] to a gelonin derivative (Figure 4). Gelonin is a ribosome-tackling toxin that cleaves a specific glycosidic bond in rRNA subsequently disrupting protein biosynthesis. This gelonin derivative lacks the translocational domain and, in contrast to related toxins, is not translocated per se into cells. In conjunction with the aptamer, however, the gelonin variant was internalized into PSMA-presenting tumor cells consequently destroying them. This unequivocally demonstrated the ability of the PSMA-aptamer to deliver a therapeutically relevant toxin as payload into target cells.

Chemotherapeutic drugs represent another class of drug molecules that had been successfully delivered into PSMA-positive human prostate cancer cells via the aptamer A10 [94, 95, 100]. Bagalkot et al. produced physical aptamer-doxorubicin conjugates not requiring any chemical modification neither of the drug nor of the aptamer [94] (Figure 4). Doxorubicin (Dox) can intercalate into the aptamer portion. Dox is an anthracycline-derived drug molecule used in chemotherapy to treat a wide range of cancers. Since this treatment causes many side effects, targeted delivery would be desirable to decrease doses and associated cytotoxicities. The authors studied *in vitro* binding and cell-specific uptake of the physical conjugates by the PSMA-positive prostate cancer cell line LNCaP to evaluate the targeted drug delivery. A significant reduction of the cell proliferation as well as enhanced cytotoxicity occurred in comparison to control cells [94].

The same group immobilized the PSMA-specific aptamer A10 on the surface of fluorescent quantum dots (QDs). Additionally, the A10-derivatized QDs were incubated with Dox yielding QD-A10-Dox conjugates [97] (Figure 4). Although not covalently linked, the aptamer-Dox-complex quenched the fluorescence of the QDs. When incubated with target cells these multifunctional nanoparticle conjugates delivered Dox into PSMA-positive cells. Cellular uptake could be detected due to the recovered fluorescence emission of the QDs that subsequently stained the corresponding cancer cells. Additionally, Dox inhibited the growth of those cells.

In related approaches, superparamagnetic iron oxide nanoparticles (SPION) or gold nanoparticles were derivatized with A10 and Dox to create multifunctional nanoparticle

conjugates for both imaging and treatment of prostate cancer cells [98, 99]. Each of the three parts of the conjugates had a distinct function: the aptamer portion served as delivery vehicle, the nanoparticle served as platform and imaging agent in magnetic resonance imaging (MRI) or computer tomography (CT), and Dox served as toxin (Figure 4).

The physical conjugates, however, were only able to carry a limited amount of cargo molecules. A new strategy with higher capacity was therefore developed by Bagalkot et al. [94]. This group produced biocompatible and biodegradable nanoparticles (NPs) derived from poly(D,L-lactic-co-glycolic acid)-block-poly(ethylene glycol) (PLGA-b-PEG) copolymers. The surface of the NPs was coated with the PSMA-specific aptamer A10. The authors encapsulated the antimetabolic chemotherapy drug docetaxel (Dtxl) into the NPs, resulting in Dtxl-containing nanoparticle-aptamer bioconjugates (Dtxl-NP-Apt). Treating LNCaP prostate cancer cells with Dtxl-NP-Apt led to a specific uptake by these cells and furthermore to a significant enhancement of cytotoxicity *in vitro* in comparison to nanoparticles lacking A10. Intratumoral injection of the bioconjugates into LNCaP-derived tumors in xenograft nude mice diminished tumor growth in five out of seven mice.

Dhar et al. encapsulated the cisplatin prodrug Pt(IV) within PLGA-PEG nanoparticles that were coated with the PSMA aptamer A10 [100] (Figure 4). After cell-specific uptake of the Pt(IV)-NP-Apt bioconjugates, Pt(IV) was converted by cytosolic reduction processes into cisplatin. Cytotoxicity assays with LNCaP cells revealed a higher *in vitro* cytotoxic activity for the Pt(IV)-NP-Apt bioconjugates with respect to free cisplatin or bioconjugates lacking the aptamer portion. Additionally, the same laboratory recently analyzed the pharmacokinetics, biodistribution, tolerability, and efficiency of the Pt(IV)-NP-Apt bioconjugates when administered *in vivo* in comparison to cisplatin [95]. The authors observed a simultaneous improvement of all these factors and proved the bioconjugates to be an efficient nontoxic item for chemotherapeutic treatment of human prostate cancer.

6.3.2. *Glycoprotein 120 (gp120)*. Envelope glycoprotein gp160 of HIV-1, consisting of two parts, gp120 (glycoprotein 120) and gp41 (glycoprotein 41), respectively, plays a crucial role in HIV-1 infection. Therefore, blocking HIV-1 entry into T cells with the aid of corresponding aptamers seems to be a therapeutically promising strategy. James et al. reported the selection of 2'-F-modified RNA aptamers that bound to gp120 with high affinity, and thereby neutralized the infectivity of R5 strains of HIV-1 in human peripheral blood mononuclear cells (PBMCs) by blocking the interaction between gp120 and CCR5 [69, 79, 80]. Zhou et al. published in 2008 the application of these RNA aptamers [101]. The authors designed aptamer-siRNA-chimeras that bound and entered target cells via gp160-mediated endocytosis. The siRNA portions were processed by dicer and specifically downregulated the mRNA levels of the HIV-transcripts *tat* and *rev*. This was the first report on aptamer-siRNA-chimera in which both parts inhibited HIV-1 functions.

One year later, the same group published the selection of new 2'-F-modified gp120-binding aptamers that blocked HIV-1 infection of cultured CEM T cells. Aptamers were endocytosed into gp160-presenting target cells [102]. The authors again took advantage of the cell-specific delivery process of gp120-binding agents. Application of optimized aptamer-siRNA-chimeras resulted in a successful downregulation of *tat*- and *rev*-mRNAs, thereby reducing HIV-1 replication and infectivity in PBMCs. Taken together, the anti-gp120 aptamers may be promising agents for an anti-HIV therapy due to their ability to inhibit HIV-1 infectivity and replication or due to their potential to escort further drugs into affected cells.

6.3.3. Nucleolin. Nucleolin, a multifunctional eukaryotic nucleolar phosphoprotein, plays a pivotal role in RNA transcription, in DNA replication, and in the regulation of different steps of ribosome biogenesis [111]. Furthermore nucleolin binds to the mRNA of Bcl-2, an antiapoptotic protein overproduced in several cancer types, and thereby increasing its half-life [112]. Nucleolin is overproduced in the cytoplasm and on the plasma membrane of certain tumor cells [113]. AS1411, a 26 nt long nucleolin-binding DNA aptamer, is taken up by nucleolin-presenting cancer cells and prevents interaction of nucleolin with the mRNA of the antiapoptotic Bcl-2. This causes destabilization of the Bcl-2 mRNA and subsequently a reduction of Bcl-2 protein which may cause apoptosis [114]. Phase I and II clinical trials have revealed the potency of AS1411 as new therapeutic agent of different cancer types, like relapsed or refractory acute myeloid leukemia or renal cell carcinoma [113].

To further improve AS1411 functions, Shieh et al. tested in 2010 whether AS1411 could deliver drugs to target cancer cells for photodynamic therapy (PDT) [103]. They created physical conjugates of AS1411 and TMPyP4 (5,10,15,20-tetrakis(1-methylpyridinium-4-yl)porphyrin). TMPyP4 is a photosensitizing porphyrin derivative that interacts with G-quadruplexes. Intracellularly, it serves as potent inhibitor of human telomerase [115]. The physical complex was escorted into MCF7 breast cancer cells and normal endothelial M10 cells via aptamer-mediated delivery and caused TMPyP4 accumulation in MCF7 cells but not in M10 cells. Subsequently, both cell types were exposed to light. The MCF7 cells were damaged to a significantly higher extent than the control cells. These results underscore the potential of AS1411 as a drug-delivery agent in cancer therapy.

6.3.4. Tyrosine-Protein Kinase-Like 7 Enzyme. Using cell-SELEX, Shangguan et al. selected numerous different aptamers that specifically recognized leukemia cells [54] (details in cell-SELEX chapter). For the cell-SELEX experiment the authors targeted cultured CCRF-CEM cells, a precursor cell line of T cell acute lymphoblastic leukemia cells (ALL). Ramos cells, a B cell line from human Burkitt's lymphoma, served for counter selection. Resulting aptamers specifically recognized CCRF-CEM cells with high affinity. Furthermore, target leukemia cells were identified when mixed with patients' human bone marrow aspirates. After identification

of PTK7 (Tyrosine-protein kinase-like 7) as target molecule of sgc8, Xiao et al. chose the aptamer sgc8c for further cell-specific internalization studies [116]. Several drug-like molecules were loaded on sgc8c, for instance, the MS2-coat protein [117] or the chemotherapeutic agents daunorubicin [118] and doxorubicin [104]. A conjugate of the latter with the aptamer sgc8 was particularly efficient in selectively killing specific cancer cells.

6.3.5. Transferrin Receptor (TfR). The enzyme replacement therapy (ERT) is a therapeutic application to replace an absent or deficient enzyme in patients by a functional counterpart. ERT is used for example in lysosomal storage diseases [119]. Receptors that underlie endocytosis are highly relevant for escort of these lysosomal enzymes into cells if coupled to appropriate delivery vehicles like aptamers. One of the mentioned receptors is TfR, ubiquitously presented on mammalian cells. It directs iron to cells via binding of transferrin-Fe³⁺, followed by cellular endocytosis and iron release into the endosomes [120]. Chen et al. used both, DNA and RNA aptamers, selectively binding to the mouse transferrin receptor (TfR) [105]. RNA aptamers for TfR were first coupled to streptavidin. The conjugates bound specifically to mouse fibroblasts (Ltk⁻ cells) but not to human 293T cells. Streptavidin could be detected in the lysosomes of Ltk⁻ cells presumably due to cellular uptake via endocytosis. After this proof of principle, a DNA aptamer was linked to α -L-iduronidase, a lysosomal enzyme [105]. Cells lacking this enzyme accumulate glycosaminoglycan. The aptamer-enzyme conjugates were internalized into mouse fibroblasts deficient for α -L-iduronidase and mediated by the TfR-specific DNA aptamer. The α -L-iduronidase was delivered into lysosomes. Inside the lysosome, the enzyme could counteract the glycosaminoglycan accumulation. The results reveal the potency of aptamers to escort lysosomal enzymes into target cells—a first step towards an aptamer-mediated enzyme replacement therapy for lysosomal storage diseases [105].

6.3.6. Mucin-1 (MUC1). Mucin-1 (MUC1) is a glycoprotein with considerable O-glycosylation within its extracellular domain. Membrane-associated glycoforms of MUC1 uniquely and abundantly line the cell surfaces of most epithelial cells, for instance, eyes, lungs and intestines. MUC1 represents an important surface biomarker for early diagnosis of cancers [121, 122]. MUC1 tumor markers undergo recycling processes, for instance, through endosomes. Therefore, these cell surface proteins may serve as entry portals for aptamers specifically binding to MUC1. Ferreira et al. reported recently the selection of MUC1 binding DNA aptamers using mimics of well-known MUC1 determinants on cancer cells as target molecules [123]. The obtained aptamers interacted specifically with cells presenting MUC1 glycosylation forms and were efficiently internalized by epithelial cancer cells. Coupling the aptamers to prodrugs like chlorin e₆ (Ce6), a heme-like agent for photodynamic therapy (PDT), selectively killed MUC1 presenting epithelial cancer cell lines [106].

6.3.7. Epidermal Growth Factor Receptor (EGFR). The epidermal growth factor receptor (EGFR) is a transmembrane

receptor with intrinsic tyrosine kinase activity. The EGFR belongs to a family of receptors known as the ErbB family (ErbB tyrosine kinase receptors). EGFR functions in a wide range of cellular processes, including cell proliferation and apoptosis [124]. It is overproduced in some types of cancer cells [125]. EGFR is activated upon binding to its ligands, for instance, EGF (epidermal growth factor) and TGF α (transforming growth factor). Subsequently, EGFR changes from an inactive monomeric form to an active homodimeric form leading to autophosphorylation and signal transduction. EGFR activation triggers internalization processes suggesting that receptor binding aptamers could become cointernalized, and thus serve as drug-delivery tools [126]. Li et al. reported in 2010 on the selection and application of aptamers that bind specifically to EGFR [107]. The authors generated conjugates of aptamers and gold nanoparticles that specifically targeted EGFR presenting A431 cells. The A431 cells internalized the aptamer-coated gold nanoparticles via EGFR-mediated endocytosis. Presumably, these aptamers are suitable for the delivery of further cargo molecules into EGFR presenting cells.

6.3.8. Immunglobulin M Heavy Chain. In 2007, Tang et al. reported on the cell-based systematic evolution of aptamers binding whole living cells [127]. Ramos cells, a B-cell Burkitt's lymphoma cell line, served as target in the cell-SELEX process. Resulting aptamers were able to identify Ramos cells within a complex biological mixture [127]. On the basis of this work, Mallikaratchy et al. identified membrane protein targets that were bound by the Ramos cell-specific aptamers [128]. By covalently cross-linking the Ramos cell-specific aptamer TD05 to its target cells, the authors could identify the molecular target of TD05 using mass spectrometry: the membrane bound IgM heavy chain. It is part of the B-cell receptor and presented in high amounts on the surface of Burkitt's lymphoma cells. Internalization of the aptamer, however, was not shown.

Nevertheless, converting these aptamers into delivery agents Wu et al. published a novel approach by modifying the end of an aptamer with a hydrophobic tail [108]. This resulted in the assembly of highly ordered micelle-like nanoparticles.

These aptamer-containing nanoparticles showed lower off rates after binding to the Ramos cells at 37°C, presumably due to multivalency. Furthermore, assembly of the aptamers to micelles allows them to enter their target cells by fusion of the micelle with the membrane of target cells. Loading the micelles with drugs could thus yield drug delivery agents. Additionally, the authors mimicked blood circulation of living organisms in a flow channel system. The aptamer-micelles were still able to recognize their target under flow conditions. These results might pave the way for further applications of aptamer-micelles as drug-delivery vehicles in a biological living system.

7. Conclusions

Since the invention of the SELEX process more than two decades ago, considerable progress has been made in the

field of aptamer technology. Several improvements, most importantly automation, have made the traditional SELEX procedure faster and more efficient and aptamers more easily available. Aptamers for all kinds of targets, ranging from small molecules to cell surface receptors have successfully been selected. Today, the selection of aptamers for whole cells is becoming increasingly popular, because their ability to specifically recognize subpopulations of cells is highly sought after. We anticipate that recent technical advances will further boost this trend. In particular, the use of FACS in the selection step represents a major technical progress, as it allows selecting aptamers and removing selection-hampering dead cells simultaneously.

The characteristics of aptamers also make them promising tools for diagnostic applications. Aptamer-based arrays, for example, could become a valuable alternative to antibody-based microarrays for protein production profiling of whole cells. These aptamer arrays might also be used to identify a specific cell type in complex cell mixtures. However, more time and scientific effort will still be required before aptamer-based products for diagnostics will be routinely applied.

Aptamers also have numerous properties that render them promising tools in therapy. As reviewed in this paper, there are aptamers for a large variety of cell-specific target molecules that play crucial roles in diseases, like viral infection, inflammation, and cancer. We have summarized examples from the recent literature illustrating aptamers that inhibit or activate cellular functions. Furthermore, we outlined how aptamers have been used to deliver therapeutic agents, such as chemotherapeutics, siRNAs or nanoparticles to cells or tissues. Aptamers are formidable delivery tools, because they are comparatively small and can easily penetrate tissues or become cointernalized with the receptors they are specifically binding to. Furthermore, aptamers can be tuned for desired applications: modified bases can be integrated to increase the *in vivo* stability or to regulate their *in vivo* function. The possibility to regulate and control the function of aptamers is crucial for a defined application in therapy. In this field, we expect further developments based on light regulation of aptamers as has been demonstrated by caged aptamers for human thrombin [129]. Another promising approach is the use of activatable [130] and deactivatable [129] aptamer probes (AAP) for the targeting of membrane proteins of living cancer cells [131].

Most of the diagnostically and therapeutically relevant aptamers mentioned in this paper have so far only been tested in animal models. There is only a limited number of clinical investigations of aptamers for therapeutic applications in humans. Some of these aptamers are binding specifically to defined proteins in cancer and eye diseases. However, even if there is only one FDA-approved aptamer-based drug so far (Macugen), some additional aptamers have already entered the phase of clinical trials. This is not very astonishing, because it often takes more than 10 years before a drug candidate becomes commercially available. The timeline for the development of aptamers as commercial therapeutics is comparable to the generation of antibody-based drugs. In 1975, the first monoclonal antibodies were

generated. In 1986, the first monoclonal antibody was approved by the FDA for therapeutic application the second one in 1994. The same holds true for aptamers—in 1990, the first aptamers were selected, in 2004, the first aptamer-based drug, Macugen, was approved by the FDA.

Finally, the selection of aptamers and their use in diagnostic and therapeutic applications is still an expanding area. Presumably, in the near future, further aptamers may become approved drugs.

References

- [1] A. D. Ellington and J. W. Szostak, "In vitro selection of RNA molecules that bind specific ligands," *Nature*, vol. 346, no. 6287, pp. 818–822, 1990.
- [2] C. Tuerk and L. Gold, "Systemic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase," *Science*, vol. 249, no. 4968, pp. 505–510, 1990.
- [3] D. L. Robertson and G. F. Joyce, "Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA," *Nature*, vol. 344, no. 6265, pp. 467–468, 1990.
- [4] L. A. Holeman, S. L. Robinson, J. W. Szostak, and C. Wilson, "Isolation and characterization of fluorophore-binding RNA aptamers," *Folding and Design*, vol. 3, no. 6, pp. 423–431, 1998.
- [5] K. Harada and A. D. Frankel, "Identification of two novel arginine binding DNAs," *EMBO Journal*, vol. 14, no. 23, pp. 5798–5811, 1995.
- [6] H. Schürer, K. Stempera, D. Knoll et al., "Aptamers that bind to the antibiotic moenomycin A," *Bioorganic and Medicinal Chemistry*, vol. 9, no. 10, pp. 2557–2563, 2001.
- [7] S. D. Mendonsa and M. T. Bowser, "In vitro selection of aptamers with affinity for neuropeptide Y using capillary electrophoresis," *Journal of the American Chemical Society*, vol. 127, no. 26, pp. 9382–9383, 2005.
- [8] S. E. Lupold, B. J. Hicke, Y. Lin, and D. S. Coffey, "Identification and characterization of nuclease-stabilized RNA molecules that bind human prostate cancer cells via the prostate-specific membrane antigen," *Cancer Research*, vol. 62, no. 14, pp. 4029–4033, 2002.
- [9] Z. Balogh, G. Lautner, V. Bardóczy, B. Komorowska, R. E. Gyurcsányi, and T. Mészáros, "Selection and versatile application of virus-specific aptamers," *FASEB Journal*, vol. 24, no. 11, pp. 4187–4195, 2010.
- [10] M. S. L. Raddatz, A. Dolf, E. Endl, P. Knolle, M. Famulok, and G. Mayer, "Enrichment of cell-targeting and population-specific aptamers by fluorescence-activated cell sorting," *Angewandte Chemie International Edition*, vol. 47, no. 28, pp. 5190–5193, 2008.
- [11] S. D. Jayasena, "Aptamers: an emerging class of molecules that rival antibodies in diagnostics," *Clinical Chemistry*, vol. 45, no. 9, pp. 1628–1650, 1999.
- [12] A. D. Keefe, S. Pai, and A. Ellington, "Aptamers as therapeutics," *Nature Reviews Drug Discovery*, vol. 9, no. 7, pp. 537–550, 2010.
- [13] P. R. Bouchard, R. M. Hutabarat, and K. M. Thompson, "Discovery and development of therapeutic aptamers," *Annual Review of Pharmacology and Toxicology*, vol. 50, pp. 237–257, 2010.
- [14] E. W. M. Ng, D. T. Shima, P. Calias, E. T. Cunningham, D. R. Guyer, and A. P. Adamis, "Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease," *Nature Reviews Drug Discovery*, vol. 5, no. 2, pp. 123–132, 2006.
- [15] G. R. Zimmermann, R. D. Jenison, C. L. Wick, J. P. Simorre, and A. Pardi, "Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA," *Nature Structural Biology*, vol. 4, no. 8, pp. 644–649, 1997.
- [16] P. L. Sazani, R. Larralde, and J. W. Szostak, "A small aptamer with strong and specific recognition of the triphosphate of ATP," *Journal of the American Chemical Society*, vol. 126, no. 27, pp. 8370–8371, 2004.
- [17] M. Famulok, J. S. Hartig, and G. Mayer, "Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy," *Chemical Reviews*, vol. 107, no. 9, pp. 3715–3743, 2007.
- [18] R. Yamamoto, M. Katahira, S. Nishikawa, T. Baba, K. Taira, and P. K. R. Kumar, "A novel RNA motif that binds efficiently and specifically to the Tat protein of HIV and inhibits the trans-activation by Tat of transcription in vitro and in vivo," *Genes to Cells*, vol. 5, no. 5, pp. 371–388, 2000.
- [19] L. Giver, D. Bartel, M. Zapp, A. Pawul, M. Green, and A. D. Ellington, "Selective optimization of the Rev-binding element of HIV-1," *Nucleic Acids Research*, vol. 21, no. 23, pp. 5509–5516, 1993.
- [20] K. A. Whitehead, R. Langer, and D. G. Anderson, "Knocking down barriers: advances in siRNA delivery," *Nature Reviews Drug Discovery*, vol. 8, no. 2, pp. 129–138, 2009.
- [21] R. Juliano, M. R. Alam, V. Dixit, and H. Kang, "Mechanisms and strategies for effective delivery of antisense and siRNA oligonucleotides," *Nucleic Acids Research*, vol. 36, no. 12, pp. 4158–4171, 2008.
- [22] M. G. Theis, A. Knorre, B. Kellersch et al., "Discriminatory aptamer reveals serum response element transcription regulated by cytohesin-2," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 31, pp. 11221–11226, 2004.
- [23] Y. Liu, C. T. Kuan, J. Mi et al., "Aptamers selected against the unglycosylated EGFRvIII ectodomain and delivered intracellularly reduce membrane-bound EGFRvIII and induce apoptosis," *Biological Chemistry*, vol. 390, no. 2, pp. 137–144, 2009.
- [24] L. Chaloin, M. J. Lehmann, G. Sczakiel, and T. Restle, "Endogenous expression of a high-affinity pseudoknot RNA aptamer suppresses replication of HIV-1," *Nucleic Acids Research*, vol. 30, no. 18, pp. 4001–4008, 2002.
- [25] K. H. Choi, M. W. Park, S. Y. Lee et al., "Intracellular expression of the T-cell factor-1 RNA aptamer as an intramer," *Molecular Cancer Therapeutics*, vol. 5, no. 9, pp. 2428–2434, 2006.
- [26] J. Mi, X. Zhang, Z. N. Rabbani et al., "H1 RNA polymerase III promoter-driven expression of an RNA aptamer leads to high-level inhibition of intracellular protein activity," *Nucleic Acids Research*, vol. 34, no. 12, pp. 3577–3584, 2006.
- [27] G. Mayer, M. Blind, W. Nagel et al., "Controlling small guanine-nucleotide-exchange factor function through cytoplasmic RNA intramers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 9, pp. 4961–4965, 2001.
- [28] M. Hafner, A. Schmitz, I. Grüne et al., "Inhibition of cytohesins by SecinH3 leads to hepatic insulin resistance," *Nature*, vol. 444, no. 7121, pp. 941–944, 2006.
- [29] A. Rhie, L. Kirby, N. Sayer et al., "Characterization of 2'-fluoro-RNA aptamers that bind preferentially to disease-associated conformations of prion protein and inhibit conversion," *Journal of Biological Chemistry*, vol. 278, no. 41, pp. 39697–39705, 2003.

