J|A|C|S COMMUNICATIONS

Published on Web 03/24/2010

A Unique Fluorescent Base Analogue for the Expansion of the Genetic Alphabet

Michiko Kimoto,[†,‡] Tsuneo Mitsui,[‡] Shigeyuki Yokoyama,[†,§] and Ichiro Hirao*[,†,‡]

RIKEN Systems and Structural Biology Center (SSBC) and TagCyx Biotechnologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, and Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received January 29, 2010; E-mail: ihirao@riken.jp

Fluorescent nucleobase analogues continue to be developed for use in a wide variety of biology and biotechnology tools as molecular probes and reporters for nucleic acids.[1] Here we present a novel fluorescent base analogue, 7-(2,2′-bithien-5-yl)-imidazo[4,5-*b*]pyridine (denoted as **Dss**). Through the use of an unnatural base pair system, the nucleoside triphosphates of **Dss** (d**Dss**TP and **Dss**TP) can be site-specifically incorporated into DNA and RNA by polymerases, opposite its pairing partner, pyrrole-2-carbaldehyde (**Pa**),[2] in DNA templates. Despite its high incorporation specificity in replication and transcription, **Dss** in oligonucleotides functions as a universal base that pairs with all four natural bases with nearly equal thermal stabilities. This unique **Dss** base would thus be useful for fluorescent base replacements at specific positions in DNA and RNA molecules with minimal functional disruption.

We previously reported a fluorescent base analogue, 2-amino-6-(2-thienyl)purine (**s**) (Figure 1A), for the expansion of the genetic alphabet by unnatural base pair systems.[3,4] The ribonucleoside triphosphate of **s** (**s**TP) is site-specifically incorporated into RNA opposite its pairing partners **Pa**[3] and imidazolin-2-one (**z**)[4] by T7 RNA polymerase. The fluorescent **s** base is useful for analyzing the local structural conformation of RNA molecules by tracking of fluorescence intensity changes.[3] However, the **s**−**Pa** and **s**−**z** pairings in replication are less specific than those in transcription. Therefore, we developed two other base pairs, 7-(2-thienyl)-imidazo[4,5-*b*]pyridine (**Ds**) and **Pa**[2b] and **Ds** and 2-nitro-4-propynylpyrrole (**Px**),[5] that exhibit high specificity and efficiency in PCR amplification and transcription. However, the fluorescence properties of **Ds** are significantly less effective than those of **s** (Figure 1B).

To improve its fluorescence, we modified the **Ds** base by attaching an extra thienyl group, affording the **Dss** nucleotides. Oligothiophene substitutions are known to increase fluorescence quantum yields and have been applied to oligonucleotide components as fluorophores.[6] **Dss** retained the same shape complementarity toward **Pa** as that of the **Ds**−**Pa** pair. We first synthesized the **Dss** base moiety (56% in three steps) and then coupled it with deoxyribose (45%) or ribose derivatives (23%), which were converted to the triphosphates for replication and transcription substrates or to the amidite for DNA chemical synthesis (see the Supporting Information).

In comparison with **s**, the **Dss** nucleotides are strongly fluorescent and slightly red-shifted (Figure 1B,C). The fluorescence of the **Dss** deoxyribonucleoside is characterized by excitation at 370 nm and emission at 442 nm with a quantum yield of 0.32 in ethanol. The fluorescence of **Dss** in single- and double-stranded DNA fragments was reduced slightly (see the Supporting Information), as shown previously for **s**.[3]

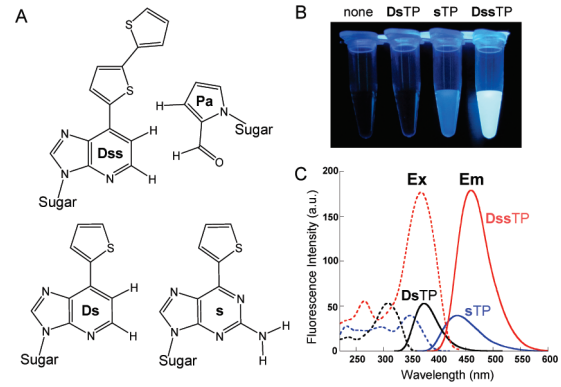

Figure 1. (A) Unnatural base pair involving fluorescent **Dss** and its pairing partner **Pa**. The **Ds** and fluorescent **s** bases also pair with **Pa**. (B) Fluorescence of **Dss**TP in comparison with that of **s**TP and **Ds**TP upon irradiation at 365 nm and (C) excitation and emission spectra of **Dss**TP, **Ds**TP, and **s**TP. Each triphosphate (10 μM) was dissolved in 10 mM phosphate buffer (pH 7.0).