- [30] N. C. Pagratis, C. Bell, Y. F. Chang et al., "Potent 2'-amino-, and 2'-fluoro-2'-deoxyribonucleotide RNA inhibitors of keratinocyte growth factor," *Nature Biotechnology*, vol. 15, no. 1, pp. 68–73, 1997.
- [31] J. Ruckman, L. S. Green, J. Beeson et al., "2'-fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165): inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain," *Journal of Biological Chemistry*, vol. 273, no. 32, pp. 20556–20567, 1998.
- [32] P. E. Burmeister, S. D. Lewis, R. F. Silva et al., "Direct in vitro selection of a 2'-O-methyl aptamer to VEGF," *Chemistry and Biology*, vol. 12, no. 1, pp. 25–33, 2005.
- [33] D. Shangguan, Z. Cao, L. Meng et al., "Cell-specific aptamer probes for membrane protein elucidation in cancer cells," *Journal of Proteome Research*, vol. 7, no. 5, pp. 2133–2139, 2008.
- [34] K. T. Guo, A. Paul, C. Schichor, G. Ziemer, and H. P. Wendel, "Cell-SELEX: novel perspectives of aptamer-based therapeutics," *International Journal of Molecular Sciences*, vol. 9, no. 4, pp. 668–678, 2008.
- [35] L. Cerchia and V. de Franciscis, "Targeting cancer cells with nucleic acid aptamers," *Trends in Biotechnology*, vol. 28, no. 10, pp. 517–525, 2010.
- [36] K. N. Morris, K. B. Jensen, C. M. Julin, M. Weil, and L. Gold, "High affinity ligands from in vitro selection: complex targets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 6, pp. 2902–2907, 1998.
- [37] D. O'Connell, A. Koenig, S. Jennings et al., "Calcium-dependent oligonucleotide antagonists specific for L-selectin," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 12, pp. 5883–5887, 1996.
- [38] B. J. Hicke, S. R. Watson, A. Koenig et al., "DNA aptamers block L-selectin function in vivo: inhibition of human lymphocyte trafficking in SCID mice," *Journal of Clinical Investigation*, vol. 98, no. 12, pp. 2688–2692, 1996.
- [39] C. H. B. Chen, G. A. Chernis, V. Q. Hoang, and R. Landgraf, "Inhibition of heregulin signaling by an aptamer that preferentially binds to the oligomeric form of human epidermal growth factor receptor-3," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 16, pp. 9226–9231, 2003.
- [40] C. M. Dollins, S. Nair, D. Boczkowski et al., "Assembling OX40 aptamers on a molecular scaffold to create a receptor-activating aptamer," *Chemistry and Biology*, vol. 15, no. 7, pp. 675–682, 2008.
- [41] A. P. Mann, A. Somasunderam, R. Nieves-Alicea et al., "Identification of thioaptamer ligand against E-selectin: potential application for inflamed vasculature targeting," *PLoS ONE*, vol. 5, no. 9, article e13050, pp. 1–11, 2010.
- [42] B. J. Hicke, C. Marion, Y. F. Chang et al., "Tenascin-C aptamers are generated using tumor cells and purified protein," *Journal of Biological Chemistry*, vol. 276, no. 52, pp. 48644–48654, 2001.
- [43] M. Homann and H. U. Göringer, "Combinatorial selection of high affinity RNA ligands to live African trypanosomes," *Nucleic Acids Research*, vol. 27, no. 9, pp. 2006–2014, 1999.
- [44] J. G. Bruno and J. L. Kiel, "In vitro selection of DNA aptamers to anthrax spores with electrochemiluminescence detection," *Biosensors and Bioelectronics*, vol. 14, no. 5, pp. 457–464, 1999.
- [45] M. Blank, T. Weinschenk, M. Priemer, and H. Schluesener, "Systematic evolution of a DNA aptamer binding to rat brain tumor microvessels: selective targeting of endothelial regulatory protein p19," *Journal of Biological Chemistry*, vol. 276, no. 19, pp. 16464–16468, 2001.
- [46] D. A. Daniels, H. Chen, B. J. Hicke, K. M. Swiderek, and L. Gold, "A tenascin-C aptamer identified by tumor cell SELEX: systematic evolution of ligands by exponential enrichment," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 15416–15421, 2003.
- [47] K. Sefah, D. Shangguan, X. Xiong, M. B. O'Donoghue, and W. Tan, "Development of DNA aptamers using Cell-SELEX," *Nature Protocols*, vol. 5, no. 6, pp. 1169–1185, 2010.
- [48] J. C. Cox, A. Hayhurst, J. Hesselberth, T. S. Bayer, G. Georgiou, and A. D. Ellington, "Automated selection of aptamers against protein targets translated in vitro: from gene to aptamer," *Nucleic Acids Research*, vol. 30, no. 20, p. e108, 2002.
- [49] R. K. Mosing and M. T. Bowser, "Isolating aptamers using capillary electrophoresis-SELEX (CE-SELEX)," *Methods in Molecular Biology*, vol. 535, pp. 33–43, 2009.
- [50] L. Cerchia, F. Ducongé, C. Pestourie et al., "Neutralizing aptamers from whole-cell SELEX inhibit the RET receptor tyrosine kinase," *PLoS Biology*, vol. 3, no. 4, article e123, 2005.
- [51] H. Zhao and F. H. Arnold, "Directed evolution converts subtilisin E into a functional equivalent of thermitase," *Protein Engineering*, vol. 12, no. 1, pp. 47–53, 1999.
- [52] M. Avci-Adali, M. Metzger, N. Perle, G. Ziemer, and H. P. Wendel, "Pitfalls of cell-systematic evolution of ligands by exponential enrichment (SELEX): existing dead cells during in vitro selection anticipate the enrichment of specific aptamers," *Oligonucleotides*, vol. 20, no. 6, pp. 317–323, 2010.
- [53] G. Mayer, M. S. L. Ahmed, A. Dolf, E. Endl, P. A. Knolle, and M. Famulok, "Fluorescence-activated cell sorting for aptamer SELEX with cell mixtures," *Nature Protocols*, vol. 5, no. 12, pp. 1993–2004, 2010.
- [54] D. Shangguan, Y. Li, Z. Tang et al., "Aptamers evolved from live cells as effective molecular probes for cancer study," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 32, pp. 11838–11843, 2006.
- [55] H. W. Chen, C. D. Medley, K. Sefah et al., "Molecular recognition of small-cell lung cancer cells using aptamers," *ChemMedChem*, vol. 3, no. 6, pp. 991–1001, 2008.
- [56] D. Shangguan, L. Meng, Z. C. Cao et al., "Identification of liver cancer-specific aptamers using whole live cells," *Analytical Chemistry*, vol. 80, no. 3, pp. 721–728, 2008.
- [57] K. Sefah, Z. W. Tang, D. H. Shangguan et al., "Molecular recognition of acute myeloid leukemia using aptamers," *Leukemia*, vol. 23, no. 2, pp. 235–244, 2009.
- [58] Z. Tang, P. Parekh, P. Turner, R. W. Moyer, and W. Tan, "Generating aptamers for recognition of virus-infected cells," *Clinical Chemistry*, vol. 55, no. 4, pp. 813–822, 2009.
- [59] C. Srisawat and D. R. Engelke, "Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins," *RNA*, vol. 7, no. 4, pp. 632–641, 2001.
- [60] T. S. Romig, C. Bell, and D. W. Drolet, "Aptamer affinity chromatography: combinatorial chemistry applied to protein purification," *Journal of Chromatography B*, vol. 731, no. 2, pp. 275–284, 1999.
- [61] J. K. Herr, J. E. Smith, C. D. Medley, D. Shangguan, and W. Tan, "Aptamer-conjugated nanoparticles for selective collection and detection of cancer cells," *Analytical Chemistry*, vol. 78, no. 9, pp. 2918–2924, 2006.
- [62] A. Rentmeister and M. Famulok, "Functional nucleic acid sensors as screening tools," in *Functional Nucleic Acids for*

- Analytical Applications*, L. Yingfu and L. Yi, Eds., p. 343, Springer, New York, NY, USA, 2003.
- [63] G. Liu, X. Mao, J. A. Phillips, H. Xu, W. Tan, and L. Zeng, "Aptamer-nanoparticle strip biosensor for sensitive detection of cancer cells," *Analytical Chemistry*, vol. 81, no. 24, pp. 10013–10018, 2009.
- [64] J. P. Overington, B. Al-Lazikani, and A. L. Hopkins, "How many drug targets are there?" *Nature Reviews Drug Discovery*, vol. 5, no. 12, pp. 993–996, 2006.
- [65] D. R. Gutsaeva, J. B. Parkerson, S. D. Yerigenahally et al., "Inhibition of cell adhesion by anti-P-selectin aptamer: a new potential therapeutic agent for sickle cell disease," *Blood*, vol. 117, no. 2, p. 727, 2011.
- [66] L. Chen, D. Q. Li, J. Zhong et al., "IL-17RA aptamer-mediated repression of IL-6 inhibits synovium inflammation in a murine model of osteoarthritis," *Osteoarthritis Cartilage*, vol. 19, no. 6, p. 711, 2011.
- [67] Y. Wang, Z. Z. Khaing, N. Li et al., "Aptamer antagonists of myelin-derived inhibitors promote axon growth," *PLoS ONE*, vol. 5, no. 3, article e9726, 2010.
- [68] S. Santulli-Marotto, S. K. Nair, C. Rusconi, B. Sullenger, and E. Gilboa, "Multivalent RNA aptamers that Inhibit CTLA-4 and enhance tumor immunity," *Cancer Research*, vol. 63, no. 21, pp. 7483–7489, 2003.
- [69] M. Khati, M. Schüman, J. Ibrahim, Q. Sattentau, S. Gordon, and W. James, "Neutralization of infectivity of diverse R5 clinical isolates of human immunodeficiency virus type 1 by gp120-binding 2'F-RNA aptamers," *Journal of Virology*, vol. 77, no. 23, pp. 12692–12698, 2003.
- [70] T. A. Springer, "Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm," *Cell*, vol. 76, no. 2, pp. 301–314, 1994.
- [71] S. R. Barthel, J. D. Gavino, L. Descheny, and C. J. Dimitroff, "Targeting selectins and selectin ligands in inflammation and cancer," *Expert Opinion on Therapeutic Targets*, vol. 11, no. 11, pp. 1473–1491, 2007.
- [72] M. C. Honorati, M. Bovara, L. Cattini, A. Piacentini, and A. Facchini, "Contribution of interleukin 17 to human cartilage degradation and synovial inflammation in osteoarthritis," *Osteoarthritis and Cartilage*, vol. 10, no. 10, pp. 799–807, 2002.
- [73] E. Zwick, J. Bange, and A. Ullrich, "Receptor tyrosine kinase signalling as a target for cancer intervention strategies," *Endocrine-Related Cancer*, vol. 8, no. 3, pp. 161–173, 2001.
- [74] H. S. Cho and D. J. Leahy, "Structure of the extracellular region of HER3 reveals an interdomain tether," *Science*, vol. 297, no. 5585, pp. 1330–1333, 2002.
- [75] G. Krähn, U. Leiter, P. Kaskel et al., "Coexpression patterns of EGFR, HER2, HER3 and HER4 in non-melanoma skin cancer," *European Journal of Cancer*, vol. 37, no. 2, pp. 251–259, 2001.
- [76] F. Hu, B. P. Liu, S. Budel et al., "Nogo-A interacts with the Nogo-66 receptor through multiple sites to create an isoform-selective subnanomolar agonist," *Journal of Neuroscience*, vol. 25, no. 22, pp. 5298–5304, 2005.
- [77] C. C. Stamper, Y. Zhang, J. F. Tobin et al., "Crystal structure of the B7-1/CTLA-4 complex that inhibits human immune responses," *Nature*, vol. 410, no. 6828, p. 608, 2001.
- [78] D. R. Leach, M. F. Krummel, and J. P. Allison, "Enhancement of antitumor immunity by CTLA-4 blockade," *Science*, vol. 271, no. 5256, pp. 1734–1736, 1996.
- [79] A. K. Dey, M. Khati, M. Tang, R. Wyatt, S. M. Lea, and W. James, "An aptamer that neutralizes R5 strains of human immunodeficiency virus type 1 blocks gp120-CCR5 interaction," *Journal of Virology*, vol. 79, no. 21, pp. 13806–13810, 2005.
- [80] A. K. Dey, C. Griffiths, S. M. Lea, and W. James, "Structural characterization of an anti-gp120 RNA aptamer that neutralizes R5 strains of HIV-1," *RNA*, vol. 11, no. 6, pp. 873–884, 2005.
- [81] J. O. McNamara, D. Kolonias, F. Pastor et al., "Multivalent 4-1BB binding aptamers costimulate CD8+ T cells and inhibit tumor growth in mice," *Journal of Clinical Investigation*, vol. 118, no. 1, pp. 376–386, 2008.
- [82] H. W. Lee, S. J. Park, B. K. Choi, H. H. Kim, K. O. Nam, and B. S. Kwon, "4-1BB promotes the survival of CD8+ T lymphocytes by increasing expression of Bcl-xL and Bfl-1," *Journal of Immunology*, vol. 169, no. 9, pp. 4882–4888, 2002.
- [83] I. Melero, W. W. Shuford, S. A. Newby et al., "Monoclonal antibodies against the 4-1BB T-cell activation molecule eradicate established tumors," *Nature Medicine*, vol. 3, no. 6, pp. 682–685, 1997.
- [84] K. F. May Jr., L. Chen, P. Zheng, and Y. Liu, "Anti-4-1BB monoclonal antibody enhances rejection of large tumor burden by promoting survival but not clonal expansion of tumor-specific CD8+ T cells," *Cancer Research*, vol. 62, no. 12, pp. 3459–3465, 2002.
- [85] A. Boltz, B. Piater, L. Toleikis et al., "Bi-specific aptamers mediating tumour cell lysis," *The Journal of Biological Chemistry*, vol. 268, no. 24, pp. 21896–21905, 2011.
- [86] J. P. Eder, G. F. Vande Woude, S. A. Boerner, and P. M. Lorusso, "Novel therapeutic inhibitors of the c-Met signaling pathway in cancer," *Clinical Cancer Research*, vol. 15, no. 7, pp. 2207–2214, 2009.
- [87] M. I. Davis, M. J. Bennett, L. M. Thomas, and P. J. Bjorkman, "Crystal structure of prostate-specific membrane antigen, a tumor marker and peptidase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 17, pp. 5981–5986, 2005.
- [88] U. Wullner, I. Neef, A. Eller, M. Kleines, M. K. Tur, and S. Barth, "Cell-specific induction of apoptosis by rationally designed bivalent aptamer-siRNA transcripts silencing eukaryotic elongation factor 2," *Current Cancer Drug Targets*, vol. 8, no. 7, pp. 554–565, 2008.
- [89] J. O. McNamara, E. R. Andrechek, Y. Wang et al., "Cell type-specific delivery of siRNAs with aptamer-siRNA chimeras," *Nature Biotechnology*, vol. 24, no. 8, pp. 1005–1015, 2006.
- [90] J. P. Dassie, X. Y. Liu, G. S. Thomas et al., "Systemic administration of optimized aptamer-siRNA chimeras promotes regression of PSMA-expressing tumors," *Nature Biotechnology*, vol. 27, no. 9, pp. 839–846, 2009.
- [91] T. C. Chu, K. Y. Twu, A. D. Ellington, and M. Levy, "Aptamer mediated siRNA delivery," *Nucleic Acids Research*, vol. 34, no. 10, article e73, 2006.
- [92] X. Ni, Y. Zhang, J. Ribas et al., "Prostate-targeted radiosensitization via aptamer-shRNA chimeras in human tumor xenografts," *The Journal of Clinical Investigation*, vol. 121, no. 6, p. 2383, 2011.
- [93] T. C. Chu, J. W. Marks, L. A. Lavery et al., "Aptamer:toxin conjugates that specifically target prostate tumor cells," *Cancer Research*, vol. 66, no. 12, pp. 5989–5992, 2006.
- [94] V. Bagalkot, O. C. Farokhzad, R. Langer, and S. Jon, "An aptamer-doxorubicin physical conjugate as a novel targeted drug-delivery platform," *Angewandte Chemie International Edition*, vol. 45, no. 48, pp. 8149–8152, 2006.

- [95] S. Dhar, N. Kolishetti, S. J. Lippard, and O. C. Farokhzad, "Targeted delivery of a cisplatin prodrug for safer and more effective prostate cancer therapy in vivo," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 5, p. 1850, 2011.
- [96] O. C. Farokhzad, J. Cheng, B. A. Teply et al., "Targeted nanoparticle-aptamer bioconjugates for cancer chemotherapy in vivo," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 16, pp. 6315–6320, 2006.
- [97] V. Bagalkot, L. Zhang, E. Levy-Nissenbaum et al., "Quantum dot-aptamer conjugates for synchronous cancer imaging, therapy, and sensing of drug delivery based on Bi-fluorescence resonance energy transfer," *Nano Letters*, vol. 7, no. 10, pp. 3065–3070, 2007.
- [98] A. Z. Wang, V. Bagalkot, C. C. Vasilidou et al., "Superparamagnetic iron oxide nanoparticle-aptamer bioconjugates for combined prostate cancer imaging and therapy," *ChemMedChem*, vol. 3, no. 9, pp. 1311–1315, 2008.
- [99] D. Kim, Y. Y. Jeong, and S. Jon, "A drug-loaded aptamer-gold nanoparticle bioconjugate for combined ct imaging and therapy of prostate cancer," *ACS Nano*, vol. 4, no. 7, pp. 3689–3696, 2010.
- [100] S. Dhar, F. X. Gu, R. Langer, O. C. Farokhzad, and S. J. Lippard, "Targeted delivery of cisplatin to prostate cancer cells by aptamer functionalized Pt(IV) prodrug-PLGA - PEG nanoparticles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 45, pp. 17356–17361, 2008.
- [101] J. Zhou, H. Li, S. Li, J. Zaia, and J. J. Rossi, "Novel dual inhibitory function aptamer-siRNA delivery system for HIV-1 therapy," *Molecular Therapy*, vol. 16, no. 8, pp. 1481–1489, 2008.
- [102] J. Zhou, P. Swiderski, H. Li et al., "Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells," *Nucleic Acids Research*, vol. 37, no. 9, pp. 3094–3109, 2009.
- [103] Y. A. Shieh, S. J. Yang, M. F. Wei, and M. J. Shieh, "Aptamer-based tumor-targeted drug delivery for photodynamic therapy," *ACS Nano*, vol. 4, no. 3, pp. 1433–1442, 2010.
- [104] Y. F. Huang, D. Shangguan, H. Liu et al., "Molecular assembly of an aptamer-drug conjugate for targeted drug delivery to tumor cells," *ChemBioChem*, vol. 10, no. 5, pp. 862–868, 2009.
- [105] C. H. B. Chen, K. R. Dellamaggiore, C. P. Ouellette et al., "Aptamer-based endocytosis of a lysosomal enzyme," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 41, pp. 15908–15913, 2008.
- [106] C. S. M. Ferreira, M. C. Cheung, S. Missailidis, S. Bisland, and J. Gariépy, "Phototoxic aptamers selectively enter and kill epithelial cancer cells," *Nucleic Acids Research*, vol. 37, no. 3, pp. 866–876, 2009.
- [107] N. Li, T. Larson, H. H. Nguyen et al., "Directed evolution of gold nanoparticle delivery to cells," *Chemical Communications*, vol. 46, no. 3, p. 392, 2010.
- [108] Y. Wu, K. Sefah, H. Liu, R. Wang, and W. Tan, "DNA aptamer-micelle as an efficient detection/delivery vehicle toward cancer cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 1, pp. 5–10, 2010.
- [109] H. Liu, P. Moy, S. Kim et al., "Monoclonal antibodies to the extracellular domain of prostate-specific membrane antigen also react with tumor vascular endothelium," *Cancer Research*, vol. 57, no. 17, pp. 3629–3634, 1997.
- [110] U. Moll, R. Lau, M. A. Sypes, M. M. Gupta, and C. W. Anderson, "DNA-PK, the DNA-activated protein kinase, is differentially expressed in normal and malignant human tissues," *Oncogene*, vol. 18, no. 20, pp. 3114–3126, 1999.
- [111] H. Ginisty, F. Amalric, and P. Bouvet, "Nucleolin functions in the first step of ribosomal RNA processing," *EMBO Journal*, vol. 17, no. 5, pp. 1476–1486, 1998.
- [112] D. Ishimaru, L. Zuraw, S. Ramalingam et al., "Mechanism of regulation of bcl-2 mRNA by nucleolin and A+U-rich element-binding factor 1 (AUF1)," *Journal of Biological Chemistry*, vol. 285, no. 35, pp. 27182–27191, 2010.
- [113] S. Soundararajan, L. Wang, V. Sridharan et al., "Plasma membrane nucleolin is a receptor for the anticancer aptamer AS1411 in MV4-11 leukemia cells," *Molecular Pharmacology*, vol. 76, no. 5, pp. 984–991, 2009.
- [114] S. Soundararajan, W. Chen, E. K. Spicer, N. Courtenay-Luck, and D. J. Fernandes, "The nucleolin targeting aptamer AS1411 destabilizes Bcl-2 messenger RNA in human breast cancer cells," *Cancer Research*, vol. 68, no. 7, pp. 2358–2365, 2008.
- [115] S. Y. Rha, E. Izbicka, R. Lawrence et al., "Effect of telomere and telomerase interactive agents on human tumor and normal cell lines," *Clinical Cancer Research*, vol. 6, no. 3, pp. 987–993, 2000.
- [116] Z. Xiao, D. Shangguan, Z. Cao, X. Fang, and W. Tan, "Cell-specific internalization study of an aptamer from whole cell selection," *Chemistry*, vol. 14, no. 6, pp. 1769–1775, 2008.
- [117] G. J. Tong, S. C. Hsiao, Z. M. Carrico, and M. B. Francis, "Viral capsid DNA aptamer conjugates as multivalent cell-targeting vehicles," *Journal of the American Chemical Society*, vol. 131, no. 31, pp. 11174–11178, 2009.
- [118] S. M. Taghdisi, K. Abnous, F. Mosaffa, and J. Behravan, "Targeted delivery of daunorubicin to T-cell acute lymphoblastic leukemia by aptamer," *Journal of Drug Targeting*, vol. 18, no. 4, pp. 277–281, 2010.
- [119] R. J. Desnick and E. H. Schuchman, "Enzyme replacement and enhancement therapies: lessons from lysosomal disorders," *Nature Reviews Genetics*, vol. 3, no. 12, pp. 954–966, 2002.
- [120] A. Dautry Varsat, A. Ciechanover, and H. F. Lodish, "pH and the recycling of transferrin during receptor-mediated endocytosis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 80, no. 8 I, pp. 2258–2262, 1983.
- [121] A. K. H. Cheng, H. Su, Y. A. Wang, and H. Z. Yu, "Aptamer-based detection of epithelial tumor marker mucin 1 with quantum dot-based fluorescence readout," *Analytical Chemistry*, vol. 81, no. 15, pp. 6130–6139, 2009.
- [122] M. A. Hollingsworth and B. J. Swanson, "Mucins in cancer: protection and control of the cell surface," *Nature Reviews Cancer*, vol. 4, no. 1, pp. 45–60, 2004.
- [123] C. S. M. Ferreira, C. S. Matthews, and S. Missailidis, "DNA aptamers that bind to MUC1 tumour marker: design and characterization of MUC1-binding single-stranded DNA aptamers," *Tumor Biology*, vol. 27, no. 6, pp. 289–301, 2006.
- [124] A. B. Singh and R. C. Harris, "Autocrine, paracrine and juxtacrine signaling by EGFR ligands," *Cellular Signalling*, vol. 17, no. 10, pp. 1183–1193, 2005.
- [125] M. Ono and M. Kuwano, "Molecular mechanisms of epidermal growth factor receptor (EGFR) activation and response to gefitinib and other EGFR-targeting drugs," *Clinical Cancer Research*, vol. 12, no. 24, pp. 7242–7251, 2006.
- [126] S. Sigismund, E. Argenzio, D. Tosoni, E. Cavallaro, S. Polo, and P. P. Di Fiore, "Clathrin-mediated internalization is

- essential for sustained EGFR signaling but dispensable for degradation,” *Developmental Cell*, vol. 15, no. 2, pp. 209–219, 2008.
- [127] Z. Tang, D. Shangguan, K. Wang et al., “Selection of aptamers for molecular recognition and characterization of cancer cells,” *Analytical Chemistry*, vol. 79, no. 13, pp. 4900–4907, 2007.
- [128] P. Mallikaratchy, Z. Tang, S. Kwame, L. Meng, D. Shangguan, and W. Tan, “Aptamer directly evolved from live cells recognizes membrane bound immunoglobulin heavy mu chain in Burkitt’s lymphoma cells,” *Molecular & Cellular Proteomics*, vol. 6, no. 12, pp. 2230–2238, 2007.
- [129] A. Heckel, M. C. R. Buff, M. S. L. Raddatz et al., “An anti-coagulant with light-triggered antidote activity,” *Angewandte Chemie International Edition*, vol. 45, no. 40, p. 6748, 2006.
- [130] A. Heckel and G. Mayer, “Light regulation of aptamer activity: an anti-thrombin aptamer with caged thymidine nucleobases,” *Journal of the American Chemical Society*, vol. 127, no. 3, pp. 822–823, 2005.
- [131] H. Shi, X. X. He, K. M. Wang et al., “Activatable aptamer probe for contrast-enhanced in vivo cancer imaging based on cell membrane protein-triggered conformation alteration,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 10, p. 3900, 2011.