The high selectivity of the **Dss**−**Pa** pairing in replication was observed by single-nucleotide insertion experiments[7] using the exonuclease-deficient Klenow fragment (KF) of *Escherichia coli* DNA polymerase I (Table 1). The incorporation efficiency of d**Dss**TP into DNA opposite **Pa** in the template was almost as high as that of dATP opposite T and 18−49-fold higher than those of

Table 1. Steady-State Kinetics of Single-Nucleotide Insertion of the Cognate **Dss**−**Pa** Pairing and Non-Cognate Pairings in Replication by KF.[a]

Entries 1-6 Primer 5′-FAM-ACTCACTATAGGGAGGAAGA
Template 3′-TATTATGCTGAGTGATATCCCTCCTTCT**N**TCTCGA

Entries 7-12 Primer 5′-FAM-ACTCACTATAGGGAGCTTCT
Template 3′-TATTATGCTGAGTGATATCCCTCGAAGA**N**AGAGCT

entry	template base (N)	nucleoside triphosphate	K_M (μM)[b]	V_{max} (% min^{-1})[b]	V_{max}/K_M (% min^{-1} M^{-1})
1	**Dss**	d**Pa**TP	150 (50)	5.9 (0.9)	3.9×10^4
2[c]	**Dss**	dATP	n.d.	n.d.	−
3[c]	**Dss**	dGTP	n.d.	n.d.	−
4[c]	**Dss**	dCTP	n.d.	n.d.	−
5[c]	**Dss**	dTTP	n.d.	n.d.	−
6	A	dTTP	0.7 (0.4)	2.8 (1.5)	4.0×10^6
7	**Pa**	d**Dss**TP	1.3 (0.6)	8.3 (3.0)	6.4×10^6
8	A	d**Dss**TP	5.0 (2.4)	0.7 (0.1)	1.5×10^5
9	G	d**Dss**TP	3.6 (0.2)	0.9 (0.03)	2.6×10^5
10	C	d**Dss**TP	7.5 (1.8)	1.0 (0.2)	1.3×10^5
11	T	d**Dss**TP	6.2 (0.1)	2.2 (0.2)	3.5×10^5
12	T	dATP	0.8 (0.4)	3.3 (1.8)	4.1×10^6

[a] See the Supporting Information for experimental details. [b] Standard deviations are given in parentheses. [c] The reaction was too slow to calculate the parameters ($V_{max} < 0.05$).

[†] RIKEN Systems and Structural Biology Center.
[‡] TagCyx Biotechnologies.
[§] The University of Tokyo.

10.1021/ja100806c

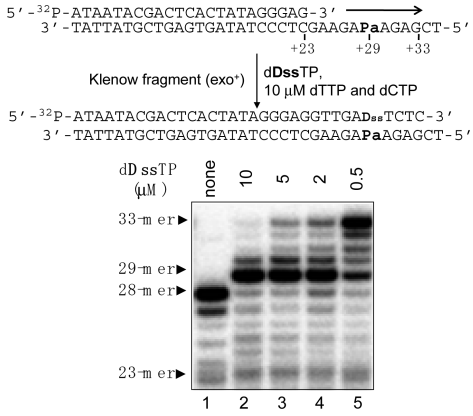

Figure 2. Primer extension involving the **Dss** incorporation opposite **Pa** by the exonuclease-proficient KF. Extension was performed at 37 °C for 5 min using the enzyme (1 unit) and 200 nM primer−template duplex.

the natural base substrates opposite **Pa**. In addition, misincorporation of the natural base substrates opposite **Dss** rarely occurred.

Primer extension involving **Dss** incorporation opposite **Pa** was examined by using the 3′ → 5′ exonuclease-proficient KF, a DNA template (35-mer) containing **Pa**, and a ^{32}P-labeled primer (Figure 2). Interestingly, primer extension including **Dss** incorporation opposite **Pa** proceeded efficiently with a lower d**Dss**TP concentration (0.5 μM d**Dss**TP and 10 μM natural base dNTPs; Figure 2, lane 5). However, equal amounts (10 μM) of d**Dss**TP and the natural dNTPs inhibited the extension after the incorporation of **Dss** at position 29 (Figure 2, lane 2). Since the primer extension was performed without dATP and dGTP, the elongation paused before position 34C of the template and yielded the 33-mer product, suggesting that no **Dss** misincorporation opposite C occurred. In addition, since a lower d**Dss**TP concentration (0.5 μM) relative to those of the natural dNTPs (10 μM) is sufficient for efficient **Dss** incorporation, **Dss** misincorporation opposite the natural bases can be eliminated.

In transcription, **Dss** was also specifically incorporated into RNA opposite **Pa** in DNA templates by T7 RNA polymerase. Transcription was performed using a conventional kit (Ampliscribe T7-Flash, Epicenter Biotechnologies) with a double-stranded DNA template (69-mer) containing **Pa** in the template strand and a triphosphate set (0.025 mM **Dss**TP and 2 mM natural base NTPs) (Figure 3). Total transcripts (52-mers) were detected by UV shadowing of the gel, and the **Dss** incorporation within the transcripts was confirmed by their fluorescence (excitation at 365 nm) on the gel. Very subtle fluorescence was observed from the transcript using the natural base template in the presence of **Dss**TP (Figure 3, lane 3), indicating

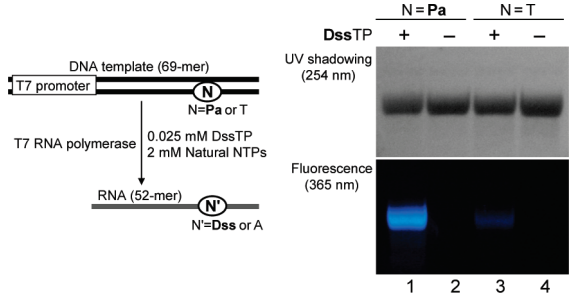

Figure 3. Dss incorporation into RNA by T7 transcription using the **Pa**-containing DNA template. Transcription was performed at 37 °C for 2 h using an Ampliscribe T7-Flash transcription kit with 200 nM DNA template.