Research Article

A ssDNA Aptamer That Blocks the Function of the Anti-FLAG M2 Antibody

Amanda S. Lakamp¹ and Michel M. Ouellette^{1,2}

¹*Eppley Institute for Research in Cancer, University of Nebraska Medical Center, 985950 Nebraska Medical Center, Omaha, NE 68198-5950, USA*

²*Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, 985950 Nebraska Medical Center, Omaha, NE 68198-5950, USA*

Correspondence should be addressed to Michel M. Ouellette, mouellet@unmc.edu

Received 21 May 2011; Revised 22 July 2011; Accepted 25 July 2011

Academic Editor: Luis A. Marky

Copyright © 2011 A. S. Lakamp and M. M. Ouellette. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Using SELEX (systematic evolution of ligands by exponential enrichment), we serendipitously discovered a ssDNA aptamer that binds selectively to the anti-FLAG M2 antibody. The aptamer consisted of two motifs (CCTTA and TGTCTWCC) separated by 2-3 bases, and the elimination of one or the other motif abrogated binding. The DNA aptamer and FLAG peptide competed for binding to the antigen-binding pocket of the M2 antibody. In addition, the aptamer eluted FLAG-tagged proteins from the antibody, suggesting a commercial application in protein purification. These findings demonstrate the feasibility of using SELEX to develop ssDNA aptamers that block the function of a specific antibody, a capability that could lead to the development of novel therapeutic modalities for patients with systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases.

1. Introduction

Antinuclear antibodies are diagnostic markers of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases [1]. In these B lymphocyte disorders, a large variety of autoantibodies are made against nuclear self-antigens, including ribonucleoproteins, nucleosomes, chromatin, and polynucleotides (RNA, ssDNA, and dsDNA). Among these, anti-DNA antibodies have been the most extensively studied [2]. Anti-DNA antibodies bind with high-affinity to either single- or double-stranded DNA and many tend to favor association with pyrimidine bases [3, 4]. Several reports have also described antinuclear antibodies cross-reacting with peptide self-antigens and depositing in the brain, kidneys, and skin [5–9]. As proposed by several investigators, this deposition may be a cause of inflammation-mediated tissue damage, especially in the kidneys where nephritis is a major source of morbidity [1, 2]. In mouse models of systemic lupus erythematosus, attempts were made to block the function of these cross-reacting antibodies using peptide aptamers, derived either from their cognate

peptide self-antigens or from phage display libraries [10, 11]. In some cases, the peptide aptamer competitively associated with the antinuclear autoantibodies, thereby preventing antibody-mediated tissue damage [10, 11]. Thus, direct antibody inhibition might be an effective therapy in patients with autoimmune diseases driven by the presence of antinuclear antibodies. Another viable approach to block antinuclear antibodies might be to use DNA aptamers, given the high-affinity of these antibodies for DNA and evidence of nucleotide base specificity. But this approach has clearly been underexplored, perhaps due to the lack of reports on the feasibility of developing DNA aptamers to block the function of specific antibodies.

An adaptive technique employed to define the sequence specificity of DNA/RNA-binding proteins is SELEX (systematic evolution of ligands by exponential enrichment). In SELEX, the protein of interest is used as a selection matrix to capture high-affinity DNA binding sites from a pool of randomized DNA molecules [12, 13]. This pool is comprised of an oligonucleotide that contains a randomized core (up to 35 bases in size) flanked by PCR priming sequences.

The randomized core is made during chemical synthesis using a mixture of all four nucleoside phosphoramidites at each of the random positions. Following their capture, the selected DNA molecules are reamplified by PCR and then further enriched through successive rounds of selection. After 4–6 rounds, the selected DNA molecules are cloned and sequenced to identify any common DNA motifs recognized by the protein of interest. SELEX can be applied to the selection of ssDNA, dsDNA, or even RNA molecules [12, 13]. It is a powerful tool that has been used to optimize nucleic acid ligands for a multitude of proteins, even some which do not normally interact with DNA or RNA. As an example, SELEX was utilized to develop RNA aptamers that bind to blood coagulation factors, including thrombin [14], Von Willebrand factor [15], and Factor IXa [16]. In all three cases, the selected RNA aptamers interacted selectively with their corresponding protein targets and, in the process, inhibited their blood coagulation activities. A second generation of aptamers was developed, and, among these, some have entered clinical trials in patients with blood coagulation disorders [15].

Using SELEX, we serendipitously discovered a ssDNA sequence that binds selectively to the M2 antibody, a commonly used reagent that recognizes the Flag epitope (DYKDDDDK). The DNA aptamer and Flag peptide competed for binding to the M2 antibody, thereby allowing the aptamer to elute Flag-tagged proteins from an immobilized M2 antibody, a commonly employed procedure in protein purification. Aside from this immediate application in protein purification, identification of this DNA aptamer demonstrates the feasibility of using SELEX to develop aptamers that block specific antibodies. Applying this approach to antinuclear autoantibodies could lead to the development of novel therapeutic strategies for patients with systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases.

2. Materials and Methods

2.1. Materials. Oligonucleotides were synthesized by the Eppley Core Facility (University of Nebraska Medical Center, Omaha, NE). Plasmid pTetFLAGhTRF2^{45–501} was a gift from Dr. Titia de Lange (Rockefeller University, New York, NY) [17]. The polynucleotide kinase and the Platinum Pfx and Taq DNA polymerases were purchased from Invitrogen (Carlsbad, CA). All other enzymes were obtained from Fermentas (Hanover, MD), New England BioLabs (Beverly, MA), Promega (Madison, WI), or Invitrogen (Carlsbad, CA). The TnT Quick coupled Transcription/Translation System was purchased from Promega (Madison, WI). The γ -[³²P]-ATP (4500 Ci/mmol) was purchased from MP Biologicals (Solon, OH), and the L-[³⁵S]-Methionine (1000 Ci/mmol) was obtained from PerkinElmer (Boston, MA). M-450 magnetic beads coated with a sheep anti-mouse IgG antibody were received from DYNAL Biotech, Inc., (Lake Success, NY). 3XFLAG peptide and all other chemicals were from Sigma-Aldrich (St. Louis, MO). The Mini-PROTEAN 4%–20% SDS-PAGE gels were purchased from Bio-Rad (Hercules, CA).

2.2. Antibodies. Mouse monoclonal antibodies against the Flag tag (IgG₁ clone M2) and Stn1/OBFC1 (IgG_{2ak} clone 3G12-1B7) were purchased from Sigma-Aldrich (St. Louis, MO). Normal mouse IgG (cat. # sc-2025) and anti-vimentin mouse monoclonal antibody (IgG₁ clone sc-6260) were obtained from Santa Cruz (Santa Cruz, CA), as was the rabbit polyclonal antibody against TRF2 (cat. # H-300). Also purchased were mouse monoclonal antibodies against PTOP (IgG₁ clone 1D8-1B6; Novus Biologicals, Littleton, CO) and TIN2 (IgG₁ clone 59B388, AbCam, Cambridge, MA).

2.3. Expression Vectors. The plasmid pcDNA3.1-Flag-Stn1 was made by insertion of the human Stn1 sequences into plasmid pcDNA3.1-Flag. The coding sequence of Stn1 was amplified from plasmid pOTB7-Stn1 (GenBank # BC017400; American Type Culture Collection, Manassas, VA) using the Platinum Pfx DNA polymerase and primers 5'-GACTGACAATTGGGTGGTATGCAGCCTGGATCCAG-CC-3' and 5'-GACTGAAGATCTTCAGAACGCTGTGTAGTAGTG-3'. The PCR product was then cut with MfeI/BglII and inserted into the EcoRI/BamHI sites of pcDNA3.1-Flag, in frame with the Flag tag. pcDNA3.1-Flag is a pcDNA3.1(-) vector encoding a Flag epitope located downstream of a T7 promoter and immediately followed by an EcoRI site. Plasmid pcDNA3.1-Flag-TRF2^{ΔB} was made from its pCMV1 equivalents by transfer of its TRF2 cassette to vector pcDNA3.1(-) (Invitrogen, Carlsbad, CA). pCMV1-Flag-TRF2^{ΔB} was in turn made by the transfer to pCMV1 of a SacII/BamHI fragment from plasmid pTetFLAGhTRF2^{45–501} (a gift from Titia de Lange, Rockefeller University, NY) [17].

2.4. Protein Expression. The Flag-Stn1 and [³⁵S]-Flag-TRF2^{ΔB} proteins were made by *in vitro* transcription/translation in a rabbit reticulocyte lysate. In a final volume of 50 μ L, one microgram of pcDNA3.1-Flag-Stn1 was transcribed/translated using the TnT Quick Coupled system, according to the manufacturer's instructions (Promega, Madison, WI). A water-programmed lysate (mock) was produced in parallel to serve as a negative control. After translation, aliquots of the two reactions were analyzed by western blotting using both the Stn1/OBFC1 antibody and anti-Flag M2 antibody. A single species of 45 kDa was detected by the two antibodies in the Stn1-programmed lysate but not the Mock lysate (See Figure S1 in Supplementary Material available online at doi: 10.4061/2011/720798). The [³⁵S]-labeled Flag-TRF2^{ΔB} protein was produced similarly with the exception that the unlabeled methionine was replaced with 2 μ L of L-[³⁵S]-methionine (1000 Ci/mmol, 10 μ Ci/ μ L).

2.5. Preparation of Beads Coated with the Anti-Flag M2 Antibody. M2 antibody-coated beads were prepared by mixing 200 μ L of M-450 magnetic beads coated with a sheep anti-mouse IgG antibody (DYNAL Biotech, Inc., Lake Success, NY) with 15 μ g of anti-Flag antibody (M2 mouse monoclonal, Stratagene, La Jolla, CA) in 5 mL of PBS containing 0.1% each of NP-40 and BSA. After overnight rotation at 4°C, M2-coated beads were washed 3 times in PBS containing 0.1% each of NP-40 and BSA, after which beads were suspended and stored in 200 μ L of the same buffer. Before use, beads

were washed 3 times with ice-cold 1X binding buffer containing 0.1% BSA. Control beads coated with normal mouse IgG (cat. # sc-2025, Santa Cruz) were prepared following the same exact protocol.

2.6. SELEX. An oligonucleotide containing a random core of 35 nucleotides flanked by PCR priming sequences was made: 5'-GCGTCGACAAGCTTTCTAGA(N)₃₅GAATTC-GGATCCCTCGAGCG-3'. In the first round of selection, 5 μ g of this randomized oligo (215 pmoles, 130 trillion molecules) was incubated with 12.5 μ L of rabbit reticulocyte lysate (programmed with either water or Flag-Stn1) and 5 μ g of sonicated denatured *E. coli* genomic DNA in a 50 μ L reaction containing 1X binding buffer (4% glycerol, 1 mM MgCl₂, 1 mM DTT, 50 mM NaCl, 10 mM Tris-HCl pH 7.5). After 30 minutes at room temperature, 10 μ L of M2-coated beads were added and the samples were rotated at room temperature for 1 hour. Magnetic beads were washed 3 times with ice-cold 1X binding buffer containing 0.1% BSA, after which the selected oligos were eluted for 10 minutes at room temperature in the presence of an excess of 3XFLAG peptide (34 μ M in 1X binding buffer). Next, the eluted DNA was reamplified by PCR using Platinum Taq DNA polymerase and primers 5'-GCGTCGACAAGCTTTCTAGA-3' (forward) and 5'-CGCTCGAGGGATCCGAATTC-3' (reverse). Aliquots taken after 10, 15, 20, and 25 cycles of PCR were resolved on a 3% agarose gel. The most optimally amplified aliquot (no smear, no supershift, within the exponential range) was selected, cut, and gel purified using the GENECLEAN III kit (MP Biologicals, Solon, OH). Next, the gel-purified products were subjected to asymmetric PCR to regenerate ssDNA molecules needed for the following rounds of SELEX. Sixteen cycles of asymmetric PCR were performed using Platinum Taq DNA polymerase and the forward primer, after which the PCR products were extracted with phenol:chloroform (1:1), chloroform only, and then ethanol precipitated. The ssDNA pellet was dissolved in water and was ready to be used in the next round of SELEX. Additional round of SELEX were done identically, except that the randomized oligo was replaced with the previously selected and reamplified ssDNA.

After the sixth round of selection, the symmetrically amplified PCR product was gel purified and then prepared for TA-cloning into vector pCR2.1-Topo (Invitrogen, Carlsbad, CA). Chemically transformed TOP10 *E. coli* cells were spread onto plates containing kanamycin (25 μ g/mL). A total of 50 white colonies were picked and sent for sequencing. Over 30 sequences were obtained for each of the two SELEX procedures performed (mock versus Flag-Stn1).

2.7. Radiolabeling of DNA Probes. In 20 μ L of forward reaction buffer (10 mM MgCl₂, 5 mM DTT, and 70 mM Tris-HCl, pH 7.6), 30 pmoles of oligonucleotide were labeled for 15 minutes at 37°C with 75 μ Ci of γ -[³²P]-ATP (4500 Ci/mole) and 10 units of T4 polynucleotide kinase (Invitrogen, Carlsbad, CA). Next, the probes were resolved by electrophoresis on an 8% polyacrylamide gel and were gel purified. Lastly, the eluted probes were desalted on a G50 spun column (GE Healthcare, Piscataway, NJ).

2.8. Electrophoretic Mobility Shift Assay (EMSA). In 25 μ L of 1X binding buffer, reactions contained 1 μ g of the indicated antibody and 80,000 cpm of the designated [³²P]-labeled oligonucleotide. Reactions were incubated at room temperature for 30 minutes, and were then loaded onto a 4% native polyacrylamide gel containing TBE buffer (45 mM Tris-borate, 2 mM EDTA, pH 8.3). Gels were run at 4°C for 1 hour at 180 Volts. Gels were then transferred to a DE81 anion exchange chromatography paper (Whatman International, Maidstone, England), dried, and exposed to a PhosphorImager screen (Molecular Dynamics, Sunnyvale, CA). In competition experiments, the indicated competitors were added 5 minutes prior to the addition of the [³²P]-labeled ABA probe. The competitors used included an increasing concentration of 3XFLAG peptide (2, 5, 10, 20, 50, 100, 200, or 500 nM) or an excess of either the ABA or CTR oligo (500 nM each).

2.9. Binding Isotherm of the ABA Probe. Experiments were performed with [³²P]-labeled ABA as described above, except that the M2 antibody was added at an increasing concentration (0, 50, 100, 200, 350, 500, 750, 1000, 2000, and 4000 ng). After exposure of the gel to a PhosphorImager screen, the amount of free [³²P]-ABA and amount of [³²P]-ABA bound to the M2 antibody were quantified by volume integration using ImageQuant program (Molecular Dynamics, Sunnyvale, CA). The amount of [³²P]-ABA bound, expressed as a fraction of the total, was plotted as a function of total amount of M2 antibody. The resulting curves were fitted by nonlinear regression to a "one-site" saturation binding curve ($[^{32}\text{P}]\text{-ABA bound} = [\text{M2}]/(K_D + [\text{M2}])$). Fitting, performed using SigmaPlot version 11.0, allowed calculation of the apparent dissociation constant (K_D).

2.10. Release of the Captured [³²P]-ABA Oligo by the 3XFLAG Peptide. The [³²P]-labeled ABA oligo was first immobilized onto the M2-coated beads. Briefly, 150 μ L of M2-coated beads were added to 375 μ L of 1X binding buffer containing 600,000 cpm of [³²P]-ABA oligo. Capture of the oligo was allowed to proceed for one hour at room temperature, after which the beads were washed 3 times with 500 μ L of ice-cold 1X binding buffer (containing 0.1% BSA). The radioactive beads were then resuspended in 450 μ L of the same buffer. To assess the ability of the 3XFLAG peptide to elute the captured oligo, an increasing concentration of 3XFLAG peptide (0, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, and 5000 nM) was added to 30 μ L of beads carrying the [³²P]-ABA oligo. After 20 minutes at room temperature, the beads were pulled down and the amounts of radioactivity remaining on the beads and in the supernatant were counted by scintillation and expressed as a fraction of the total.

2.11. Inhibition of the Capture of [³⁵S]-Labeled Flag-TRF2^{ΔB}. Prior to these experiments, beads were blocked at 4°C for 30 minutes with 5% milk in 1X binding buffer to reduce nonspecific binding. In 1X binding buffer, M2-coated beads were first incubated with the indicated competitor (30 μ M each 3XFLAG, 3XABA, or 3XCTR) or with no competitor (No Comp). After 10 minutes at room temperature, 2–10 μ L

of [^{35}S]-labeled FLAG-TRF2 $^{\Delta\text{B}}$ were added. After 1 hour of rotation at room temperature, each reaction was washed 3 times with 500 μL of ice-cold 1X binding buffer (containing 0.1% BSA), after which the amount of [^{35}S]-Flag-TRF2 $^{\Delta\text{B}}$ captured was quantified. In one experiment, the amount captured was quantified by SDS-PAGE electrophoresis and exposure to a PhosphorImager screen. In a second experiment done in triplicate, the amount of [^{35}S]-Flag-TRF2 $^{\Delta\text{B}}$ captured was quantified by scintillation counting. In both experiments, beads coated with normal mouse IgG were included as a negative control for the capture (IgG).

2.12. Release of Precaptured [^{35}S]-Labeled Flag-TRF2 $^{\Delta\text{B}}$. The [^{35}S]-labeled Flag-TRF2 $^{\Delta\text{B}}$ protein was first immobilized onto the M2-coated beads. Briefly, 80 μL of M2-coated beads were added to 800 μL of 1X binding buffer containing 32 μL of [^{35}S]-Flag-TRF2 $^{\Delta\text{B}}$. Capture of the protein was allowed to proceed for one hour at room temperature, after which the beads were washed 3 times with 500 μL of ice-cold 1X binding buffer (containing 0.1% BSA). The radioactive beads were then resuspended in 80 μL of the same buffer. To assess the ability of the 3XABA oligo to elute the captured protein, five microliters of these beads were mixed with 45 μL of 1X binding buffer containing the indicated competitor (30 μM each 3XFLAG, 3XABA, or 3XCTR) or no competitor (No Comp). After 20 minutes at room temperature, the beads were pulled down and the amount of radioactivity in the supernatants was quantified by SDS-PAGE electrophoresis and exposure to a PhosphorImager screen. In an experiment done in triplicate, the amount of [^{35}S]-Flag-TRF2 $^{\Delta\text{B}}$ released was also quantified by scintillation counting. In both experiments, beads boiled to release all of the captured [^{35}S]-labeled protein were included as positive control for the elution (Total).

3. Results

3.1. SELEX Driven by the Anti-Flag M2 Antibody. SELEX was used to characterize the DNA-binding specificity of a human Flag-tagged Stn1 protein [18, 19], in this case produced by *in vitro* translation in a rabbit reticulocyte cell-free system. The *in vitro* translated Flag-Stn1 was mixed with a pool of randomized ssDNA molecules and then, the anti-Flag M2 antibody was used to immunoprecipitate the Flag-Stn1/DNA complexes to select for high-affinity targets. After six successive rounds of selection, the selected ssDNA molecules were cloned and sequenced. Analysis of a related unselected random pool revealed a slight bias for G-rich sequences, but no evidence of any recurrent motifs [20]. In contrast, analysis of the Flag-Stn1 selected oligonucleotides revealed a bipartite consensus motif made of two distinct elements: CCTTA and TGTCTWCC (where W = A/T; Figure 1(a)). These elements were separated by 2-3 bases of poorly conserved sequences. However, the same bipartite consensus motif was also produced when SELEX was performed with the same antibody but using a mock-programmed reticulocyte lysate as a control source of proteins (Figure 1(b)). Based on these results, we concluded that the bipartite consensus had been selected, not by the

Flag-Stn1 protein, but by the M2 antibody itself or proteins of the lysate that the antibody cross-reacted with.

3.2. The M2 Antibody Binds the Consensus ssDNA Motifs. To test the possibility of direct interactions between the bipartite consensus and M2 antibody, EMSA was performed. A series of equal length [^{32}P]-labeled ssDNA probes containing or lacking each element of the consensus were made (Figure 2(a)). Probes AB, BA, and ABA carried both the CCTTA and TGTCTWCC motifs, whereas probe 2XB and 4XA consisted of multiple copies of one or the other element. The unrelated CTR probe was used as negative control. Figure 2(b) shows that the M2 antibody could form a stable complex with the ABA probe and to a lesser extent, the AB probe. None of the other probes interacted with the M2 antibody. Also noteworthy, neither the BA probe lacking the CCTTA element nor the AA probe lacking the TGTCTWCC element interacted with the antibody, thereby indicating that both elements are essential for binding. Figure 2(b) also shows that among 7 different antibodies, the M2 antibody only interacted with probe ABA. These results show that probe ABA binds very selectively to the M2 antibody, thereby suggesting that these interactions are involving the variable regions of the antibody, whether part of the light chain, heavy chain, or both.

To measure the affinity of the M2 antibody for probe ABA, the apparent dissociation constant (K_D) of the complex was determined. A binding isotherm was generated by EMSA using a constant amount of [^{32}P]-ABA incubated with an increasing concentration of M2 antibody (Figure 2(c)). Plotting the amount of ABA bound as a function of the total concentration of antibody generated a binding isotherm. Fitting of this isotherm to a "one site" saturation binding curve allowed for calculation of the K_D . In this case, the apparent K_D was determined to have a value of 80 ± 7 nM (Figure 2(d)). The value is comparable to the apparent K_D of other previously reported DNA aptamer/antibody complexes [21]. This result shows that the M2 antibody binds with high affinity to its preferred DNA ligand, the ABA probe.

3.3. The Flag Peptide Competes with the Binding of the ABA Oligo to the M2 Antibody. The M2 antibody binds to both the Flag peptide (DYKDDDDK) and the ABA probe, so we sought to determine whether the two ligands could compete for each other. In a first series of experiments, we asked if an excess of 3XFLAG peptide (MDYKDHGDYKDH-DIDYKDDDDK) could block the binding of [^{32}P]-labeled ABA to the M2 antibody. In the EMSA shown in Figure 3(a), an increasing concentration of 3XFLAG peptide was incubated with the M2 antibody prior to adding the [^{32}P]-ABA. At concentrations of 50 nM and higher, the 3XFLAG peptide could completely block the binding of ABA to the M2 antibody. A similar inhibition was produced following the inclusion of an excess of unlabeled ABA oligo, but not after the addition of the CTR oligo. These results show that the 3XFLAG peptide competes with the binding of probe ABA to the M2 antibody.

In a second series of experiments, we asked whether an excess of 3XFLAG peptide could elute a [^{32}P]-ABA oligo

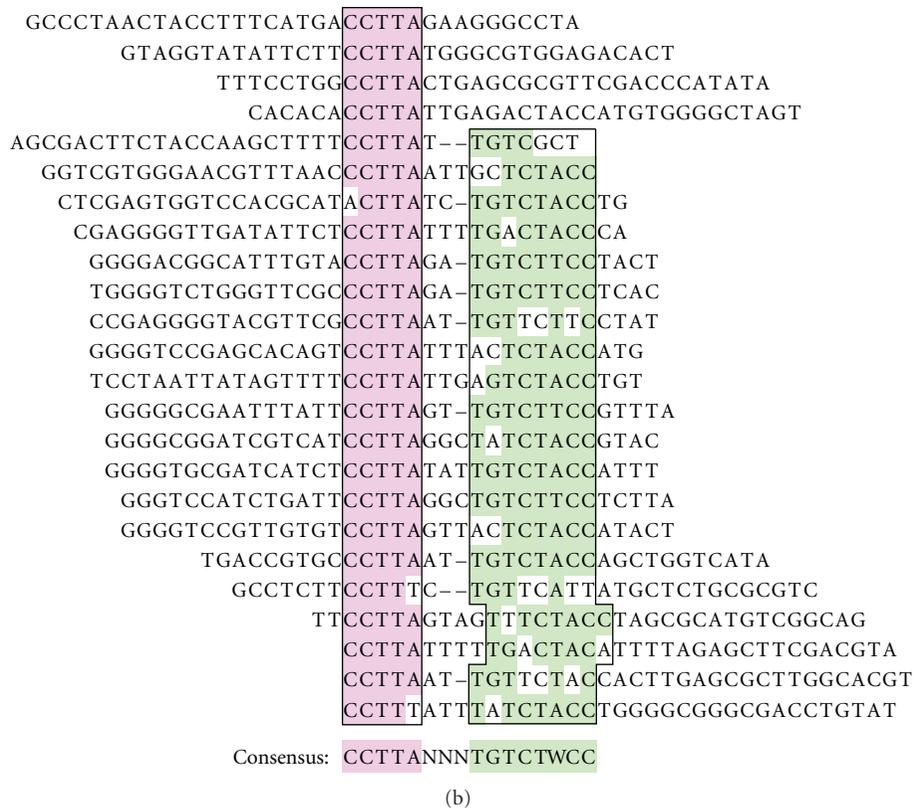
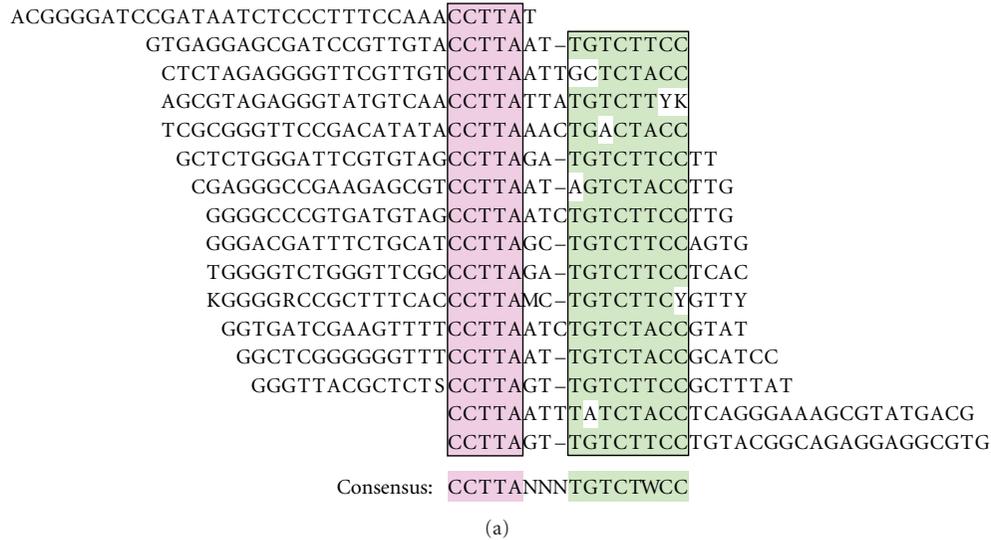


FIGURE 1: SELEX with or without the Flag-Stn1 protein yields the same bipartite consensus. A randomized pool of ssDNA molecules was incubated with an *in vitro* translation system programmed with either a Flag-Stn1 plasmid (Flag-Stn1 present) or water (Flag-Stn1 absent), after which protein/DNA complexes were recovered by immunoprecipitation using the anti-Flag M2 antibody. After six successive rounds of selection, the remaining oligonucleotides were cloned and sequenced. (a) Oligonucleotides selected in the presence of the Flag-Stn1 protein. The selected ssDNA molecules share a common bipartite consensus, which consists of a CCTTA (pink) and TGTCTWCC (green) motifs. Nonnucleotide letters denote ambiguous sequencing reads (R = A/G, Y = C/T, S = C/G, W = A/T, K = T/G, M = A/C). (b) Oligonucleotides selected in the absence of the Flag-Stn1 protein. The same bipartite consensus is shared among the ssDNA molecules selected by the mock-programmed *in vitro* translation system.

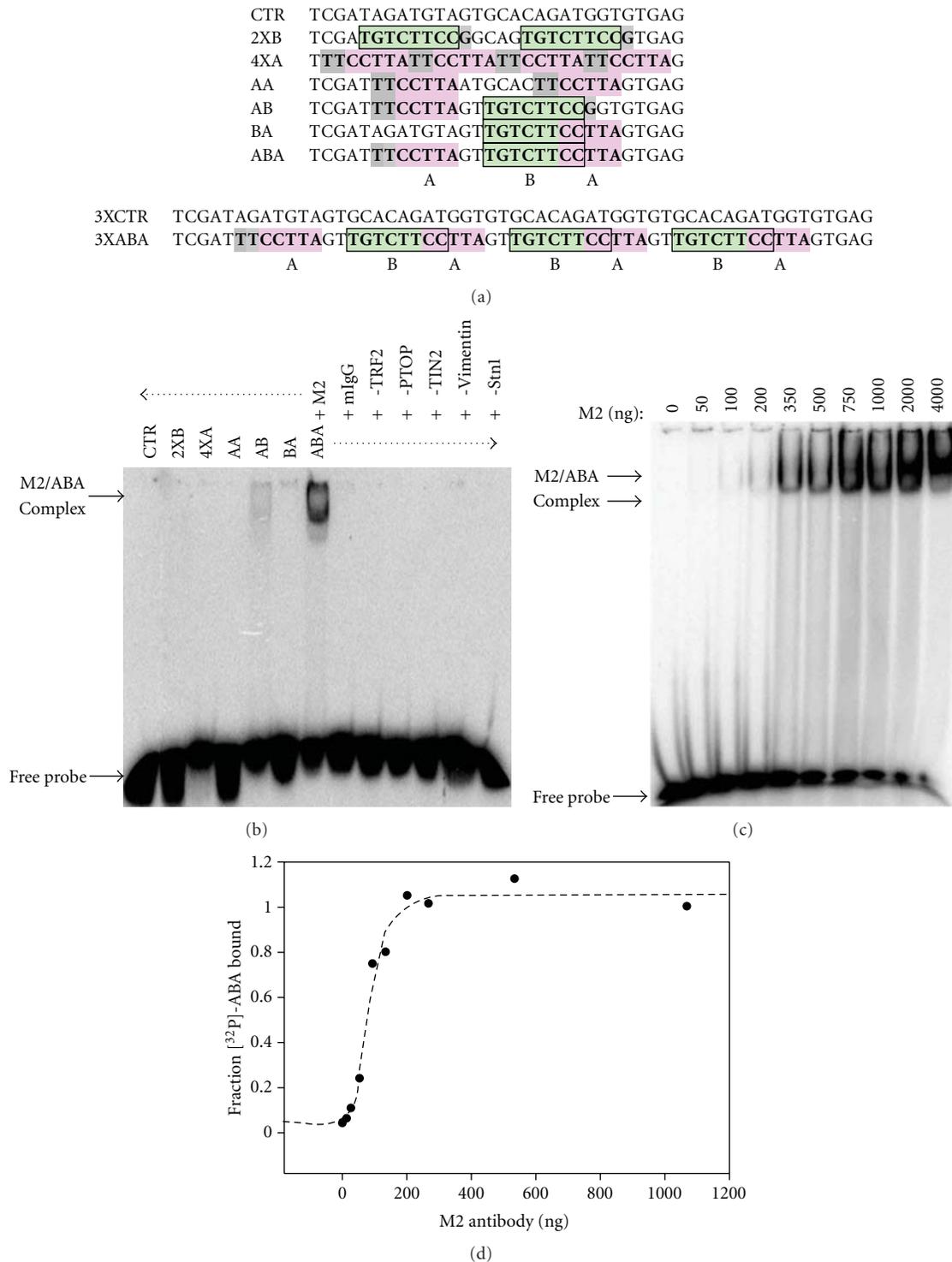


FIGURE 2: The bipartite ssDNA consensus binds directly to the anti-Flag M2 antibody. (a) Sequence and graphical representation of the probes used. All probes were labeled at the 5'-end with [32 P]. Probes were designed to carry none, one, or both of the identified motifs. Motif A = CCTTA (pink). Motif B = TGTCTWCC (green). (b) The anti-Flag M2 antibody binds selectively to probe ABA. The indicated probes (80,000 cpm) were incubated with the listed antibodies (1 μ g), after which protein/DNA complexes were resolved by electrophoresis in a native polyacrylamide gel. (c) Titration of the M2 antibody. A constant amount of probe ABA was incubated with an increasing concentration of M2 antibody, and the protein/DNA complexes were resolved by native electrophoresis. At the higher antibody concentrations, a larger protein/DNA complex is observed (top arrow), which may represent the product of antibody oligomerization. (d) Binding isotherm of probe ABA interacting with the anti-Flag M2 antibody. The scatter plot shows the amount of ABA bound (amount present in both protein/DNA complexes) as a function of antibody concentration. Dotted line shows fitting of the data to a "one site" saturation binding curve. Nonlinear regression of the data allowed for the determination of the dissociation constant ($K_D = 80 \pm 7$ nM). Error on the value of the K_D represents the 95% confidence interval of the best-fit curve.

already bound to the M2 antibody. For this purpose, [^{32}P]-ABA was first captured by magnetic beads coated with the M2 antibody, after which an increasing concentration of 3XFLAG peptide was added. In Figure 3(b), the amount of ABA bound to the beads and the amount released from the beads were plotted as a function of the concentration of the 3XFLAG peptide. The data show that the ABA oligo was very efficiently released by the addition of the 3XFLAG peptide. A small fraction of oligo, representing 36%, could not be displaced, but the other fraction was almost completely released by 2000 nM of 3XFLAG peptide. Fitting the data to a “one site” competition curve allowed for the calculation of an IC_{50} . In agreement with the values of the K_D , the IC_{50} was determined to be 96 nM (95% confidence interval: 72–127 nM). These results show that the 3XFLAG peptide can elute the ABA oligo when bound to the M2 antibody.

3.4. The 3XABA Oligo Blocks the Association of Flag-Tagged Proteins with the M2 Antibody. If the ABA oligo and the Flag peptide compete for binding to the M2 antibody, then an excess of ABA oligo could potentially allow the elution of Flag-tagged proteins from the M2-coated beads, a commonly used procedure in protein purification. To address this possibility, we first determined whether an excess of ABA oligo could block binding of [S^{35}]-labeled Flag-TRF2 $^{\Delta\text{B}}$ to M2-coated beads. In these experiments, the M2-coated beads were first incubated with the competitors (30 μM of 3XFLAG peptide, 3XABA or 3XCTR oligos) or with no competitor (No Comp). Then, the [S^{35}]-labeled Flag-TRF2 $^{\Delta\text{B}}$ was added and the amount of [S^{35}] captured by the beads was measured by either SDS-PAGE electrophoresis (Figure 4(a)) or scintillation counting (Figure 4(b)). As Figures 4(a)–4(b) show, the 3XFLAG peptide abolished the capture of the [S^{35}]-labeled protein whereas the 3XCTR oligo had no effect. Although not as active as the 3XFLAG peptide, the 3XABA oligo inhibited the capture of Flag-TRF2 $^{\Delta\text{B}}$ by 62%. The shorter ABA aptamer was also tested but found to be less active (Figure S2), with ABA blocking the capture by 45% only.

Next, we asked whether an excess of 3XABA oligo could elute an [S^{35}]-labeled Flag-TRF2 $^{\Delta\text{B}}$ protein already bound to the M2-coated beads. In these experiments, the [S^{35}]-labeled protein was first captured by the M2-coated beads. Next, the radioactive beads were exposed to the competitors (30 μM of 3XFLAG peptide, 3XABA or 3XCTR oligos) or to no competitor (No Comp), and the amount of [S^{35}] released was quantified by either SDS-PAGE electrophoresis (Figure 4(c)) or scintillation counting (Figure 4(d)). As Figures 4(c)–4(d) show, the 3XFLAG peptide caused the release of 54% of the captured protein whereas the 3XCTR oligo had no activity. The 3XABA oligo was almost as effective as the 3XFLAG peptide, causing the release of 36% of the bound protein.

4. Discussion

The initial intent of this study was to use SELEX to characterize the ssDNA-binding specificity of Flag-tagged Stn1, a subunit of the CST complex [18, 19]. Instead, the SELEX procedure enriched for ssDNA molecules that bind with high

affinity to the anti-Flag M2 antibody. How was this selection by the antibody even possible? First, it may be that the *in vitro* translated Flag-Stn1 protein did not possess the minimum required DNA-binding activity. Improper folding could have contributed to this lack of activity. Alternatively, it may be that the OB-fold domain (oligonucleotide/oligosaccharide binding) of Stn1 is involved in mediating protein-protein interactions rather than DNA-binding [18, 19]. Second, the Flag epitope (DYKDDDDYK) is highly negatively charged, making it possible for an optimized DNA molecule to mimic the electrostatic signature of the peptide. Third, ssDNA is far more flexible than dsDNA, and in that respect, could more readily adapt to the positively charged surface of the antigen-binding pocket of the M2 antibody [22]. In other words, we propose that selection mediated by the antibody would not have been possible if the Flag epitope had been neutral or positively charged, if SELEX had been performed using dsDNA, or if the Flag-tagged Stn1 protein had exhibited a higher affinity for ssDNA. In fact, we have successfully used SELEX under similar conditions to define the ssDNA binding specificity of POT1, a protein that recognizes telomeric DNA. This SELEX, also performed using the M2 antibody, almost exclusively selected for telomeric sequences with no evidence of the bipartite motif detected (Choi, K. H. and Ouellette, M. M., manuscript in preparation). The selection by SELEX of off-target aptamers is not uncommon, especially when selecting single-stranded nucleic acids [23]. Off-target aptamers that bind to the selection matrix (e.g. avidin, M2 antibody), as we have observed here, have also been reported by others [23]. Hence, caution and experimental controls are necessary, especially when selecting ssDNA molecules for their binding to a protein of unknown functionality. Under these conditions, an important control to perform, as we have done here, is a mock selection to verify that the molecules were indeed selected by the protein of interest.

Our findings show that the identified bipartite consensus binds selectively and with high affinity to the M2 antibody. An important finding was the specificity of this interaction. For example, the consensus did not bind to any of the other antibodies tested, many of which belonging to the same isotype as the M2 antibody (mouse IgG $_1$). This specificity implied that the interactions involved the variable regions of the antibody, which together form the antigen-binding pocket. This assumption was confirmed by the results of the competition experiments. As Figure 4 shows, the ssDNA consensus could very efficiently displace the Flag peptide from the antigen-binding pocket of the M2 antibody. Conversely, the Flag peptide could similarly displace the DNA aptamer from the antibody as well (see Figure 3). This competition between the Flag peptide and DNA aptamer implies that the two are recognized by overlapping binding determinants at the surface of the antibody. A shared feature of the Flag peptide and ABA aptamer is their negative electrostatic charges. On the surface of the antibody, these negative charges on the two ligands are likely to interact with some of the same binding determinants. The binding pocket of the M2 antibody appears to be specific only for the first 4 amino acids of the Flag peptide, namely, DYKD [22]. Two clusters of positively charged residues (lysines, arginine, histidine) are

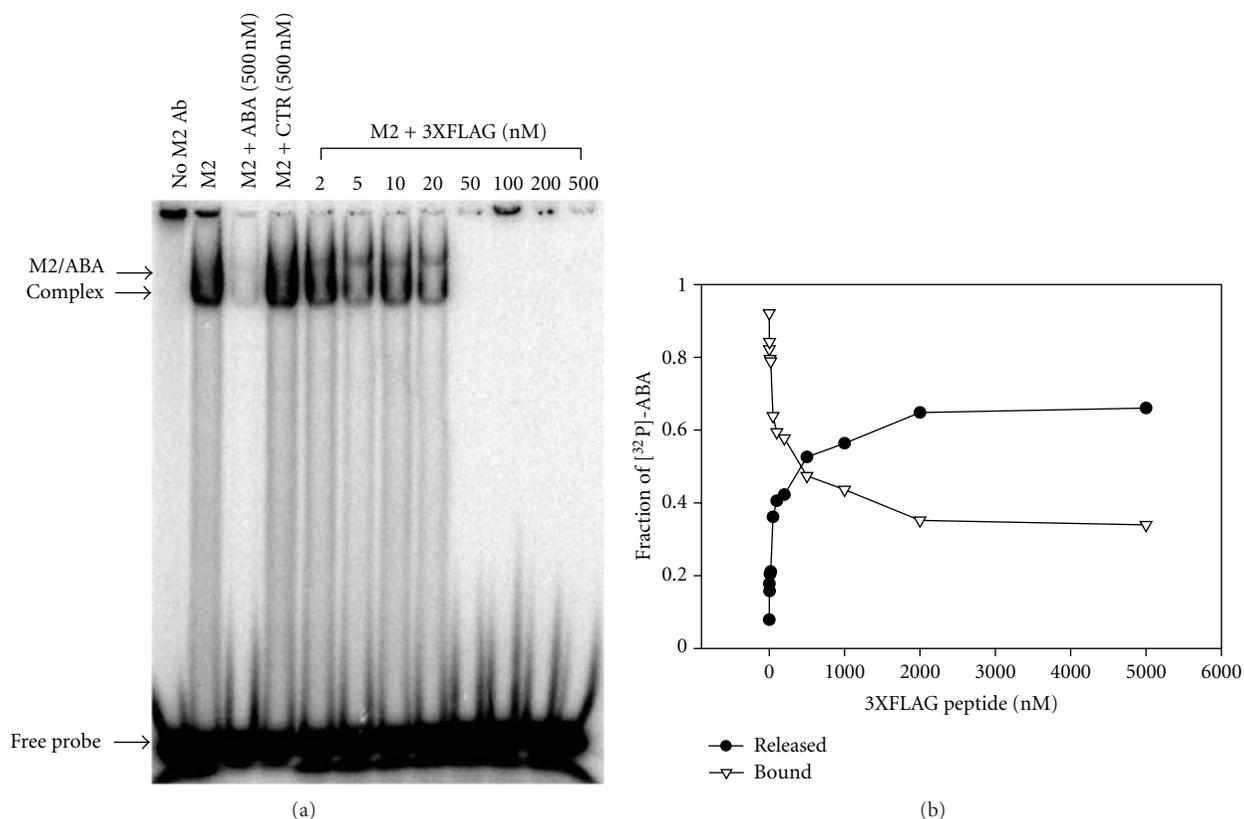


FIGURE 3: The 3XFLAG peptide blocks the binding of ABA to the M2 antibody. (a) The 3XFLAG peptide prevents formation of the ABA/M2 complex. The M2 antibody was incubated with an increasing concentration of 3XFLAG peptide after which the $[^{32}\text{P}]\text{-ABA}$ probe was added. Protein/DNA complexes were resolved by electrophoresis in a native polyacrylamide gel. (b) The 3XFLAG peptide elutes the ABA probe already bound to the M2 antibody. The $[^{32}\text{P}]\text{-ABA}$ was first captured by magnetic beads coated with the M2 antibody. The beads were then incubated with an increasing concentration of 3XFLAG peptide and the amount of $[^{32}\text{P}]\text{-ABA}$ released and remaining on the beads was counted by scintillation ($n = 1$).

present in the binding pocket, each capable of coordinating an aspartate within the DYKD motif. The same two clusters could potentially interact with phosphate groups present within the backbone of the ABA oligo. Also important for the binding are two tyrosine residues, each one capable of forming stacking interactions with the tyrosine of DYKD. When bound to the ABA oligo, the same tyrosines could provide stacking interactions to some of the nucleotide bases. Yet, these limited interactions would not suffice to explain the sequence specificity observed. The selected consensus consisted of two separate motifs (CCTTA and TGTCTWCC) and mutating one or the other was sufficient to abrogate binding. To allow recognition of this large bipartite consensus, other residues must be implicated that make contact with several of the nucleotide bases. Hence, it must be that the Flag peptide interacts with a subset only of all the residues implicated in the binding of ABA. If so, then how could the Flag peptide be so efficient at displacing ABA from the M2 antibody? It is possible that the binding determinants that the two ligands share are especially critical for the binding of ABA. A second possibility would be that the binding of the Flag peptide causes changes in the conformation of the antibody, and that these allosteric alterations are incompatible with the binding of ABA. How the ABA aptamer is able to mimic the structure

of the Flag peptide remains to be determined, but this form of mimicry between peptide and DNA has previously been reported to play an important role in autoimmune diseases and may also have important applications in drug design [5–9]. The ABA oligo and Flag peptide offer an ideal system with which to study the structural basis of this form of molecular mimicry.

One commonly employed procedure in protein purification is the capture of Flag-tagged proteins by the M2 antibody and their subsequent release by incubation with an excess of 3XFLAG peptide [24]. This procedure allows the elution of the purified proteins under non-denaturing conditions. However, the procedure leaves the eluted protein in solution with an excess of 3XFLAG peptide, which can pose a problem if the Flag tag needs to remain functional. A DNA aptamer capable of eluting the Flag-tagged proteins, such as 3XABA, could alleviate these potential drawbacks. Once the protein is eluted, the DNA aptamer could conveniently be inactivated by an antisense oligo or eliminated using trace amounts of DNase. Alternatively, if the aptamer is biotinylated, its removal could be accomplished using streptavidin-coated beads. Finally, a DNA aptamer would also offer the possibility of using mutated oligos as negative controls or for the differential elution of captured proteins.

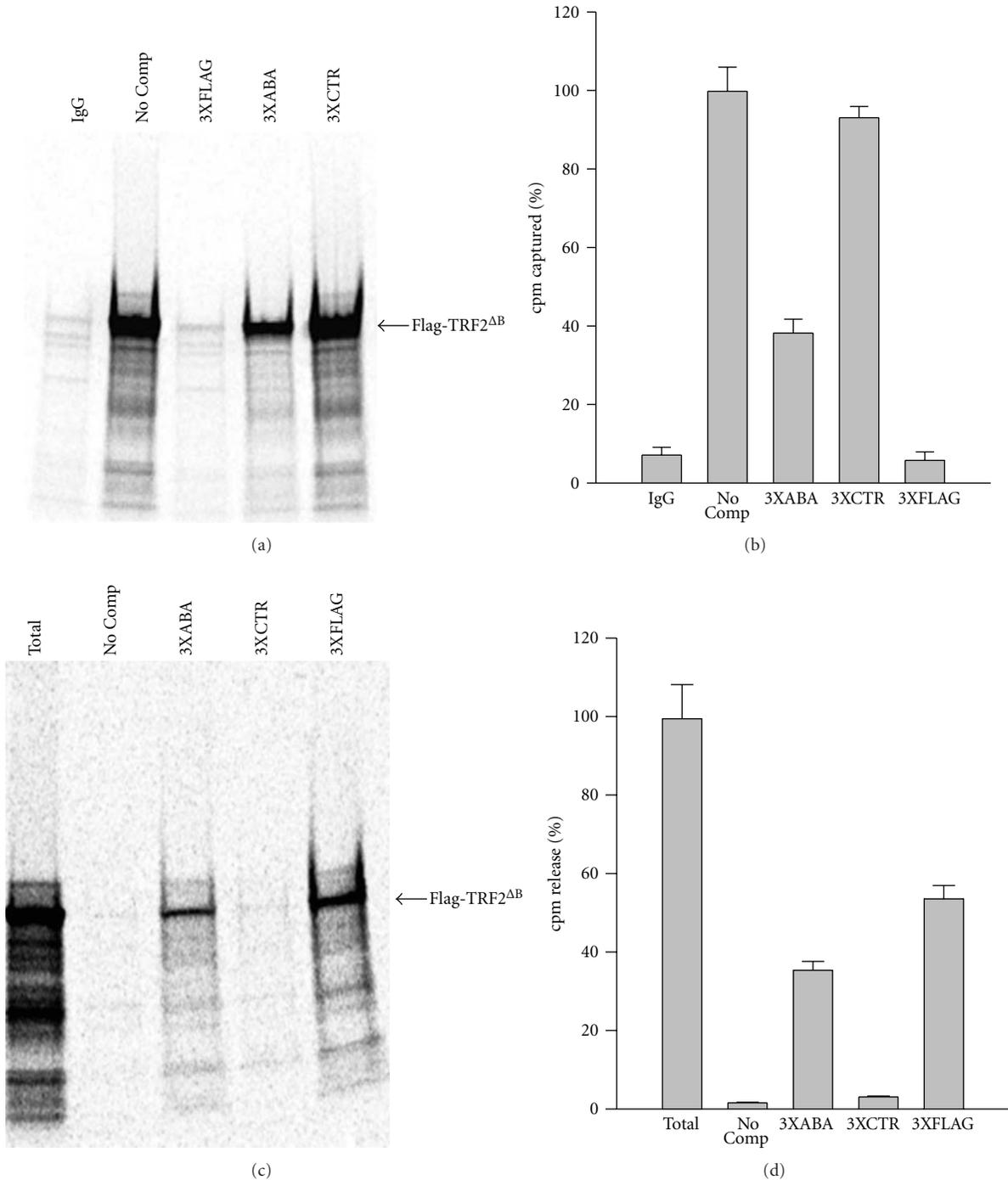


FIGURE 4: 3XABA oligonucleotide blocks the interaction of Flag-tagged proteins with the M2 antibody. (a)-(b) The 3XABA oligo blocks the binding of Flag-TRF2^{ΔB} to M2-coated beads. Magnetic beads coated with the M2 antibody were incubated in the absence (No Comp) or presence of the indicated competitor (3XABA, 3XCTR, 3XFLAG). *In vitro* translated [³⁵S]-labeled Flag-TRF2^{ΔB} was then added and the amount captured by the beads was determined by SDS-PAGE electrophoresis and exposure to a PhosphorImager cassette (a). In a second experiment done in triplicate, the amount of [³⁵S]-labeled protein captured was counted by scintillation (b). The amount of [³⁵S]-labeled protein captured in the absence of competitor (No Comp) was arbitrarily set to 100%. In both experiments, beads coated with normal mouse IgG were included as negative control for the capture (IgG). Data represent the mean ± S.D. (*n* = 3). (c)-(d) The 3XABA oligo elutes the Flag-TRF2^{ΔB} proteins already bound to M2-coated beads. The [³⁵S]-Flag-TRF2^{ΔB} protein was first captured by magnetic beads coated with the M2 antibody. The beads were then incubated in the absence (No comp) or presence of the indicated competitor (3XABA, 3XCTR, 3XFLAG). The amount of [³⁵S]-Flag-TRF2^{ΔB} released was determined by SDS-PAGE electrophoresis and exposure to a PhosphorImager cassette (c). In a second experiment done in triplicate, the amount of [³⁵S]-labeled protein released was counted by scintillation (d). The amount of [³⁵S]-labeled protein released by the boiling (total) was arbitrarily set to 100%. In both experiments, beads boiled to release all of the captured [³⁵S]-labeled protein were included as positive control for the elution (Total). Data represent the mean ± S.D. (*n* = 3).