Table 2. T_m Values for DNA Duplexes Containing **Dss**[a]

5′-GGTAACN$_1$ATGCG-3′
3′-CCATTGN$_2$TACGC-5′

N_1−N_2	T_m (°C)	N_1−N_2	T_m (°C)
Dss−**Pa**	44.8	**Dss**−A	45.3
Dss−G	45.9	**Dss**−C	44.9
Dss−T	43.9	A−T	48.6

[a] Determined in 100 mM NaCl, 10 mM sodium phosphate (pH 7.0), and 0.1 mM EDTA with 5 μM DNA duplex.

that the **Dss** misincorporation opposite the natural bases (~0.2% misincorporation per base) was limited under these conditions.

Duplex formation involving **Dss** was thermodynamically unique. In contrast to the enzymatic specificity of the **Dss**−**Pa** pairing, **Dss** pairs with all four natural bases as well as **Pa** with equal thermal stabilities (T_m = 43.9−45.9 °C) in duplex DNAs (12-mers) (Table 2). The T_m values of the 12-mer duplexes containing these cognate **Dss**−**Pa** and noncognate **Dss**−natural base pairs were 3−5 °C lower than that containing the A−T pair (T_m = 48.6 °C). Thus, **Dss** acts as a universal base for duplex formation in addition to the specific pairing between **Dss** and **Pa** in replication and transcription.

In summary, a strongly fluorescent base analogue, **Dss**, has been developed. The triphosphates d**Dss**TP and **Dss**TP can be site-specifically incorporated into DNA and RNA opposite **Pa** in templates by replication and transcription. In addition, the **Dss** base functions as a universal base and would be useful for fluorescent labeling of the duplex and any region of nucleic acids by replacement of the natural bases. As we previously demonstrated with the fluorescent **s** base,[3] **Dss** could also be applied to local structural analyses of functional nucleic acids through detection of fluorescence quenching by stacking with neighboring bases. Applications of the site-specific **Dss** labeling of DNA and RNA molecules with long chains (>100-mer) are in progress.

Acknowledgment. This work was supported by Grants-in-Aid for Scientific Research (KAKENHI 19201046 to I.H., 20710176 to M.K.) and by the Targeted Proteins Research Program and the RIKEN Structural Genomics/Proteomics Initiative from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Supporting Information Available: Details of **Dss** chemical synthesis and physical and biochemical experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Asseline, U. *Curr. Org. Chem.* **2006**, *10*, 491. (b) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265. (c) Zhao, Y.; Knee, J. L.; Baranger, A. M. *Biorg. Chem.* **2008**, *36*, 271. (d) Krueger, A. T.; Kool, E. T. *J. Am. Chem. Soc.* **2008**, *130*, 3989. (e) Börjesson, K.; Preus, S.; El-Sagheer, A. H.; Brown, T.; Albinsson, B.; Wilhelmsson, L. M. *J. Am. Chem. Soc.* **2009**, *131*, 4288. (f) Xie, Y.; Dix, A. V.; Tor, Y. *J. Am. Chem. Soc.* **2009**, *131*, 17605. (g) Zhao, L.; Xia, T. *Methods* **2009**, *49*, 128.

(2) (a) Mitsui, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S. *J. Am. Chem. Soc.* **2003**, *125*, 5298. (b) Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. *Nat. Methods* **2006**, *3*, 729.

(3) Kimoto, M.; Mitsui, T.; Harada, Y.; Sato, A.; Yokoyama, S.; Hirao, I. *Nucleic Acids Res.* **2007**, *35*, 5360.

(4) Hirao, I.; Harada, Y.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Yokoyama, S. *J. Am. Chem. Soc.* **2004**, *126*, 13298.

(5) Kimoto, M.; Kawai, R.; Mitsui, T.; Yokoyama, S.; Hirao, I. *Nucleic Acids Res.* **2009**, *37*, e14.

(6) (a) Garcia, P.; Pernaut, J. M.; Hapiot, P.; Wintgens, V.; Valat, P.; Garmier, F.; Delabouglise, D. *J. Phys. Chem.* **1993**, *97*, 513. (b) Strässler, C.; Davis, N. E.; Kool, E. T. *Helv. Chim. Acta* **1999**, *82*, 2160.

(7) (a) Goodman, M. F.; Creighton, S.; Bloom, L. B.; Petruska, J. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 83. (b) Kimoto, M.; Yokoyama, S.; Hirao, I. *Biotechnol. Lett.* **2004**, *26*, 999.

JA100806C