These results described here may also have implications for the treatment of autoimmune diseases driven by antinuclear antibodies. Previous attempts to block the function of these autoantibodies have been made with peptide mimics, either isolated from phage display libraries or derived from peptide self-antigens with which these antinuclear antibodies cross-react [10, 11]. These results described therein show that SELEX can be used to identify DNA aptamers that block the function of a particular antibody. Could this approach be used to develop aptamers that block the function of antinuclear autoantibodies? In certain autoimmune diseases, blocking the function of antinuclear autoantibodies has been shown to limit tissue damage [10, 11]. The sequence specificity of antinuclear autoantibodies has not been systematically investigated [3, 4]. SELEX would represent an ideal method for characterizing the sequence specificity of antinuclear autoantibodies. The SELEX data could serve as starting point for the design of second-generation aptamers that might be useful to block the function of these autoantibodies. The development of these DNA aptamers could therefore represent an alternative approach to the treatment of patients afflicted with systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases.

5. Conclusion

In summary, SELEX identified a DNA aptamer that binds directly to the antigen-binding pocket of the anti-Flag M2 antibody. The DNA aptamer and Flag peptide competed for binding to the M2 antibody and the aptamer eluted Flag-tagged proteins from an immobilized M2 antibody. This new reagent therefore offers an alternative method for the elution of Flag-tagged proteins bound to the M2 antibody, a commonly employed procedure in protein purification. Aside from this immediate application in protein purification, identification of this bipartite consensus demonstrates the feasibility of using SELEX to develop DNA aptamers that block specific antibodies. Applying this approach to autoantibodies, particularly the antinuclear antibodies, could lead to the development of novel therapeutic strategies for patients with systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases.

Acknowledgments

The authors wish to thank Titia de Lange (Rockefeller University, New York, NY) for the pTetFLAGhTRF2^{45–501} plasmid. Appreciation is extended to Asserewou Honoré Etekpou for the quality of his technical assistance. They also wish to acknowledge the UNMC Eppley Cancer Center Molecular Biology Core and the UNMC DNA Sequencing Core Facility. This paper was supported by Grants from the National Institute of Health (P50CA127297, P30CA036727) and a GAANN fellowship from US Department of Education (P200A090064) to A. S. Lakamp.

References

[1] R. Lyons, S. Narain, C. Nichols, M. Satoh, and W. H. Reeves, "Effective use of autoantibody tests in the diagnosis of systemic

- autoimmune disease," *Annals of the New York Academy of Sciences*, vol. 1050, pp. 217–228, 2005.
- [2] Y. J. Jang and B. D. Stollar, "Anti-DNA antibodies: aspects of structure and pathogenicity," *Cellular and Molecular Life Sciences*, vol. 60, no. 2, pp. 309–320, 2003.
- [3] D. S. Pisetsky and S. A. Caster, "Binding specificity of a monoclonal anti-DNA antibody," *Molecular Immunology*, vol. 19, no. 5, pp. 645–650, 1982.
- [4] T. Sasaki, T. Murayoi, Y. Sekiguchi, E. Tamate, K. Yoshinaga, and Y. Kitagawa, "Monoclonal human anti-DNA antibodies from EB virus-transformed lymphocytes of systemic lupus erythematosus (SLE) patients," *Journal of Clinical Immunology*, vol. 5, no. 4, pp. 246–253, 1985.
- [5] L. A. DeGiorgio, K. N. Konstantinov, S. C. Lee, J. A. Hardin, B. T. Volpe, and B. Diamond, "A subset of lupus anti-DNA antibodies cross-reacts with the NR2 glutamate receptor in systemic lupus erythematosus," *Nature Medicine*, vol. 7, no. 11, pp. 1189–1193, 2001.
- [6] B. Deocharan, X. Qing, E. Beger, and C. Putterman, "Antigenic triggers and molecular targets for anti-double-stranded DNA antibodies," *Lupus*, vol. 11, no. 12, pp. 865–871, 2002.
- [7] T. W. Faust, E. H. Chang, C. Kowal et al., "Neurotoxic lupus autoantibodies alter brain function through two distinct mechanisms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 43, pp. 18569–18574, 2010.
- [8] J. B. Lefkowitz, M. Kiehl, J. Rubenstein et al., "Heterogeneity and clinical significance of glomerular-binding antibodies in systemic lupus erythematosus," *Journal of Clinical Investigation*, vol. 98, no. 6, pp. 1373–1380, 1996.
- [9] J. S. Rice, C. Kowal, B. T. Volpe, L. A. DeGiorgio, and B. Diamond, "Molecular mimicry: anti-DNA antibodies bind microbial and nonnucleic acid self-antigens," *Current Topics in Microbiology and Immunology*, vol. 296, pp. 137–151, 2005.
- [10] B. Gaynor, C. Putterman, P. Valadon, L. Spatz, M. D. Scharff, and B. Diamond, "Peptide inhibition of glomerular deposition of an anti-DNA antibody," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 5, pp. 1955–1960, 1997.
- [11] H. B. Lee, B. A. Diamond, and M. D. Blaufox, "In vivo detection of deposition of radiolabeled lupus antikidney antibody and its inhibition by soluble antigen," *Journal of Nuclear Medicine*, vol. 42, no. 1, pp. 138–140, 2001.
- [12] M. M. Ouellette and W. E. Wright, "Use of reiterative selection for defining protein-nucleic acid interactions," *Current Opinion in Biotechnology*, vol. 6, no. 1, pp. 65–72, 1995.
- [13] B. Zimmermann, I. Bilusic, C. Lorenz, and R. Schroeder, "Genomic SELEX: a discovery tool for genomic aptamers," *Methods*, vol. 52, no. 2, pp. 125–132, 2010.
- [14] D. M. Tasset, M. F. Kubik, and W. Steiner, "Oligonucleotide inhibitors of human thrombin that bind distinct epitopes," *Journal of Molecular Biology*, vol. 272, no. 5, pp. 688–698, 1997.
- [15] R. C. Becker, S. Oney, K. C. Becker, and B. Sullenger, "Antidote-controlled antithrombotic therapy targeting factor IXa and von willebrand factor," *Annals of the New York Academy of Sciences*, vol. 1175, pp. 61–70, 2009.
- [16] C. P. Rusconi, E. Scardino, J. Layzer et al., "RNA aptamers as reversible antagonists of coagulation factor IXa," *Nature*, vol. 419, no. 6902, pp. 90–94, 2002.
- [17] B. van Steensel, A. Smogorzewska, and T. de Lange, "TRF2 protects human telomeres from end-to-end fusions," *Cell*, vol. 92, no. 3, pp. 401–413, 1998.
- [18] C. M. Price, K. A. Boltz, M. F. Chaiken, J. A. Stewart, M. A. Beilstein, and D. E. Shippen, "Evolution of CST function in

- telomere maintenance,” *Cell Cycle*, vol. 9, no. 16, pp. 3157–3165, 2010.
- [19] Y. Miyake, M. Nakamura, A. Nabetani et al., “RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway,” *Molecular Cell*, vol. 36, no. 2, pp. 193–206, 2009.
- [20] A. Prakash, A. Natarajan, L. A. Marky, M. M. Ouellette, and G. E. Borgstahl, “Identification of the DNA-binding domains of human replication protein A that recognize G-quadruplex DNA,” *Journal of Nucleic Acids*, vol. 2011, Article ID 896947, 14 pages, 2011.
- [21] S. D. Mendonsa and M. T. Bowser, “In vitro evolution of functional DNA using capillary electrophoresis,” *Journal of the American Chemical Society*, vol. 126, no. 1, pp. 20–21, 2004.
- [22] T. P. Roosild, S. Castronovo, and S. Choe, “Structure of anti-FLAG M2 Fab domain and its use in the stabilization of engineered membrane proteins,” *Acta Crystallographica Section F*, vol. 62, no. 9, pp. 835–839, 2006.
- [23] G. F. Joyce, “Directed evolution of nucleic acid enzymes,” *Annual Review of Biochemistry*, vol. 73, pp. 791–836, 2004.
- [24] A. Einhauer and A. Jungbauer, “The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins,” *Journal of Biochemical and Biophysical Methods*, vol. 49, pp. 455–465, 2001.

Review Article

Non-B DNA Secondary Structures and Their Resolution by RecQ Helicases

Sudha Sharma

Department of Biochemistry and Molecular Biology, College of Medicine, Howard University, 520 W Street, NW, Suite 3424A, Washington, DC 20059, USA

Correspondence should be addressed to Sudha Sharma, sudha.sharma@howard.edu

Received 25 May 2011; Accepted 25 July 2011

Academic Editor: Souvik Maiti

Copyright © 2011 Sudha Sharma. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In addition to the canonical B-form structure first described by Watson and Crick, DNA can adopt a number of alternative structures. These non-B-form DNA secondary structures form spontaneously on tracts of repeat sequences that are abundant in genomes. In addition, structured forms of DNA with intrastrand pairing may arise on single-stranded DNA produced transiently during various cellular processes. Such secondary structures have a range of biological functions but also induce genetic instability. Increasing evidence suggests that genomic instabilities induced by non-B DNA secondary structures result in predisposition to diseases. Secondary DNA structures also represent a new class of molecular targets for DNA-interactive compounds that might be useful for targeting telomeres and transcriptional control. The equilibrium between the duplex DNA and formation of multistranded non-B-form structures is partly dependent upon the helicases that unwind (resolve) these alternate DNA structures. With special focus on tetraplex, triplex, and cruciform, this paper summarizes the incidence of non-B DNA structures and their association with genomic instability and emphasizes the roles of RecQ-like DNA helicases in genome maintenance by resolution of DNA secondary structures. In future, RecQ helicases are anticipated to be additional molecular targets for cancer chemotherapeutics.

1. Introduction

Structure of the right-handed B-form DNA has been known since 1953 [1]. Instead of being a conformationally homogeneous molecule, DNA has the capability of adopting several types of conformations as dictated by its sequence [2]. As early as 1957, association of ribonucleic poly-A and poly-U polymers into three-stranded complexes was revealed using sedimentation coefficient and optical absorption measurements [3]. It was later shown by atomic resolution single-crystal X-ray diffraction analysis that the DNA hexamer d(CpGpCpGpCpG) forms a left-handed conformation (Z-DNA) with altered helical parameters relative to the right-handed B-form [4]. This was followed by the identification of cruciform structures formed by inverted repeats [5, 6]. Finally, guanine-rich motifs in DNA were discovered to form parallel four-stranded complexes called tetraplex, G-quadruplex, or G4 DNA [7]. More than ten different DNA conformations have now been discovered [8], and these are often referred to as secondary structures, alternative DNA,

or non-B DNA. A non-B database has been developed for prediction of alternative DNA structures including Z-DNA motifs, quadruplex-forming motifs, inverted repeats, mirror repeats and direct repeats, and their associated subsets of cruciforms, triplex, and slipped structures, respectively [9].

Non-B DNA structures are functional genomic elements that play a variety of roles in the cell [10]. These include gene function and regulation [11], immune response [12], telomere maintenance [13], recombination [14], antigenic variation in human pathogens [15], and the generation of genomic diversity [16]. The DNA secondary structures are suggested to be involved in regulation at both transcriptional and translational levels; however, when the subtle balance between the replication, transcriptional, and repair machinery is impaired, these secondary structures may induce genetic instability. Alternate structure-forming sequences are known to be unstable and represent hotspots for deletion or recombination in bacteria, yeast, and mammals [17–20]. This genetic instability has generally been related to DNA replication because non-B structures cause DNA polymerase

pausing *in vitro* and replication fork pausing *in vivo* [21]. Slow replication was observed in an inverted repeat sequence in *Escherichia coli* [22], and inverted repeats lead to deletions or chromosomal rearrangements more frequently in yeast that are deficient in DNA polymerase activity [23, 24]. Slow progression of the replication fork could facilitate formation of secondary structures at long tracts of single-stranded DNA in the lagging-strand template [25]. These secondary structures pose obstacle to replication fork progression causing fork arrest and/or collapse ultimately leading to double-strand breaks (DSBs) and genome rearrangements [26, 27]. The formation of alternative DNA structures can also activate nucleotide excision and SOS pathways resulting in segments of single-stranded DNA (ssDNA) [28]. Such ssDNA regions can be converted to DSBs during replication and lead to mutations through mechanisms such as homologous recombination or nonhomologous end joining [29]. Conditions that favor the structural transitions from B-DNA to non-B DNA lead to genetic instability in model systems [27]. Alternative structure-mediated mutagenesis has been implicated in the incidence of gross rearrangements and deletions as well as point mutations [30–32]. There is significant circumstantial evidence for the involvement of DNA secondary structures in association with genetic instability leading to human disease [33, 34].

The vastness of mutagenic capability would be predicted to reduce the prevalence of secondary structure-forming tracts in genomes; however they are abundant and often enriched in the regulatory regions of genes [9]. With such an array of challenging sequence elements, it is evident that cells have developed the capacity for controlling the potential of these sequences for genome destabilization. Among several elucidated mechanisms to resolve secondary DNA structures, RecQ helicases represent an important class of enzymatic activities that are utilized to counteract such challenge to genomic stability.

2. RecQ Helicases

The RecQ family represents one of the most highly conserved groups of DNA helicases [35–39]. Bacteria and budding yeast have one RecQ homolog, RecQ and SGS1, respectively. The RecQ helicase family has 5 homologs in the human genome: *RECQ1*, *WRN*, *BLM*, *RecQ4*, and *RecQ5 β* (Figure 1). RecQ helicases share a centrally located helicase domain that couples nucleotide hydrolysis to DNA unwinding and defines the RecQ family. Many, but not all, RecQ helicases contain additional conserved RQC (RecQ C-terminal) and HRDC (helicase and RNaseD C-terminal) domains, which are implicated in protein interactions [40, 41] and DNA binding [42]. Eukaryotic RecQ helicases have, in addition, further N- and C-terminal extensions that are involved in protein-protein interactions and postulated to lend unique functional characteristics to each helicase [43, 44]. Certain RecQ homologues also have strongly acidic regions that have been shown to mediate interaction with single-strand DNA-binding proteins such as RPA [37]. Furthermore, nuclear

localization signal (NLS) has been identified for several RecQ proteins (Figure 1) [37].

RecQ helicases unwind DNA duplex with 3'-5' polarity in a reaction that is dependent of NTP hydrolysis [37]. Curiously, several RecQ proteins have been demonstrated to also have the ability to promote annealing of complementary strands in a reaction that is inhibited by the presence of ATP [37]. It was confirmed that ATP binding to the protein modulates oligomeric state of RECQ1 and regulates these apparently conflicting biochemical activities [45, 46].

RecQ helicases are remarkable among all DNA helicases for two primary reasons. First, in addition to unwinding duplex DNA, they are capable of unwinding a variety of DNA substrates containing noncanonical structures including forked duplexes, displacement loops (D-loops; an intermediate in homologous recombination reactions), triple helices, 3- or 4-way junctions, and G-quadruplex DNA [37, 47]. In fact, in many instances, they prefer these substrates to standard duplex DNA. Second, germline mutations in three human RecQ helicase homologs WRN, BLM, and RECQ4, which are located on chromosomes 8p-12, 15q-26.1, and 8q-24.3, respectively, give rise to rare genetic disorders of Werner, Bloom, and Rothmund-Thomson/RAPADILINO/Baller-Gerold syndromes, respectively, all of which are characterized by chromosomal instability and predisposition to cancer [48–51]. Distinct clinical features of these disorders indicate that human RecQ homologs perform unique cellular functions. Cellular studies point to critical requirement of RECQ1 and RecQ5 β [52, 53], but defects in these have not been associated with a human disease yet. As reviewed in next sections, collective biochemical, cellular, and genetic findings signify a pivotal role of RecQ helicases in resolution of non-B DNA structures and genome maintenance.

3. Prevalence, Consequences, and Unraveling of Non-B Secondary DNA Structures

Genomic maintenance entails highly regulated interaction of intrinsic factors such as the nature of sequence or the action of DNA replication and repair proteins and extrinsic factors such as environmental mutagens. Repetitive sequences in the genomes have the propensity to form complex secondary structures which could lead to diverse types of genomic instability. One of the common mechanisms of alternative structure-induced instability is obstruction of replication fork progression leading to fork stalling and/or collapse [54]. RecQ helicases are proposed as genome caretakers and guardians of DNA replication forks [55]. The following sections summarize impact of certain specific non-B DNA structures on genomic stability and review the roles of RecQ helicases in resolving these structures.

3.1. Cruciforms

3.1.1. Cruciform DNA Structures and Their Occurrence in Genomes. A cruciform structure is formed by intrastrand base pairing of inverted repeat sequences and is characterized

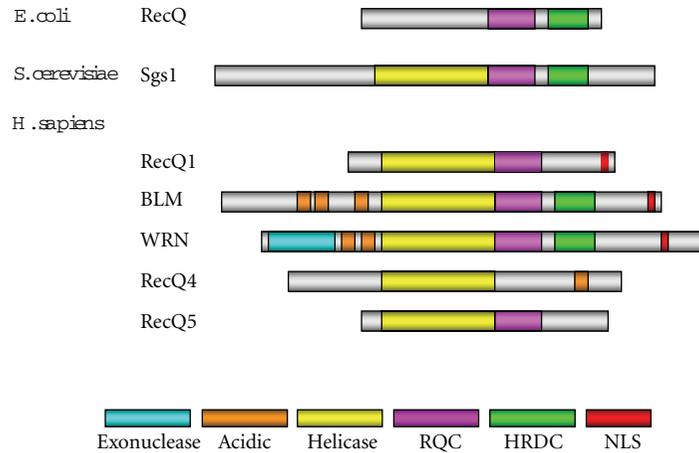


FIGURE 1: The RecQ helicase family. Proteins are aligned by their conserved helicase domain. Conserved domains and motifs in each group are shown by different colors as depicted at the bottom.

by the presence of a four-way junction in which two of the branches are hairpin structures formed on each strand of the inverted repeat [5] (Figure 2(a)). The bases located between the inverted repeats do not self-pair and instead form the apical loops of the hairpins; however the overall structure is stabilized by the free energy of negative supercoiling [56]. Cruciform is structurally similar to a Holliday Junction (HJ) recombination intermediate [57]. In fact, cruciform structures formed by extrusion of an inverted repeat sequence in supercoiled plasmids have been extensively used to study the mechanistic properties of HJ-resolving enzymes [58, 59].

The existence of cruciforms has been demonstrated *in vitro* [60] and *in vivo* [61–63]. Cruciform structures have been reported in the genome of *E. coli* [64] as well as mammalian cells [63]. Cruciform-forming inverted repeat sequences have been found at the operator and transcription termination regions [65], as well as at the replication origin region [63, 66]. The distribution of such sequences often overlaps with chromosomal regions prone to gross rearrangements [67]. Because cruciform structures are energetically unfavorable, they are thought to form transiently *in vivo* as stable structures. The action of cellular factors such as junction specific nucleases, binding proteins, and DNA helicases is suggested to affect the equilibrium and the rate of formation of cruciform structures *in vivo* [68].

3.1.2. Cruciforms and Genomic Instability. Palindromes, a specific type of inverted repeat separated by only very few base pairs, are poorly tolerated in *E. coli* cells and are underrepresented in the *S. cerevisiae* and human genomes [69, 70] presumably due to their tendency to form hairpin and cruciform structures, which could be recognized and cleaved by a nuclease [71] or could affect or slow down DNA replication [72]. Cruciform structures are found in mutagenic hotspots, and their presence has been suggested to be etiologic in causing rearrangements and chromosomal instability in humans [2, 73, 74]. The AT-rich palindromic

repeats involved in the recurrent t(11;22) constitutional translocation favor adopting a cruciform structure *in vitro* and involve frequent DSBs [75, 76]. Direct *Alu* repeats artificially inserted in an inverted orientation in the yeast genome undergo DSBs and enter a break-fusion cycle resulting in dicentric chromosomes [77]. It has been proposed that the break might be caused by a cruciform-specific resolution activity similar to HJ resolvase, which generally is thought to act on intermediates produced through homologous recombination in the repair of DSBs or stalled replication [67]. However, recent studies show that nearby inverted repeats in budding and fission yeasts recombine spontaneously and frequently to form dicentric and acentric chromosomes independent of DSB formation, possibly by a replication mechanism involving template switching [78, 79].

3.1.3. Cruciform Resolvases. Four-way DNA joint molecules, termed HJs, are key intermediates in recombination [80]. Proteins with the enzymatic ability to cleave synthetic HJs *in vitro* have been termed HJ “resolvases,” and these DNA junction-resolving enzymes exhibit considerable selectivity for the structure of their substrates [81]. *E. coli* RuvC and its associated proteins RuvA and RuvB constitute the archetypal resolvase system [82]. RuvC is a dimeric protein that promotes HJ resolution by introducing a pair of symmetrically related nicks in two diametrically opposed strands across the junction point [82]. Ongoing search for the eukaryotic equivalent of bacterial RuvC HJ resolvase has led to the discovery of a number of DNA endonucleases, including Mus81-Mms4/EME1 [83], Slx-Slx4/BTBD12/MUS312 [84–86], XPF-ERCC1 [87], and Yen1/GEN1 [88, 89]. Furthermore, MUS81-EME1 also forms part of a larger nuclease complex containing SLX1-SLX4 and XPF-ERCC1 raising the possibility that these nucleases cooperate to process HJs [81]. Thus, it appears that eukaryotes possess alternative, and mechanistically varied, ways to process HJs, perhaps reflecting the critical importance of this step for cell viability and mutation avoidance.

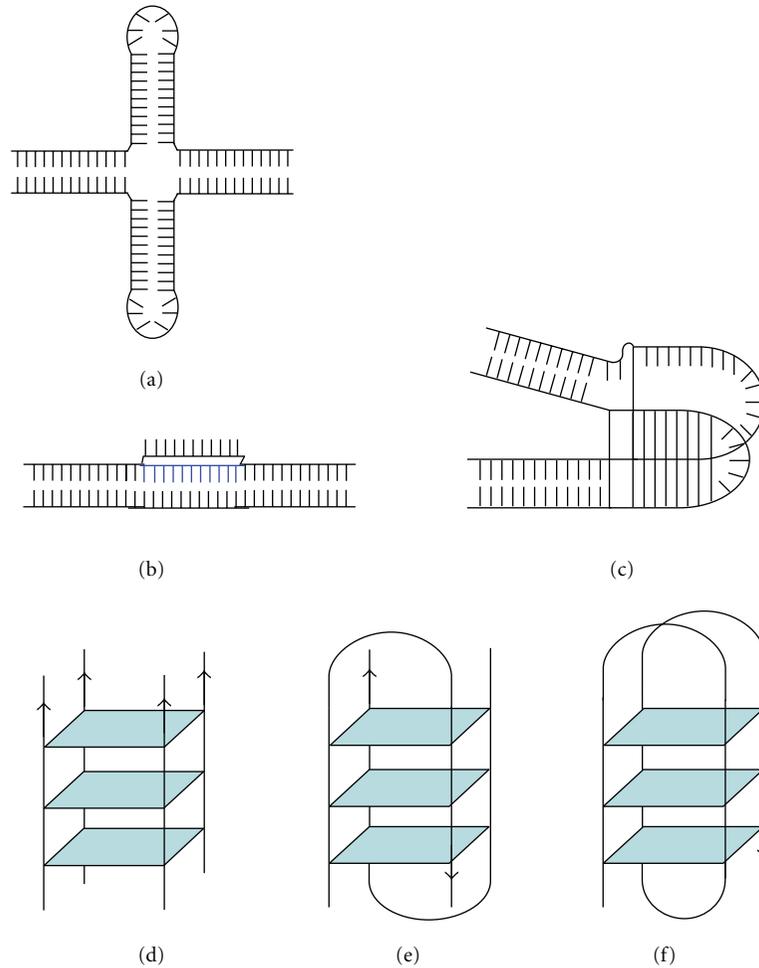


FIGURE 2: Schematic representation of certain non-B-form DNA structures. DNA can assume various alternate conformations depending upon the sequences. (a) Cruciform structures are formed at inverted repeats and also as intermediates of homologous recombination pathway, (b) Intermolecular triplex are formed by a triplex forming oligonucleotide (TFO, shown in blue) which binds to the purine-rich strand of the target duplex through the major groove. (c) Intramolecular triplex DNA structures can form at homopurine·homopyrimidine sequences with mirror symmetry, where a single-stranded region can bind in the major groove of the underlying DNA duplex to form a three-stranded helix. (d) G4 DNA is formed via parallel arrangement of four G-rich DNA strands; green boxes represent four guanine bases in planar arrangement via Hoogsteen base pairing. (e) Intermolecular G4 conformation is formed by DNA sequences with G-rich repeats forming hairpins that dimerize to stabilize bimolecular structure. (f) Intramolecular G4 DNA (or fold-over G quadruplex) is formed by single DNA strand with either four G-rich repeats or longer G tract that can fold upon themselves to form the G4 structure.

3.1.4. Metabolism of Cruciform-Like Structures by RecQ Helicases. Cellular and biochemical studies have established that RecQ helicases are vital in the metabolism of cruciform-like structures [90]. Loss of SGS1 in *S. cerevisiae* results in the accumulation of HR-dependent replication intermediates that resemble HJs [91]. In humans, DNA processing defects during replication and/or recombination have been suggested to contribute to the molecular pathology of Werner and Bloom syndromes [92]. Werner syndrome cells fail to resolve recombination intermediates [93], and the expression of wild-type WRN protein or RusA, a bacterial enzyme that cleaves four-way junctions, was both shown to rescue the WRN recombination defect and to improve cell survival following DNA damage [94]. Cytogenetic phenotype of elevated sister chromatid exchanges (SCEs) in Bloom syndrome cells has suggested hyperrecombination or

aberrant resolution of DNA recombination intermediates in the absence of active BLM protein [95]. Thus, a potential role for both WRN and BLM would be to prevent DNA structures including HJ that arise at blocked or collapsed replication forks from being processed into mature recombinants [55].

Indeed, RecQ helicases preferentially resolve 4-way HJs which are formally analogous to cruciform structures [37]. Several RecQ helicase proteins, including human BLM, WRN, RECQ5, and RECQ1, and the yeast homolog SGS1 were shown to selectively bind HJ structures and to promote ATP-dependent branch migration *in vitro* [44, 45, 96–99]. Furthermore, BLM [97], WRN [96], RECQ5 [44], and RECQ1 [100] are capable of branch-migrating HJs over several kilobases which is remarkable given that these helicases normally display poor processivity in the absence of RPA [37]. The bacterial HJ core recognition protein RuvA inhibits

HJ branch migration by BLM, WRN, RECQ1 or RECQ5 β suggesting that these RecQ helicases specifically recognize the HJ core where they initiate unwinding [37]. It is conceivable that RecQ helicases may promote branch migration with a mechanism similar to the oligomeric RuvAB branch migration motor wherein two RuvA tetramers bind the junction and promote the loading of two RuvB hexamers on the two arms of the junction [101]. This notion is supported by the fact that BLM, WRN, and RECQ1 form oligomeric structures in solution [102]. The WRN protein binds HJ as an oligomer [103], and an N-terminal fragment of BLM is known to form hexamers and dodecamers [104]. Remarkably, the N-terminal region (residues 1–56) of RECQ1 was found to be essential for both oligomerization and HJ resolution activity [105].

Slipped strand hairpins or cruciforms with single-stranded regions are also formed at tracts of trinucleotide repeats during replication and stall replication forks [10]. A structure-specific nuclease, Flap endonuclease 1 (FEN-1), has been implicated in resolution of such structures [54, 106, 107]. The nuclease activity of FEN-1 is robustly stimulated by physical interaction with WRN [108, 109]. WRN and FEN-1 directly interact at the sites of arrested replication forks suggesting that the formation of a functional FEN-1/WRN complex is important for resolving stalled DNA replication forks [110]. Cleavage of an HJ intermediate of fork regression by FEN-1 requires WRN branch fork migration [110]; WRN helicase activity initiating from the HJ core provides a suitable DNA molecule with a free 5' ssDNA end on which FEN-1 can load to ultimately catalyze structure-specific cleavage of the unwound 5' ssDNA arm [111]. However, stimulation of FEN-1 is mediated by direct protein-protein interaction but does not require WRN catalytic activity [108, 112]. In fact, expression of a conserved noncatalytic C-terminal domain of WRN necessary and sufficient for the physical and functional interaction with FEN-1 is sufficient to rescue the yeast *dna2-1* mutant phenotypes [113]. The conserved C-terminal in BLM was subsequently also found to mediate a physical and functional interaction with FEN-1 [109], and the phenotypes of yeast *dna2* mutants can be rescued by expression of BLM [114]. Importantly, BLM stimulated FEN-1 cleavage of foldback flaps, bubbles, or triplet repeats in a helicase-dependent manner [115, 116]. Thus, WRN and BLM helicases likely act as very effective remover of structures that inhibit FEN-1 and thereby prevent duplications, expansions, and other genome disruption [111].

Cruciform metabolism is also among critical functions of certain RecQ family helicases that are mediated by the species-specific interaction with topoisomerase III homologs [90]; the Rmi1 protein serves as an additional component of the heterotrimeric functional complex [117]. The major role of the conserved helicase-topoisomerase complex is to catenate or decatenate dsDNA, resulting in the resolution or “dissolution” of double HJs [118, 119]. Double HJs are shown to exist *in vivo* and are thought to arise when both ends of a DSB invade a homologous sequence at the final steps of homologous recombination [120, 121]. Human and *Drosophila* BLM proteins, but not other RecQ helicases, together with topoisomerase III α , have the ability

to catalyze double HJ dissolution on model DNA substrates in a reaction that requires BLM-mediated ATP hydrolysis and the active-site tyrosine residue of topoisomerase III α [118, 122]. BLAP75/RMI promotes BLM-dependent dissolution reaction by recruiting topoisomerase III α to the double HJ [123, 124]. Notably, this functional interaction is highly specific, as the BLAP-75 topoisomerase III α pair has no effect on either WRN or *E. coli* RecQ helicase activity, and *E. coli* Top3 cannot substitute for topoisomerase III α in the enhancement of the BLM helicase activity [125]. Dissolution of dHJs in *S. cerevisiae* is performed by the SGS1-Top3-Rmi1 complex [126]. Recent data are consistent with SGS1 and Top3 acting together *in vivo* because cells lacking SGS1 or Top3 exhibited persistent HJ-containing DNA structures following exposure to DNA damage [127].

Collectively, RecQ helicases constitute a remarkable group of enzymes that promote resolution of HJs via non-resolvase mechanisms, and this is believed to be one of their critical functions in genome maintenance.

3.2. Triplex

3.2.1. Triplex Structures and Their Occurrence in Genomes. Naturally occurring homopurine/homopyrimidine sequences can fold into triplex configuration by binding a third strand of DNA or RNA in the major groove of Watson-Crick duplex DNA through Hoogsteen or reversed Hoogsteen hydrogen bonds [128] (Figure 2(c)). Intermolecular triplexes are formed when the triplex-forming strand originates from a second DNA molecule, for example, triplex-forming oligonucleotides (TFOs) [128] (Figure 2(b)). Intramolecular triplexes are the major elements of H-DNAs in which the third strand is provided by one of the strands of the same duplex DNA molecule at homopurine:homopyrimidine sequences with mirror symmetry [129]. Unfavourable charge repulsion between the three negatively charged DNA strands contributes to the low stability of triplexes under physiological conditions. At physiological pH, triplex formation usually involves a purine-rich third strand that is antiparallel to the complementary strand and is stabilized by negative supercoiling, modification with phosphorothioate groups, or polyvalent cations such as Mg²⁺ or polyamines such as spermine and spermidine [129].

Triplexes have been shown to exist in chromosomes and nuclei, and the existence of H-DNA structures has been evidenced both *in vitro* and *in vivo* [130]. Triplex formation *in vivo* is supported by the identification of mammalian proteins that bind specifically to them [131, 132] and to the polypyrimidine [133] and polypurine single strands [18]. H-DNA conformations have been identified *in vivo* by using triplex-specific monoclonal antibodies [134, 135] and fluorescent “in situ nondenaturing” hybridization [136]. The presence of an H-DNA conformation *in vitro* has been demonstrated in constructs containing the sequences of interest from *E. coli* and mammalian genomic DNA or by using chemicals that modify nucleotides specifically in single-stranded DNA or double-stranded DNA [137, 138]. The sequence-specific DNA recognition and binding

characteristics of synthetic TFOs have been extensively studied because of their potential applications in genome modification and therapy [139]. Most annotated genes in both the mouse and human genomes are predicted to contain at least one unique potential TFO binding site [140]. Similarly, naturally occurring sequences capable of adopting H-DNA structures are very abundant in mammalian cells (~1 in every 50,000 bp in humans) [129, 141]. Majority of polypurine·polypyrimidine sequences are located in introns, promoters and 5' or 3' untranslated regions and are enriched in genes involved in cell signaling and cell communication [142]. Importantly, H-DNA structure-forming sequences are found flanking protooncogenes [143, 144].

3.2.2. Triplex and Genomic Instability. Naturally occurring triplexes are sources of genomic instability, and TFO can induce targeted mutagenesis, recombination, or DNA repair, and can inhibit proliferation and induce apoptosis in cultured cells [14, 31]. Genomic instability of human DNA sequences that can form triplexes is associated with the etiology of several diseases including neurological disorders [34]. For example, the triplex-forming potential of a (GAA)_n repeat has been correlated with the genomic instability and reduced frataxin gene expression in Friedreich's ataxia, a triplet repeat disorder [145]. The repeated sequence was shown to inhibit DNA polymerization *in vitro* and progression of replication forks *in vivo* suggesting that the triplex formation by the Friedreich's ataxia (GAA)_n repeat inhibits DNA replication [146, 147]. In addition to posing block to replication progression [148], naturally occurring triplex-forming sequences have been shown to interfere with transcription [144]. Many breakpoints on the translocated *c-myc* gene in Burkitt's lymphoma and t(12;15) BALB/c plasmacytomas are clustered around the H-DNA-forming sequences in the promoter regions [149]. Indeed, the naturally occurring H-DNA structure-forming sequence from the human *c-MYC* gene was shown to induce DSBs within these sequences in mammalian cells [32] and cause genomic instability in mice [149]. Collectively, these studies imply that the triplex structures result in fragile sites or mutation hotspots causing DSBs and subsequent translocation of the gene [129]. As part of the mechanisms whereby cells prevent the deleterious effects of alternate DNA structures, triplex formation *in vivo* is likely to be at least partly inhibited by destabilizing proteins or helicases.

3.2.3. Triplex Unwinding by RecQ Helicases. Triplex-forming sequences have been demonstrated to block replication *in vitro* [148]. Purified recombinant *E. coli* RecQ protein partially alleviates triplex formation and facilitates fork progression through triplex-forming DNA *in vitro*; loss of RecQ significantly increases the mutations caused by triplex-forming DNA *in vivo* [150]. RecQ-deficient *E. coli* utilize RecG helicase for fork regression upon encountering triplex structures and thereby restart replication [150]. A RecG equivalent in human is not identified yet, but human BLM and WRN helicases can unwind a DNA triple helix [151] and also catalyze fork regression [152]. *In vitro* studies with triple

helices formed by a pyrimidine motif third strand demonstrated that WRN and BLM catalyze triplex unwinding in nucleoside triphosphate hydrolysis-dependent reaction and require a free 3'-ssDNA overhang attached to the third strand [151]. Triplex unwinding by BLM and WRN does not require a single strand:double strand fork at the junction of the third strand and the triplex [151] indicating that the unwinding is promoted by inherent structural elements of the triplex substrate; this might also be facilitated by the oligomeric structures of these helicases [102]. More recently it was demonstrated that DHX9 (nuclear DNA helicase II (NDH II) or RNA helicase A (RHA)) protein, a superfamily 2 helicase, preferentially unwinds intermolecular triplex DNA substrates *in vitro* with a specific 3'-5' polarity with respect to the displaced third strand (similar to BLM and WRN helicase); this activity required a 3'-ssDNA overhang on the third strand and was dependent on ATP hydrolysis [153]. In contrast, a 5'-ssDNA overhang on the triplex-forming oligonucleotide is required to unwind the third strand of the triplex structure by FANCI, a superfamily 2 helicase that unwinds DNA in the 5'-3' direction [154]. This is consistent with the observations that FANCI requires a preexisting 5'-ssDNA to unwind conventional B-form duplex DNA substrates [155] and G-quadruplex DNA substrates [156]. Triplex unwinding by other human RecQ proteins has not been reported yet; however, examination of individual RecQ homolog might reveal differential preference for non-B DNA structures [105].

Identification and characterization of triplex unwinding helicases have signified the critical importance of triplex resolution for genomic stability. This is highlighted by the fact that mutations in WRN, BLM, FANCI lead to genomic instability and certain cancer in humans [157], whereas homozygous DHX9 knockout mice are embryonic lethal [158]. Cellular DNA metabolic processes involve transient formation of ssDNA which can possibly interact with other strands to form secondary and tertiary structures [159]. If triplexes are not resolved, they can potentially interfere with processes such as DNA replication, recombination, and repair [129]. Further investigations to uncover the roles of helicases in resolving triplex DNA structures are necessary for understanding the cellular mechanism(s) for genome stability maintenance.

3.3. G4 DNA

3.3.1. G4 DNA Structures and Their Occurrence in Genomes. G-quadruplexes form *in vitro* in guanine-rich sequences that contain four tracts of at least three guanines separated by other bases and are stabilized by G-quartets [160]. The G-quartets arise from the association of four guanines into a cyclic Hoogsteen hydrogen-bonding arrangement in which each guanine base makes two hydrogen bonds with its neighbor using different hydrogen-bonding positions to the canonical Watson-Crick base pairing. The planar G-quartets stack on top of each other, giving rise to four-stranded helical structures (Figure 2(d)). These structures, called G-tetraplex, G-quadruplex, or G4 DNA, may involve intramolecular or

intermolecular interactions, and the phosphodiester backbones of the four participating strands may be in parallel or antiparallel orientation [161] (Figures 2(e) and 2(f)). The formation of G4 structures is strongly dependent on monovalent cations such as K^+ and Na^+ and, hence, physiological buffer conditions favor their formation, and it has been suggested that G4 DNA may be routinely assembled and disassembled within cells [162].

The human genome contains nearly 376 000 distinct sites with the potential to form G4 DNA [163, 164], and the evidence for *in vivo* formation of G4 DNA has emerged in recent years [165]. Notably, G4 DNA has been observed by electron microscopy from transcribed human G-rich DNA arrays in bacteria [166] and by immunochemistry at the end of the ciliate *Oxytricha* telomeres [167]. The G-rich chromosomal domains predicted to form G4 DNA include four classes of repetitive regions: telomeres, rDNA, immunoglobulin heavy chain switch region, and G-rich minisatellites [168]. Replication, recombination, transcription, and telomeric DNA elongation involve steps in which two strands of duplex DNA can be unwound transiently, providing an opportunity for the G-rich strand to form quadruplex structures during these DNA metabolic events [165]. Formation of G4 DNA modulates key cellular processes such as immunoglobulin gene rearrangement, promoter activation, and telomere maintenance [169].

3.3.2. G4 DNA and Genomic Instability. A direct link between potential G4-forming sequences and genomic instability has been provided by genetic studies in model organisms. *DOG-1* (deletions of guanine-rich DNA), an ortholog of mammalian FANCF helicase, is essential for the stability of G-tracts in the genome of *C. elegans* [170, 171]. Worms defective in *DOG-1* accumulate deletions in regions of the genome containing long G-tracts [171] whereas the introduction of a G-quadruplex-forming DNA sequence into *C. elegans* is highly mutagenic and is removed from genomes lacking *DOG-1* [170]. FANCF is one of the 13 known genes which lead to Fanconi anemia, and cells from patients lacking functional FANCF accumulate large genomic DNA deletions that map to potential G4-forming sequences [172]. Moreover, FANCF preferentially unwinds G quadruplexes over other DNA substrates *in vitro* suggesting that the FANCF helicase, like *DOG-1*, functions in resolving potential replication impediments caused by DNA G-quadruplexes [156]. The RTEL (regulator of telomere length) helicase, another *DOG-1* homolog, has a very clear role at telomeres in mice, and *RTEL*-deficient embryonic stem cells exhibit chromosome-end fusions lacking detectable telomere signals [173]. It is suggested that G-quadruplexes impose a structural barrier to DNA replication and various nucleic acid processing enzymes and are a potential source of genetic instability if not resolved. Identification of several DNA helicases that efficiently unwind and disrupt G4 DNA indicates that eukaryotic cells possess the mechanism for resolution of G4 DNA structures.

3.3.3. Resolution of G4 DNA by RecQ Helicases. RecQ family members are prominent in that they preferentially unwind

tetraplex DNA [37]. The *E. coli* RecQ [174], yeast SGS1 [175], and human WRN [176] and BLM [177] proteins have been demonstrated to melt synthetic G4 DNA constructs. Both SGS1 and BLM unwind G4 DNA with at least 15-fold preference relative to duplex substrates [175, 178, 179] and HJ structures [178]. This substrate preference correlates with the binding affinity and maps to the conserved RQC region of the RecQ proteins [41]. The G4 DNA unwinding activity is proposed to contribute in the maintenance of two G-rich genomic domains, rDNA and telomeres. SGS1 is required for recombination-mediated lengthening of telomeres in telomerase-deficient *S. cerevisiae* [180–182]. Furthermore, *SGS1*-deficient cells are characterized by nucleolar fragmentation and production of rDNA circles suggesting a role of SGS1 in rDNA metabolism [183, 184]. A possible role of WRN in rDNA metabolism is indicated by the fact that a significant fraction of WRN is nucleolar [185]. Notably, cells from Werner syndrome patients show premature senescence and accelerated rates of telomere shortening [186]. WRN helicase was shown to be necessary for preventing dramatic telomere loss during lagging-strand replication of the G-rich strand and the consequent accumulation of chromosome aberrations such as chromosome fusions [187]. Consistent with a role in telomere maintenance, the WRN helicase is localized to telomeres, possibly via its interaction with TRF2 which also binds BLM [188, 189].

By resolving the tetraplex and other non-B DNA structures, RecQ proteins might clear the way for DNA polymerase during replication or repair synthesis. In support of this hypothesis, Kamath-Loeb et al. demonstrated that physical association of DNA polymerase δ with WRN enables unwinding of tetraplex (and hairpin) structures by the helicase and allows polymerase to pass through the roadblock [190, 191]. WRN was also shown to physically interact with p50 subunit of human pol δ which constitutes the active dimeric core of the enzyme with p125 subunit [192]. Thus a possible function of WRN (and presumably other RecQ proteins) might be the recruitment of this polymerase to the complex secondary structures and restoration of stalled DNA synthesis. Indeed, stimulation of DNA polymerase activity of pol δ by BLM and stimulation of BLM helicase activity by pol δ have been demonstrated [193]. Cellular phenotypes of genetic mutants and the demonstration of robust G4-unwinding activity *in vitro* support the notion that failure to unwind G4 DNA contributes in part to the genetic instability observed in Bloom and Werner syndrome cells.

Yet, RecQ proteins are not the only helicases known to resolve G4 structures. Besides FANCF and its orthologs, Pif1 (petite integration frequency 1), a 5'-3' helicase, processes G4-forming sequences *in vivo* and *in vitro* [194]. Human Pif1 helicase has been shown to bind and unwind G-quadruplex DNA [195]. In yeast, the involvement of Pif1 in telomere stability has been well established [196], and the association of hPif1 with telomeres and telomerase [197] indicates that hPif1 is a telomere G4 DNA-binding protein. Using genome-wide chromatin immunoprecipitation and *Pif1*-deficient cells, Zakian group has recently demonstrated that G4 motifs are a significant subset of the *in vivo* binding sites of the *S. cerevisiae* Pif1, and DNA replication through

G4 motifs is promoted by the *S. cerevisiae* Pif1 DNA helicase [198]. The G4 DNA resolving activity of mammalian Pif1 is of questionable significance as such since the *Pif1*-null mice are normal [199], a contrasting condition with WRN or BLM-deficient cells where genomic instability can be readily detected. It is conceivable that Pif1 activity is normally unnecessary, with sufficient G4 resolvase activity provided by other helicases (e.g., WRN, BLM, and FANCD1). It is possible that a requirement for Pif1 in mammalian cells would be obvious when one or more of these other G4 resolvase systems are compromised.

G4 resolution is, nevertheless, not a common characteristic of all RecQ helicases. Recently, it was demonstrated that RECQ1 does not unwind G-quadruplex substrates [105]. The inability to resolve this particular form of alternate DNA structure distinguishes RECQ1 from WRN, BLM, SGS1, or *E. coli* RecQ helicases which proficiently unwind a variety of G-quadruplex DNA substrates [45]. Furthermore, the telomere lengths of *RECQ1* wild-type, knockout, or heterozygous mouse cells show no significant difference suggesting minimum to no role of RECQ1 in telomere maintenance [200]. However, RECQ1 was purified with human telomeric chromatin specifically in cells that use a recombination mediated pathway known as Alternative Lengthening of Telomeres (ALTs) for telomere maintenance [201]. It is possible that RECQ1 plays indirect role in telomere metabolism via its interacting partners. Supporting this notion, recent evidence suggests that SGS1 regulates processing of telomeres by the 5'-3' exonuclease, EXO1 [202]. Interestingly, RECQ1 and EXO1 exhibit physical and functional interaction in human cells [203]; however, it remains to be tested whether they collaborate in a complex for accurate processing of chromosome ends.

Regardless, it has been demonstrated through various studies that certain RecQ helicases are crucial for the metabolism of G4 DNA structures at specific genome locations such as telomeres and rDNA. Mutant phenotypes in yeast and humans affirm vital importance of this function of RecQ helicases in genome maintenance.

4. Concluding Remarks and Outlook

Proficiency of RecQ helicases in unwinding alternate DNA structures has implicated them as roadblock removers for replication fork progression since the DNA sequences that can form unusual, non-B-form structures have been shown to block polymerases *in vitro* [21]. It has been proposed that at least one function of the RecQ DNA helicases is to prevent aberrant deleterious recombinogenic pathways when replication is perturbed by DNA damage, alternate DNA structure, or impaired DNA synthesis [204]. The processing of aberrant DNA structures by RecQ helicases is likely to counter their potential toxicity incurred by recombinogenic pathways [205] (Figure 3). The ability of helicases to unwind non-B DNA structures would be expected to increase access to repair and replication proteins. The RecQ helicases work in close coordination with other proteins (e.g., topoisomerases) to resolve various secondary structures.

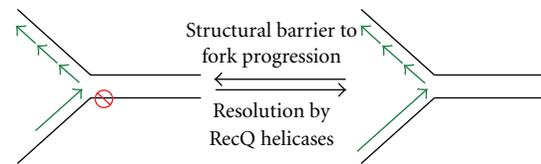


FIGURE 3: Proposed function of secondary structure resolution by RecQ helicases. Alternate DNA structures (indicated by red symbol) act as natural barrier to replication fork progression and can lead to fork stalling, collapse and ensue recombinogenic processing. Resolution by RecQ helicases is proposed to serve as roadblock remover for smooth progression of replication fork to avoid mutagenesis and/or genome rearrangements.

Associations of RecQ helicases with proteins critical for various steps of DNA replication and repair [37] suggest that RecQ helicases might cooperate with them to ensure faithful progression of replication forks through natural impediments by non-B DNA structures. It is likely that there is competition between the proteins which promote non-B secondary structure formation and RecQ helicases *in vivo*. Future studies will uncover how the activities of RecQ-helicases are controlled/regulated (via protein-protein interactions, posttranslational modifications, etc.) for maintaining genomic stability in general and preventing non-B DNA structure-induced instability in particular.

As noted above, RecQ proteins exhibit subtle but noteworthy differences among themselves with respect to their ability to unwind or preference for certain alternate DNA structures. Evidently, current data implies that individual human RecQ homologs are uniquely required to unwind specific DNA structures *in vivo*. Further investigation is essential to elucidate cellular environment, genomic contexts, and/or protein factors that license a specific RecQ protein to metabolize specific DNA structures *in vivo*. It is clear that non-B structures both perform physiological roles and potentiate genomic instability. Analyses of the mutation spectrum and genomic rearrangements in RecQ-deficient cells will illustrate significance of RecQ helicases in underlying mutational mechanisms associated with non-B DNA structures [206].

An opportunistic aspect of the unique nature of naturally occurring non-B DNA conformations is to use them as potential target for cancer therapy since these sequence-specific structures are proposed to affect gene expression and telomere activation, respectively [207, 208]. Gene expression of oncogenes could be selectively inhibited by using chemicals (drugs) or small molecules targeted to specific non-B DNA conformation present in their regulatory regions [129, 209]; stabilizing the secondary structures would be predicted to prevent access of nucleic acid binding proteins and interfere with critical cellular processes. Considering the demonstrated roles of RecQ helicases in resolving such non-B DNA structures, specific inhibition of RecQ and other non-B resolving helicases via small molecules [210], DNA-binding drugs or gene silencing might be a promising strategy to explore for anticancer therapy [207]. Development of new methodologies to investigate specific functions of non-B

DNA structures and identification of novel structure-specific DNA helicases involved in resolution of such secondary structures will certainly expand the array of molecular targets available for drug development and therapeutic intervention.

Acknowledgment

This work is made possible in part by NIH/NCRR Grant 2 G12 RR003048 and NIH/NIGMS Grant 7SC1GM093999-02.

References

- [1] J. D. Watson and F. H. C. Crick, "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid," *Nature*, vol. 171, no. 4356, pp. 737–738, 1953.
- [2] A. Bacolla and R. D. Wells, "Non-B DNA conformations, genomic rearrangements, and human disease," *Journal of Biological Chemistry*, vol. 279, no. 46, pp. 47411–47414, 2004.
- [3] G. Felsenfeld, D. R. Davies, and A. Rich, "Formation of a three-stranded polynucleotide molecule," *Journal of the American Chemical Society*, vol. 79, no. 8, pp. 2023–2024, 1957.
- [4] A. H. J. Wang, G. J. Quigley, F. J. Kolpak et al., "Molecular structure of a left-handed double helical DNA fragment at atomic resolution," *Nature*, vol. 282, no. 5740, pp. 680–686, 1979.
- [5] D. M. Lilley, "The inverted repeat as a recognizable structural feature in supercoiled DNA molecules," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 77, no. 11, pp. 6468–6472, 1980.
- [6] N. Panayotatos and R. D. Wells, "Cruciform structures in supercoiled DNA," *Nature*, vol. 289, no. 5797, pp. 466–469, 1981.
- [7] D. Sen and W. Gilbert, "Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis," *Nature*, vol. 334, no. 6180, pp. 364–366, 1988.
- [8] D. Svozil, J. Kalina, M. Omelka, and B. Schneider, "DNA conformations and their sequence preferences," *Nucleic Acids Research*, vol. 36, no. 11, pp. 3690–3706, 2008.
- [9] R. Z. Cer, K. H. Bruce, U. S. Mudunuri et al., "Non-B DB: a database of predicted non-B DNA-forming motifs in mammalian genomes," *Nucleic Acids Research*, vol. 39, supplement 1, pp. D383–D391, 2011.
- [10] J. Zhao, A. Bacolla, G. Wang, and K. M. Vasquez, "Non-B DNA structure-induced genetic instability and evolution," *Cellular and Molecular Life Sciences*, vol. 67, no. 1, pp. 43–62, 2010.
- [11] B. P. Belotserkovskii, R. Liu, S. Tornaletti, M. M. Krasilnikova, S. M. Mirkin, and P. C. Hanawalt, "Mechanisms and implications of transcription blockage by guanine-rich DNA sequences," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 29, pp. 12816–12821, 2010.
- [12] D. A. Collier, J. A. Griffin, and R. D. Wells, "Non-B right-handed DNA conformations of homopurine-homopyrimidine sequences in the murine immunoglobulin C(α) switch region," *Journal of Biological Chemistry*, vol. 263, no. 15, pp. 7397–7405, 1988.
- [13] S. Neidle and G. N. Parkinson, "The structure of telomeric DNA," *Current Opinion in Structural Biology*, vol. 13, no. 3, pp. 275–283, 2003.
- [14] J. Y. Chin, E. B. Schleifman, and P. M. Glazer, "Repair and recombination induced by triple helix DNA," *Frontiers in Bioscience*, vol. 12, pp. 4288–4297, 2007.
- [15] J. T. Meier, M. I. Simon, and A. G. Barbour, "Antigenic variation is associated with DNA rearrangements in a relapsing fever *Borrelia*," *Cell*, vol. 41, no. 2, pp. 403–409, 1985.
- [16] S. S. Smith, "Evolutionary expansion of structurally complex DNA sequences," *Cancer Genomics and Proteomics*, vol. 7, no. 4, pp. 207–215, 2010.
- [17] G. A. Cromie, C. B. Millar, K. H. Schmidt, and D. R. Leach, "Palindromes as substrates for multiple pathways of recombination in *Escherichia coli*," *Genetics*, vol. 154, no. 2, pp. 513–522, 2000.
- [18] A. Collick, J. Drew, J. Penberth et al., "Instability of long inverted repeats within mouse transgenes," *EMBO Journal*, vol. 15, no. 5, pp. 1163–1171, 1996.
- [19] D. A. Gordenin, K. S. Lobachev, N. P. Degtyareva, A. L. Malkova, E. Perkins, and M. A. Resnick, "Inverted DNA repeats: a source of eukaryotic genomic instability," *Molecular and Cellular Biology*, vol. 13, no. 9, pp. 5315–5322, 1993.
- [20] D. Branzei and M. Foiani, "Leaping forks at inverted repeats," *Genes and Development*, vol. 24, no. 1, pp. 5–9, 2010.
- [21] E. V. Mirkin and S. M. Mirkin, "Replication fork stalling at natural impediments," *Microbiology and Molecular Biology Reviews*, vol. 71, no. 1, pp. 13–35, 2007.
- [22] D. R. Leach, "Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair," *BioEssays*, vol. 16, no. 12, pp. 893–900, 1994.
- [23] B. Ruskin and G. R. Fink, "Mutations in POL1 increase the mitotic instability of tandem inverted repeats in *Saccharomyces cerevisiae*," *Genetics*, vol. 134, no. 1, pp. 43–56, 1993.
- [24] F. J. Lemoine, N. P. Degtyareva, K. Lobachev, and T. D. Petes, "Chromosomal translocations in yeast induced by low levels of DNA polymerase: a model for chromosome fragile sites," *Cell*, vol. 120, no. 5, pp. 587–598, 2005.
- [25] R. R. Sinden, V. I. Hashem, and W. A. Rosche, "DNA-Directed mutations. Leading and lagging strand specificity," *Annals of the New York Academy of Sciences*, vol. 870, pp. 173–189, 1999.
- [26] G. Wang and K. M. Vasquez, "Non-B DNA structure-induced genetic instability," *Mutation Research*, vol. 598, no. 1-2, pp. 103–119, 2006.
- [27] A. Bacolla, M. Wojciechowska, B. Kosmider, J. E. Larson, and R. D. Wells, "The involvement of non-B DNA structures in gross chromosomal rearrangements," *DNA Repair*, vol. 5, no. 9-10, pp. 1161–1170, 2006.
- [28] A. Bacolla, A. Jaworski, T. D. Connors, and R. D. Wells, "PKD1 unusual DNA conformations are recognized by nucleotide excision repair," *Journal of Biological Chemistry*, vol. 276, no. 21, pp. 18597–18604, 2001.
- [29] E. Petermann and T. Helleday, "Pathways of mammalian replication fork restart," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 10, pp. 683–687, 2010.
- [30] A. Bacolla, A. Jaworski, J. E. Larson et al., "Breakpoints of gross deletions coincide with non-B DNA conformations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 39, pp. 14162–14167, 2004.
- [31] G. Wang, M. M. Seidman, and P. M. Glazer, "Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair," *Science*, vol. 271, no. 5250, pp. 802–805, 1996.

- [32] G. Wang and K. M. Vasquez, "Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 37, pp. 13448–13453, 2004.
- [33] R. D. Wells, "Discovery of the role of non-B DNA structures in mutagenesis and human genomic disorders," *The Journal of Biological Chemistry*, vol. 284, no. 14, pp. 8997–9009, 2009.
- [34] A. Bacolla and R. D. Wells, "Non-B DNA conformations as determinants of mutagenesis and human disease," *Molecular Carcinogenesis*, vol. 48, no. 4, pp. 273–285, 2009.
- [35] V. A. Bohr, "Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance," *Trends in Biochemical Sciences*, vol. 33, no. 12, pp. 609–620, 2008.
- [36] I. D. Hickson, "RecQ helicases: caretakers of the genome," *Nature Reviews Cancer*, vol. 3, no. 3, pp. 169–178, 2003.
- [37] S. Sharma, K. M. Doherty, and R. M. Brosh Jr., "Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability," *Biochemical Journal*, vol. 398, no. 3, pp. 319–337, 2006.
- [38] K. J. Ouyang, L. L. Woo, and N. A. Ellis, "Homologous recombination and maintenance of genome integrity: cancer and aging through the prism of human RecQ helicases," *Mechanisms of Ageing and Development*, vol. 129, no. 7–8, pp. 425–440, 2008.
- [39] J. A. Cobb, L. Bjergbaek, and S. M. Gasser, "RecQ helicases: at the heart of genetic stability," *FEBS Letters*, vol. 529, no. 1, pp. 43–48, 2002.
- [40] J. W. Lee, R. Kusumoto, K. M. Doherty et al., "Modulation of Werner syndrome protein function by a single mutation in the conserved RecQ domain," *Journal of Biological Chemistry*, vol. 280, no. 47, pp. 39627–39636, 2005.
- [41] M. D. Huber, M. L. Duquette, J. C. Shiels, and N. Maizels, "A conserved G4 DNA binding domain in RecQ family helicases," *Journal of Molecular Biology*, vol. 358, no. 4, pp. 1071–1080, 2006.
- [42] D. A. Bernstein and J. L. Keck, "Conferring substrate specificity to DNA helicases: role of the RecQ HRDC domain," *Structure*, vol. 13, no. 8, pp. 1173–1182, 2005.
- [43] L. Wu and I. D. Hickson, "DNA helicases required for homologous recombination and repair of damaged replication forks," *Annual Review of Genetics*, vol. 40, pp. 279–306, 2006.
- [44] P. L. Garcia, Y. Liu, J. Jiricny, S. C. West, and P. Janscak, "Human RECQ51 β , a protein with DNA helicase and strand-annealing activities in a single polypeptide," *EMBO Journal*, vol. 23, no. 14, pp. 2882–2891, 2004.
- [45] S. Sharma, J. A. Sommers, S. Choudhary et al., "Biochemical analysis of the DNA unwinding and strand annealing activities catalyzed by human RECQ1," *Journal of Biological Chemistry*, vol. 280, no. 30, pp. 28072–28084, 2005.
- [46] L. Muzzolini, F. Beuron, A. Patwardhan et al., "Different quaternary structures of human RECQ1 are associated with its dual enzymatic activity," *PLoS Biology*, vol. 5, no. 2, article e20, 2007.
- [47] C. Z. Bachrati and I. D. Hickson, "Analysis of the DNA unwinding activity of RecQ family helicases," *Methods in Enzymology*, vol. 409, pp. 86–100, 2006.
- [48] C. Z. Bachrati and I. D. Hickson, "RecQ helicases: suppressors of tumorigenesis and premature aging," *Biochemical Journal*, vol. 374, no. 3, pp. 577–606, 2003.
- [49] C. E. Yu, J. Oshima, Y. H. Fu et al., "Positional cloning of the Werner's syndrome gene," *Science*, vol. 272, no. 5259, pp. 258–262, 1996.
- [50] N. A. Ellis, J. Groden, T. Z. Ye et al., "The Bloom's syndrome gene product is homologous to RecQ helicases," *Cell*, vol. 83, no. 4, pp. 655–666, 1995.
- [51] S. Kitao, A. Shimamoto, M. Goto et al., "Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome," *Nature Genetics*, vol. 22, no. 1, pp. 82–84, 1999.
- [52] S. Sharma and R. M. Brosh Jr., "Unique and important consequences of RECQ1 deficiency in mammalian cells," *Cell Cycle*, vol. 7, no. 8, pp. 989–1000, 2008.
- [53] Y. Hu, S. Raynard, M. G. Sehorn et al., "RECQL5/Recq15 helicase regulates homologous recombination and suppresses tumor formation via disruption of Rad51 presynaptic filaments," *Genes and Development*, vol. 21, no. 23, pp. 3073–3084, 2007.
- [54] R. D. Wells, R. Dere, M. L. Hebert, M. Napierala, and L. S. Son, "Advances in mechanisms of genetic instability related to hereditary neurological diseases," *Nucleic Acids Research*, vol. 33, no. 12, pp. 3785–3798, 2005.
- [55] C. Z. Bachrati and I. D. Hickson, "RecQ helicases: guardian angels of the DNA replication fork," *Chromosoma*, vol. 117, no. 3, pp. 219–233, 2008.
- [56] E. A. Oussatcheva, J. Pavlicek, O. F. Sankey, R. R. Sinden, Y. L. Lyubchenko, and V. N. Potaman, "Influence of global DNA topology on cruciform formation in supercoiled DNA," *Journal of Molecular Biology*, vol. 338, no. 4, pp. 735–743, 2004.
- [57] G. R. Smith, "Meeting DNA palindromes head-to-head," *Genes and Development*, vol. 22, no. 19, pp. 2612–2620, 2008.
- [58] R. R. Sinden and D. E. Pettijohn, "Cruciform transitions in DNA," *Journal of Biological Chemistry*, vol. 259, no. 10, pp. 6593–6600, 1984.
- [59] J. M. Fogg and D. M. Lilley, "Ensuring productive resolution by the junction-resolving enzyme RuvC: large enhancement of the second-strand cleavage rate," *Biochemistry*, vol. 39, no. 51, pp. 16125–16134, 2000.
- [60] K. M. Karrer and J. G. Gall, "The macronuclear ribosomal DNA of *Tetrahymena pyriformis* is a palindrome," *Journal of Molecular Biology*, vol. 104, no. 2, pp. 421–453, 1976.
- [61] R. R. Sinden, S. S. Broyles, and D. E. Pettijohn, "Perfect palindromic lac operator DNA sequence exists as a stable cruciform structure in supercoiled DNA in vitro but not in vivo," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 80, no. 7 I, pp. 1797–1801, 1983.
- [62] N. Panayotatos and A. Fontaine, "A native cruciform DNA structure probed in bacteria by recombinant T7 endonuclease," *Journal of Biological Chemistry*, vol. 262, no. 23, pp. 11364–11368, 1987.
- [63] M. Zannis-Hadjopoulos, L. Frappier, M. Khoury, and G. B. Price, "Effect of anti-cruciform DNA monoclonal antibodies on DNA replication," *EMBO Journal*, vol. 7, no. 6, pp. 1837–1844, 1988.
- [64] J. A. McClellan, P. Boublikova, E. Palecek, and D. M. Lilley, "Superhelical torsion in cellular DNA responds directly to environmental and genetic factors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 21, pp. 8373–8377, 1990.
- [65] M. Rosenberg and D. Court, "Regulatory sequences involved in the promotion and termination of RNA transcription," *Annual Review of Genetics*, vol. 13, pp. 319–353, 1979.
- [66] D. Lockshon and D. A. Galloway, "Sequence and structural requirements of a herpes simplex viral DNA replication origin," *Molecular and Cellular Biology*, vol. 8, no. 10, pp. 4018–4027, 1988.

- [67] K. S. Lobachev, A. Rattray, and V. Narayanan, "Hairpin- and cruciform-mediated chromosome breakage: causes and consequences in eukaryotic cells," *Frontiers in Bioscience*, vol. 12, pp. 4208–4220, 2007.
- [68] K. Mizuuchi, M. Mizuuchi, and M. Gellert, "Cruciform structures in palindromic DNA are favored by DNA supercoiling," *Journal of Molecular Biology*, vol. 156, no. 2, pp. 229–243, 1982.
- [69] K. S. Lobachev, J. E. Stenger, O. G. Kozyreva, J. Jurka, D. A. Gordenin, and M. A. Resnick, "Inverted Alu repeats unstable in yeast are excluded from the human genome," *EMBO Journal*, vol. 19, no. 14, pp. 3822–3830, 2000.
- [70] J. E. Stenger, K. S. Lobachev, D. Gordenin, T. A. Darden, J. Jurka, and M. A. Resnick, "Biased distribution of inverted and direct Alus in the human genome: implications for insertion, exclusion, and genome stability," *Genome Research*, vol. 11, no. 1, pp. 12–27, 2001.
- [71] D. R. Leach and F. W. Stahl, "Viability of λ phages carrying a perfect palindrome in the absence of recombination nucleases," *Nature*, vol. 305, no. 5933, pp. 448–451, 1983.
- [72] A. Ahmed and L. Podemski, "Observations on template switching during DNA replication through long inverted repeats," *Gene*, vol. 223, no. 1–2, pp. 187–194, 1998.
- [73] S. C. Raghavan, P. C. Swanson, X. Wu, C. L. Hsieh, and M. R. Lieber, "A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex," *Nature*, vol. 428, no. 6978, pp. 88–93, 2004.
- [74] H. Kurahashi, H. Inagaki, T. Ohye, H. Kogo, T. Kato, and B. S. Emanuel, "Chromosomal translocations mediated by palindromic DNA," *Cell Cycle*, vol. 5, no. 12, pp. 1297–1303, 2006.
- [75] H. Kurahashi, H. Inagaki, T. Ohye et al., "The constitutional t(11;22): implications for a novel mechanism responsible for gross chromosomal rearrangements," *Clinical Genetics*, vol. 78, no. 4, pp. 299–309, 2010.
- [76] H. Kurahashi, H. Inagaki, K. Yamada et al., "Cruciform DNA structure underlies the etiology for palindrome-mediated human chromosomal translocations," *Journal of Biological Chemistry*, vol. 279, no. 34, pp. 35377–35383, 2004.
- [77] K. S. Lobachev, D. A. Gordenin, and M. A. Resnick, "The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements," *Cell*, vol. 108, no. 2, pp. 183–193, 2002.
- [78] A. L. Paek, S. Kaochar, H. Jones, A. Elezaby, L. Shanks, and T. Weinert, "Fusion of nearby inverted repeats by a replication-based mechanism leads to formation of dicentric and acentric chromosomes that cause genome instability in budding yeast," *Genes and Development*, vol. 23, no. 24, pp. 2861–2875, 2009.
- [79] K. Mizuno, S. Lambert, G. Baldacci, J. M. Murray, and A. M. Carr, "Nearby inverted repeats fuse to generate acentric and dicentric palindromic chromosomes by a replication template exchange mechanism," *Genes and Development*, vol. 23, no. 24, pp. 2876–2886, 2009.
- [80] J. E. Haber, G. Ira, A. Malkova, and N. Sugawara, "Repairing a double-strand chromosome break by homologous recombination: revisiting Robin Holliday's model," *Philosophical Transactions of the Royal Society B*, vol. 359, no. 1441, pp. 79–86, 2004.
- [81] E. K. Schwartz and W. D. Heyer, "Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes," *Chromosoma*, vol. 120, no. 2, pp. 109–127, 2011.
- [82] S. C. West, "Processing of recombination intermediates by the RuvABC proteins," *Annual Review of Genetics*, vol. 31, pp. 213–244, 1997.
- [83] M. N. Boddy, P. H. Gaillard, W. H. McDonald, P. Shanahan, J. R. Yates III, and P. Russell, "Mus81-Eme1 are essential components of a Holliday junction resolvase," *Cell*, vol. 107, no. 4, pp. 537–548, 2001.
- [84] S. Fekairi, S. Scaglione, C. Chahwan et al., "Human SLX4 Is a holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases," *Cell*, vol. 138, no. 1, pp. 78–89, 2009.
- [85] J. M. Svendsen, A. Smogorzewska, M. E. Sowa et al., "Mammalian BTBD12/SLX4 assembles a holliday junction resolvase and is required for DNA repair," *Cell*, vol. 138, no. 1, pp. 63–77, 2009.
- [86] S. L. Andersen, D. T. Bergstralh, K. P. Kohl, J. R. LaRocque, C. B. Moore, and J. Sekelsky, "Drosophila MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination," *Molecular Cell*, vol. 35, no. 1, pp. 128–135, 2009.
- [87] A. Z. Al-minawi, N. Saleh-gohari, and T. Helleday, "The ERCC1/XPF endonuclease is required for efficient single-strand annealing and gene conversion in mammalian cells," *Nucleic Acids Research*, vol. 36, no. 1, pp. 1–9, 2008.
- [88] S. C. Ip, U. Rass, M. G. Blanco, H. R. Flynn, J. M. Skehel, and S. C. West, "Identification of Holliday junction resolvases from humans and yeast," *Nature*, vol. 456, no. 7220, pp. 357–361, 2008.
- [89] J. M. Svendsen and J. W. Harper, "GEN1/Yen1 and the SLX4 complex: solutions to the problem of Holliday junction resolution," *Genes and Development*, vol. 24, no. 6, pp. 521–536, 2010.
- [90] K. A. Bernstein, S. Gangloff, and R. Rothstein, "The RecQ DNA helicases in DNA repair," *Annual Review of Genetics*, vol. 44, pp. 393–417, 2010.
- [91] G. Liberi, G. Maffioletti, C. Lucca et al., "Rad51-dependent DNA structures accumulate at damaged replication forks in sgs1 mutants defective in the yeast ortholog of BLM RecQ helicase," *Genes and Development*, vol. 19, no. 3, pp. 339–350, 2005.
- [92] K. Hanada and I. D. Hickson, "Molecular genetics of RecQ helicase disorders," *Cellular and Molecular Life Sciences*, vol. 64, no. 17, pp. 2306–2322, 2007.
- [93] P. R. Prince, M. J. Emond, and R. J. Monnat Jr., "Loss of werner syndrome protein function promotes aberrant mitotic recombination," *Genes and Development*, vol. 15, no. 8, pp. 933–938, 2001.
- [94] Y. Saintigny, K. Makienko, C. Swanson, M. J. Emond, and R. J. Monnat Jr., "Homologous recombination resolution defect in Werner syndrome," *Molecular and Cellular Biology*, vol. 22, no. 20, pp. 6971–6978, 2002.
- [95] M. Payne and I. D. Hickson, "Genomic instability and cancer: lessons from analysis of Bloom's syndrome," *Biochemical Society Transactions*, vol. 37, no. 3, pp. 553–559, 2009.
- [96] A. Constantinou, M. Tarsounas, J. K. Karow et al., "Werner's syndrome protein (WRN) migrates Holliday junctions and co-localizes with RPA upon replication arrest," *EMBO Reports*, vol. 1, no. 1, pp. 80–84, 2000.
- [97] J. K. Karow, A. Constantinou, J. L. Li, S. C. West, and I. D. Hickson, "The Bloom's syndrome gene product promotes branch migration of Holliday junctions," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 12, pp. 6504–6508, 2000.

- [98] R. J. Bennett, J. L. Keck, and J. C. Wang, "Binding specificity determines polarity of DNA unwinding by the Sgs1 protein of *S. cerevisiae*," *Journal of Molecular Biology*, vol. 289, no. 2, pp. 235–248, 1999.
- [99] F. G. Harmon and S. C. Kowalczykowski, "RecQ helicase, in concert with RecA and SSB proteins, initiates and disrupts DNA recombination," *Genes and Development*, vol. 12, no. 8, pp. 1134–1144, 1998.
- [100] D. V. Bugreev, R. M. Brosh Jr., and A. V. Mazin, "RECQ1 possesses DNA branch migration activity," *Journal of Biological Chemistry*, vol. 283, no. 29, pp. 20231–20242, 2008.
- [101] P. Xie, "Model for RuvAB-mediated branch migration of Holliday junctions," *Journal of Theoretical Biology*, vol. 249, no. 3, pp. 566–573, 2007.
- [102] A. Vindigni, F. Marino, and O. Gileadi, "Probing the structural basis of RecQ helicase function," *Biophysical Chemistry*, vol. 149, no. 3, pp. 67–77, 2010.
- [103] S. A. Compton, G. Tolun, A. S. Kamath-Loeb, L. A. Loeb, and J. D. Griffith, "The Werner syndrome protein binds replication fork and Holliday junction DNAs as an oligomer," *Journal of Biological Chemistry*, vol. 283, no. 36, pp. 24478–24483, 2008.
- [104] J. K. Karow, R. H. Newman, P. S. Freemont, and I. D. Hickson, "Oligomeric ring structure of the Bloom's syndrome helicase," *Current Biology*, vol. 9, no. 11, pp. 597–600, 1999.
- [105] V. Popuri, C. Z. Bachrati, L. Muzzolini et al., "The human RecQ helicases, BLM and RECQ1, display distinct DNA substrate specificities," *Journal of Biological Chemistry*, vol. 283, no. 26, pp. 17766–17776, 2008.
- [106] Y. Liu, R. Prasad, W. A. Beard et al., "Coordination between polymerase β and FEN1 can modulate CAG repeat expansion," *Journal of Biological Chemistry*, vol. 284, no. 41, pp. 28352–28366, 2009.
- [107] P. Singh, L. Zheng, V. Chavez, J. Qiu, and B. Shen, "Concerted action of exonuclease and gap-dependent endonuclease activities of FEN-1 contributes to the resolution of triplet repeat sequences (CTG)_n- and (GAA)_n-derived secondary structures formed during maturation of Okazaki fragments," *Journal of Biological Chemistry*, vol. 282, no. 6, pp. 3465–3477, 2007.
- [108] R. M. Brosh Jr., C. von Kobbe, J. A. Sommers et al., "Werner syndrome protein interacts with human flap endonuclease 1 and stimulates its cleavage activity," *EMBO Journal*, vol. 20, no. 20, pp. 5791–5801, 2001.
- [109] S. Sharma, J. A. Sommers, L. Wu, V. A. Bohr, I. D. Hickson, and R. M. Brosh, "Stimulation of flap endonuclease-1 by the Bloom's syndrome protein," *Journal of Biological Chemistry*, vol. 279, no. 11, pp. 9847–9856, 2004.
- [110] S. Sharma, M. Otterlei, J. A. Sommers et al., "WRN helicase and FEN-1 form a complex upon replication arrest and together process branch-migrating DNA structures associated with the replication fork," *Molecular Biology of the Cell*, vol. 15, no. 2, pp. 734–750, 2004.
- [111] S. Sharma, J. A. Sommers, and R. M. Brosh Jr., "Processing of DNA replication and repair intermediates by the concerted action of RecQ helicases and Rad2 structure-specific nucleases," *Protein and Peptide Letters*, vol. 15, no. 1, pp. 89–102, 2008.
- [112] S. Sharma, J. A. Sommers, R. K. Gary, E. Friedrich-Heineken, U. Hübscher, and R. M. Brosh Jr., "The interaction site of Flap Endonuclease-1 with WRN helicase suggests a coordination of WRN and PCNA," *Nucleic Acids Research*, vol. 33, no. 21, pp. 6769–6781, 2005.
- [113] S. Sharma, J. A. Sommers, and R. M. Brosh Jr., "In vivo function of the conserved non-catalytic domain of Werner syndrome helicase in DNA replication," *Human Molecular Genetics*, vol. 13, no. 19, pp. 2247–2261, 2004.
- [114] O. Imamura and J. L. Campbell, "The human Bloom syndrome gene suppresses the DNA replication and repair defects of yeast *dna2* mutants," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 14, pp. 8193–8198, 2003.
- [115] J. D. Bartos, W. Wang, J. E. Pike, and R. A. Bambara, "Mechanisms by which bloom protein can disrupt recombination intermediates of Okazaki fragment maturation," *Journal of Biological Chemistry*, vol. 281, no. 43, pp. 32227–32239, 2006.
- [116] W. Wang and R. A. Bambara, "Human bloom protein stimulates flap endonuclease 1 activity by resolving DNA secondary structure," *Journal of Biological Chemistry*, vol. 280, no. 7, pp. 5391–5399, 2005.
- [117] K. A. Hoadley, D. Xu, Y. Xue, K. A. Satyshur, W. Wang, and J. L. Keck, "Structure and cellular roles of the RMI core complex from the bloom syndrome dissolvasome," *Structure*, vol. 18, no. 9, pp. 1149–1158, 2010.
- [118] L. Wu and I. D. Hickson, "The Bloom's syndrome helicase suppresses crossing over during homologous recombination," *Nature*, vol. 426, no. 6968, pp. 870–874, 2003.
- [119] G. Ira, A. Malkova, G. Liberi, M. Foiani, and J. E. Haber, "Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast," *Cell*, vol. 115, no. 4, pp. 401–411, 2003.
- [120] A. Schwacha and N. Kleckner, "Identification of double Holliday junctions as intermediates in meiotic recombination," *Cell*, vol. 83, no. 5, pp. 783–791, 1995.
- [121] M. Bzymek, N. H. Thayer, S. D. Oh, N. Kleckner, and N. Hunter, "Double holliday junctions are intermediates of DNA break repair," *Nature*, vol. 464, no. 7290, pp. 937–941, 2010.
- [122] D. Johnson-Schlitz and W. R. Engels, "Template disruptions and failure of double Holliday junction dissolution during double-strand break repair in *Drosophila* BLM mutants," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 45, pp. 16840–16845, 2006.
- [123] S. Raynard, W. Bussen, and P. Sung, "A double holliday junction dissolvasome comprising BLM, topoisomerase III α , and BLAP75," *Journal of Biological Chemistry*, vol. 281, no. 20, pp. 13861–13864, 2006.
- [124] L. Wu, C. Z. Bachrati, J. Ou et al., "BLAP75/RMI1 promotes the BLM-dependent dissolution of homologous recombination intermediates," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 11, pp. 4068–4073, 2006.
- [125] W. Bussen, S. Raynard, V. Busygina, A. K. Singh, and P. Sung, "Holliday junction processing activity of the BLM-Topo III α -BLAP75 complex," *Journal of Biological Chemistry*, vol. 282, no. 43, pp. 31484–31492, 2007.
- [126] P. Cejka, J. L. Plank, C. Z. Bachrati, I. D. Hickson, and S. C. Kowalczykowski, "Rmi1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3," *Nature Structural and Molecular Biology*, vol. 17, no. 11, pp. 1377–1382, 2010.
- [127] H. W. Mankouri, T. M. Ashton, and I. D. Hickson, "Holliday junction-containing DNA structures persist in cells lacking Sgs1 or Top3 following exposure to DNA damage," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 12, pp. 4944–4949, 2011.

- [128] M. D. Frank-Kamenetskii and S. M. Mirkin, "Triplex DNA structures," *Annual Review of Biochemistry*, vol. 64, pp. 65–95, 1995.
- [129] A. Jain, G. Wang, and K. M. Vasquez, "DNA triple helices: biological consequences and therapeutic potential," *Biochimie*, vol. 90, no. 8, pp. 1117–1130, 2008.
- [130] R. Zain and J. S. Sun, "Do natural DNA triple-helical structures occur and function in vivo?" *Cellular and Molecular Life Sciences*, vol. 60, no. 5, pp. 862–870, 2003.
- [131] L. D. Nelson, M. Musso, and M. W. van Dyke, "The yeast STM1 gene encodes a purine motif triple helical DNA-binding protein," *Journal of Biological Chemistry*, vol. 275, no. 8, pp. 5573–5581, 2000.
- [132] M. Musso, G. Bianchi-Scarrà, and M. W. van Dyke, "The yeast CDP1 gene encodes a triple-helical DNA-binding protein," *Nucleic Acids Research*, vol. 28, no. 21, pp. 4090–4096, 2000.
- [133] I. García-Bassets, M. Ortiz-Lombardía, S. Pagans et al., "The identification of nuclear proteins that bind the homopyrimidine strand of d(GATC)_n DNA sequences, but not the homopurine strand," *Nucleic Acids Research*, vol. 27, no. 16, pp. 3267–3275, 1999.
- [134] J. S. Lee, G. D. Burkholder, L. J. Latimer, B. L. Haug, and R. P. Braun, "A monoclonal antibody to triplex DNA binds to eucaryotic chromosomes," *Nucleic Acids Research*, vol. 15, no. 3, pp. 1047–1061, 1987.
- [135] S. C. Raghavan, P. Chastain, J. S. Lee et al., "Evidence for a triplex DNA conformation at the bcl-2 major breakpoint region of the t(14;18) translocation," *Journal of Biological Chemistry*, vol. 280, no. 24, pp. 22749–22760, 2005.
- [136] M. Ohno, T. Fukagawa, J. S. Lee, and T. Ikemura, "Triplex-forming DNAs in the human interphase nucleus visualized in situ by polypurine/polypyrimidine DNA probes and antitriplex antibodies," *Chromosoma*, vol. 111, no. 3, pp. 201–213, 2002.
- [137] S. M. Mirkin and M. D. Frank-Kamenetskii, "H-DNA and related structures," *Annual Review of Biophysics and Biomolecular Structure*, vol. 23, pp. 541–576, 1994.
- [138] S. C. Raghavan, A. Tsai, C. L. Hsieh, and M. R. Lieber, "Analysis of non-B DNA structure at chromosomal sites in the mammalian genome," *Methods in Enzymology*, vol. 409, pp. 301–316, 2006.
- [139] M. M. Seidman and P. M. Glazer, "The potential for gene repair via triple helix formation," *Journal of Clinical Investigation*, vol. 112, no. 4, pp. 487–494, 2003.
- [140] S. S. Gaddis, Q. Wu, H. D. Thames et al., "A web-based search engine for triplex-forming oligonucleotide target sequences," *Oligonucleotides*, vol. 16, no. 2, pp. 196–201, 2006.
- [141] G. P. Schroth and P. S. Ho, "Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA," *Nucleic Acids Research*, vol. 23, no. 11, pp. 1977–1983, 1995.
- [142] A. Bacolla, J. R. Collins, B. Gold et al., "Long homopurine·homopyrimidine sequences are characteristic of genes expressed in brain and the pseudoautosomal region," *Nucleic Acids Research*, vol. 34, no. 9, pp. 2663–2675, 2006.
- [143] C. Mayfield, S. Ebbinghaus, J. Gee et al., "Triplex formation by the human Ha-ras promoter inhibits Sp1 binding and in vitro transcription," *Journal of Biological Chemistry*, vol. 269, no. 27, pp. 18232–18238, 1994.
- [144] B. P. Belotserkovskii, E. de Silva, S. Tornaletti, G. Wang, K. M. Vasquez, and P. C. Hanawalt, "A triplex-forming sequence from the human c-MYC promoter interferes with DNA transcription," *Journal of Biological Chemistry*, vol. 282, no. 44, pp. 32433–32441, 2007.
- [145] A. M. Gacy, G. M. Goellner, C. Spiro et al., "GAA instability in Friedreich's Ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases," *Molecular Cell*, vol. 1, no. 4, pp. 583–593, 1998.
- [146] K. Ohshima, L. Montermini, R. D. Wells, and M. Pandolfo, "Inhibitory effects of expanded GAA·TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication in vivo," *Journal of Biological Chemistry*, vol. 273, no. 23, pp. 14588–14595, 1998.
- [147] M. M. Krasilnikova and S. M. Mirkin, "Replication stalling at Friedreich's ataxia (GAA)_n repeats in vivo," *Molecular and Cellular Biology*, vol. 24, no. 6, pp. 2286–2295, 2004.
- [148] H. P. Patel, L. Lu, R. T. Blaszak, and J. J. Bissler, "PKD1 intron 21: triplex DNA formation and effect on replication," *Nucleic Acids Research*, vol. 32, no. 4, pp. 1460–1468, 2004.
- [149] G. Wang, S. Carbajal, J. Vijg, J. DiGiovanni, and K. M. Vasquez, "DNA structure-induced genomic instability in vivo," *Journal of the National Cancer Institute*, vol. 100, no. 24, pp. 1815–1817, 2008.
- [150] B. P. Dixon, L. Lu, A. Chu, and J. J. Bissler, "RecQ and RecG helicases have distinct roles in maintaining the stability of polypurine·polypyrimidine sequences," *Mutation Research*, vol. 643, no. 1–2, pp. 20–28, 2008.
- [151] R. M. Brosh Jr., A. Majumdar, S. Desai, I. D. Hickson, V. A. Bohr, and M. M. Seidman, "Unwinding of a DNA triple helix by the Werner and Bloom syndrome helicases," *Journal of Biological Chemistry*, vol. 276, no. 5, pp. 3024–3030, 2001.
- [152] A. Machwe, L. Xiao, J. Groden, and D. K. Orren, "The Werner and Bloom syndrome proteins catalyze regression of a model replication fork," *Biochemistry*, vol. 45, no. 47, pp. 13939–13946, 2006.
- [153] A. Jain, A. Bacolla, P. Chakraborty, F. Grosse, and K. M. Vasquez, "Human DHX9 helicase unwinds triple-helical DNA structures," *Biochemistry*, vol. 49, no. 33, pp. 6992–6999, 2010.
- [154] J. A. Sommers, N. Rawtani, R. Gupta et al., "FANCI uses its motor ATPase to destabilize protein-DNA complexes, unwind triplexes, and inhibit RAD51 strand exchange," *Journal of Biological Chemistry*, vol. 284, no. 12, pp. 7505–7517, 2009.
- [155] R. Gupta, S. Sharma, J. A. Sommers, Z. Jin, S. B. Cantor, and R. M. Brosh Jr., "Analysis of the DNA substrate specificity of the human BACH1 helicase associated with breast cancer," *Journal of Biological Chemistry*, vol. 280, no. 27, pp. 25450–25460, 2005.
- [156] Y. Wu, S.-Y. Kazuo, and R. M. Brosh Jr., "FANCI helicase defective in Fanconi anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability," *Molecular and Cellular Biology*, vol. 28, no. 12, pp. 4116–4128, 2008.
- [157] M. Uhring and A. Poterszman, "DNA helicases and human diseases," *Medicine/Sciences*, vol. 22, no. 12, pp. 1087–1094, 2006.
- [158] C. G. Lee, V. da Costa Soares, C. Newberger, K. Manova, E. Lacy, and J. Hurwitz, "RNA helicase A is essential for normal gastrulation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 23, pp. 13709–13713, 1998.
- [159] S. M. Mirkin, "Discovery of alternative DNA structures: a heroic decade (1979–1989)," *Frontiers in Bioscience*, vol. 13, no. 3, pp. 1064–1071, 2008.

- [160] R. I. Mathad and D. Yang, "G-quadruplex structures and G-quadruplex-interactive compounds," *Methods in Molecular Biology*, vol. 735, pp. 77–96, 2011.
- [161] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, and S. Neidle, "Quadruplex DNA: sequence, topology and structure," *Nucleic Acids Research*, vol. 34, no. 19, pp. 5402–5415, 2006.
- [162] J. Zhou, G. Yuan, J. Liu, and C. G. Zhan, "Formation and stability of G-quadruplexes self-assembled from guanine-rich strands," *Chemistry*, vol. 13, no. 3, pp. 945–949, 2007.
- [163] A. K. Todd, M. Johnston, and S. Neidle, "Highly prevalent putative quadruplex sequence motifs in human DNA," *Nucleic Acids Research*, vol. 33, no. 9, pp. 2901–2907, 2005.
- [164] J. L. Huppert and S. Balasubramanian, "Prevalence of quadruplexes in the human genome," *Nucleic Acids Research*, vol. 33, no. 9, pp. 2908–2916, 2005.
- [165] H. J. Lipps and D. Rhodes, "G-quadruplex structures: in vivo evidence and function," *Trends in Cell Biology*, vol. 19, no. 8, pp. 414–422, 2009.
- [166] M. L. Duquette, P. Handa, J. A. Vincent, A. F. Taylor, and N. Maizels, "Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA," *Genes and Development*, vol. 18, no. 13, pp. 1618–1629, 2004.
- [167] K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, and H. J. Lipps, "Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo," *Nature Structural and Molecular Biology*, vol. 12, no. 10, pp. 847–854, 2005.
- [168] S. G. Hershman, Q. Chen, J. Y. Lee et al., "Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*," *Nucleic Acids Research*, vol. 36, no. 1, pp. 144–156, 2008.
- [169] N. Maizels, "Dynamic roles for G4 DNA in the biology of eukaryotic cells," *Nature Structural and Molecular Biology*, vol. 13, no. 12, pp. 1055–1059, 2006.
- [170] E. Kruisselbrink, V. Guryev, K. Brouwer, D. B. Pontier, E. Cuppen, and M. Tijsterman, "Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCD1-defective *C. elegans*," *Current Biology*, vol. 18, no. 12, pp. 900–905, 2008.
- [171] I. Cheung, M. Schertzer, A. Rose, and P. M. Lansdorp, "Disruption of dog-1 in *Caenorhabditis elegans* triggers deletions upstream of guanine-rich DNA," *Nature Genetics*, vol. 31, no. 4, pp. 405–409, 2002.
- [172] T. B. London, L. J. Barber, G. Mosedale et al., "FANCD1 is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts," *Journal of Biological Chemistry*, vol. 283, no. 52, pp. 36132–36139, 2008.
- [173] H. Ding, M. Schertzer, X. Wu et al., "Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein," *Cell*, vol. 117, no. 7, pp. 873–886, 2004.
- [174] X. Wu and N. Maizels, "Substrate-specific inhibition of RecQ helicase," *Nucleic Acids Research*, vol. 29, no. 8, pp. 1765–1771, 2001.
- [175] H. Sun, R. J. Bennett, and N. Maizels, "The *Saccharomyces cerevisiae* Sgs1 helicase efficiently unwinds G-G paired DNAs," *Nucleic Acids Research*, vol. 27, no. 9, pp. 1978–1984, 1999.
- [176] M. Fry and L. A. Loeb, "Human Werner syndrome DNA helicase unwinds tetrahedral structures of the fragile X syndrome repeat sequence d(CGG)_n," *Journal of Biological Chemistry*, vol. 274, no. 18, pp. 12797–12802, 1999.
- [177] H. Sun, J. K. Karow, I. D. Hickson, and N. Maizels, "The Bloom's syndrome helicase unwinds G4 DNA," *Journal of Biological Chemistry*, vol. 273, no. 42, pp. 27587–27592, 1998.
- [178] M. D. Huber, D. C. Lee, and N. Maizels, "G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition," *Nucleic Acids Research*, vol. 30, no. 18, pp. 3954–3961, 2002.
- [179] P. Mohaghegh, J. K. Karow, R. M. Brosh Jr., V. A. Bohr, and I. D. Hickson, "The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases," *Nucleic Acids Research*, vol. 29, no. 13, pp. 2843–2849, 2001.
- [180] H. Cohen and D. A. Sinclair, "Recombination-mediated lengthening of terminal telomeric repeats requires the Sgs1 DNA helicase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 6, pp. 3174–3179, 2001.
- [181] P. H. Huang, F. E. Pryde, D. Lester et al., "SGS1 is required for telomere elongation in the absence of telomerase," *Current Biology*, vol. 11, no. 2, pp. 125–129, 2001.
- [182] F. B. Johnson, R. A. Marciniak, M. McVey, S. A. Stewart, W. C. Hahn, and L. Guarente, "The *Saccharomyces cerevisiae* WRN homolog Sgs1p participates in telomere maintenance in cells lacking telomerase," *EMBO Journal*, vol. 20, no. 4, pp. 905–913, 2001.
- [183] V. Kaliraman and S. J. Brill, "Role of SGS1 and SLX4 in maintaining rDNA structure in *Saccharomyces cerevisiae*," *Current Genetics*, vol. 41, no. 6, pp. 389–400, 2002.
- [184] G. Versini, I. Comet, M. Wu, L. Hoopes, E. Schwob, and P. Pasero, "The yeast Sgs1 helicase is differentially required for genomic and ribosomal DNA replication," *EMBO Journal*, vol. 22, no. 8, pp. 1939–1949, 2003.
- [185] R. A. Marciniak, D. B. Lombard, F. B. Johnson, and L. Guarente, "Nucleolar localization of the Werner Syndrome protein in human cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 12, pp. 6887–6892, 1998.
- [186] S. L. Ding and C. Y. Shen, "Model of human aging: recent findings on Werner's and Hutchinson-Gilford progeria syndromes," *Clinical Interventions in Aging*, vol. 3, no. 3, pp. 431–444, 2008.
- [187] L. Crabbe, R. E. Verdun, C. I. Hagglblom, and J. Karlseder, "Defective telomere lagging strand synthesis in cells lacking WRN helicase activity," *Science*, vol. 306, no. 5703, pp. 1951–1953, 2004.
- [188] P. L. Opresko, M. Otterlei, J. Graakjær et al., "The werner syndrome helicase and exonuclease cooperate to resolve telomeric D loops in a manner regulated by TRF1 and TRF2," *Molecular Cell*, vol. 14, no. 6, pp. 763–774, 2004.
- [189] P. L. Opresko, C. von Kobbe, J. P. Laine, J. Harrigan, I. D. Hickson, and V. A. Bohr, "Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom syndrome helicases," *Journal of Biological Chemistry*, vol. 277, no. 43, pp. 41110–41119, 2002.
- [190] A. S. Kamath-Loeb, L. A. Loeb, E. Johansson, P. M. Burgers, and M. Fry, "Interactions between the Werner Syndrome Helicase and DNA Polymerase δ Specifically Facilitate Copying of Tetraplex and Hairpin Structures of the d(CGG)_n Trinucleotide Repeat Sequence," *Journal of Biological Chemistry*, vol. 276, no. 19, pp. 16439–16446, 2001.
- [191] A. S. Kamath-Loeb, E. Johansson, P. M. Burgers, and L. A. Loeb, "Functional interaction between the Werner syndrome protein and DNA polymerase δ ," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 9, pp. 4603–4608, 2000.
- [192] A. M. Szekely, F. Bleichert, A. Nümann et al., "Werner protein protects nonproliferating cells from oxidative DNA damage,"

- Molecular and Cellular Biology*, vol. 25, no. 23, pp. 10492–10506, 2005.
- [193] N. Selak, C. Z. Bachrati, I. Shevelev et al., “The Bloom’s syndrome helicase (BLM) interacts physically and functionally with p12, the smallest subunit of human DNA polymerase δ ,” *Nucleic Acids Research*, vol. 36, no. 16, pp. 5166–5179, 2008.
- [194] C. Ribeyre, J. Lopes, J.-B. Boulé et al., “The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo,” *PLoS Genetics*, vol. 5, no. 5, Article ID e1000475, 2009.
- [195] C. M. Sanders, “Human Pif1 helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity,” *Biochemical Journal*, vol. 430, no. 1, pp. 119–128, 2010.
- [196] J. B. Boulé, L. R. Vega, and V. A. Zakian, “The yeast Pif1p helicase removes telomerase from telomeric DNA,” *Nature*, vol. 438, no. 7064, pp. 57–61, 2005.
- [197] D. H. Zhang, B. Zhou, Y. Huang, L. X. Xu, and J. Q. Zhou, “The human Pif1 helicase, a potential *Escherichia coli* RecD homologue, inhibits telomerase activity,” *Nucleic Acids Research*, vol. 34, no. 5, pp. 1393–1404, 2006.
- [198] K. Paeschke, J. A. Capra, and V. A. Zakian, “DNA replication through g-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 dna helicase,” *Cell*, vol. 145, no. 5, pp. 678–691, 2011.
- [199] B. E. Snow, M. Mateyak, J. Paderova et al., “Murine Pif1 interacts with telomerase and is dispensable for telomere function in vivo,” *Molecular and Cellular Biology*, vol. 27, no. 3, pp. 1017–1026, 2007.
- [200] S. Sharma, D. J. Stumpo, A. S. Balajee et al., “RECQL, a member of the RecQ family of DNA helicases, suppresses chromosomal instability,” *Molecular and Cellular Biology*, vol. 27, no. 5, pp. 1784–1794, 2007.
- [201] J. Déjardin and R. E. Kingston, “Purification of proteins associated with specific genomic loci,” *Cell*, vol. 136, no. 1, pp. 175–186, 2009.
- [202] J. R. Lydeard, Z. Lipkin-Moore, S. Jain, V. V. Eapen, and J. E. Haber, “Sgs1 and exo1 redundantly inhibit break-induced replication and de novo telomere addition at broken chromosome ends,” *PLoS Genetics*, vol. 6, no. 5, Article ID e1000973, 2010.
- [203] K. M. Doherty, S. Sharma, L. A. Uzdilla et al., “RECQ1 helicase interacts with human mismatch repair factors that regulate genetic recombination,” *Journal of Biological Chemistry*, vol. 280, no. 30, pp. 28085–28094, 2005.
- [204] W. K. Chu and I. D. Hickson, “RecQ helicases: multifunctional genome caretakers,” *Nature Reviews Cancer*, vol. 9, no. 9, pp. 644–654, 2009.
- [205] J.-M. Chen, D. N. Cooper, C. Férec, H. Kehrer-Sawatzki, and G. P. Patrinos, “Genomic rearrangements in inherited disease and cancer,” *Seminars in Cancer Biology*, vol. 20, no. 4, pp. 222–233, 2010.
- [206] A. Bacolla, G. Wang, A. Jain et al., “Non-B DNA-forming sequences and WRN deficiency independently increase the frequency of base substitution in human cells,” *Journal of Biological Chemistry*, vol. 286, no. 12, pp. 10017–10026, 2011.
- [207] S. Sharma, K. M. Doherty, and R. M. Brosh Jr., “DNA helicases as targets for anti-cancer drugs,” *Current Medicinal Chemistry*, vol. 5, no. 3, pp. 183–199, 2005.
- [208] L. A. Christensen, R. A. Finch, A. J. Booker, and K. M. Vasquez, “Targeting oncogenes to improve breast cancer chemotherapy,” *Cancer Research*, vol. 66, no. 8, pp. 4089–4094, 2006.
- [209] S. Balasubramanian, L. H. Hurley, and S. Neidle, “Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?” *Nature Reviews Drug Discovery*, vol. 10, no. 4, pp. 261–275, 2011.
- [210] M. Aggarwal, J. A. Sommers, R. H. Shoemaker, and R. M. Brosh Jr., “Inhibition of helicase activity by a small molecule impairs Werner syndrome helicase (WRN) function in the cellular response to DNA damage or replication stress,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 4, pp. 1525–1530, 2011.