Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron 63 (2007) 3528–3537
Tetrahedron

Characterization of fluorescent, unnatural base pairs

Tsuneo Mitsui,[a] Michiko Kimoto,[a] Rie Kawai,[a] Shigeyuki Yokoyama[a,b,c,*] and Ichiro Hirao[a,*]

[a]*Protein Research Group, RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan*
[b]*Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[c]*RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan*

Received 7 June 2006; revised 10 November 2006; accepted 10 November 2006
Available online 3 February 2007

Abstract—Site-specific fluorescent labeling of DNA and RNA molecules is an important tool for investigations in many areas, including analytical and material sciences and medical applications. We report herein an unnatural base pair system between fluorescent base analogs—2-amino-6-(2-thienyl)purine (**s**) and 2-amino-6-(2-thiazolyl)purine (**v**), and their fluorophore-linked pairing partners of 2-oxopyridine (**y**) derivatives—5-(3-amino-1-propynyl)-**y** linked with 5-carboxyfluorescein (FAM-**y**) and 5-[3-(6-aminohexanamido)-1-propynyl]-**y** linked with FAM (FAM-hx-**y**). These fluorescent nucleoside 5′-triphosphates are enzymatically incorporated into RNA at desired positions by T7 transcription using the unnatural base pair system. The fluorescent **s** and **v** bases in nucleic acids are useful for analyzing the local structural dynamics, and their combinations with FAM-**y** and FAM-hx-**y** provide a unique fluorescence resonance energy transfer (FRET) system.

1. Introduction

Fluorescent labeling of nucleic acids has become a powerful tool for analyzing the dynamics of their higher-ordered structures and for genetic analyses and diagnostics.[1] The intensity of a fluorescent residue at a specific position in DNA and RNA molecules is sensitively altered, depending on the base sequences and structural changes, by responding to other adjacent fluorophores and quenchers. For example, a fluorescent base analog, 2-aminopurine (AP), in a double-stranded DNA fragment facilitates the detection of netropsin binding, which causes strong quenching of the AP fluorescence.[2] Site-specific incorporation of AP into a trans-activator responsive region (TAR) of HIV-1 is useful for evaluating potential TAR inhibitors; the binding of the potential inhibitors to the TAR sequence containing AP increases the fluorescence intensity.[3] Furthermore, molecular beacons by fluorescence resonance energy transfer (FRET) are also useful; a hairpin DNA containing a fluorophore and a quencher at both termini increases the fluorescence intensity upon duplex formation with the complementary strand.[4] A more complex example is the detection of the stepwise movement of an aminoacyl-tRNA into the ribosome by single-molecule fluorescence resonance energy transfer (smFRET), using the fluorescently-labeled tRNA and ribosome.[5] Thus, the interest in novel nucleotide derivatives with higher fluorescence sensitivity and selectivity has increased,[6] allowing with the hope that fluorescent nucleotides would facilitate DNA/RNA-based fluorescence technology.

For the fluorescent labeling of nucleic acids, there are two classes of nucleotide analogs: one includes fluorescent base analogs, such as AP, and the other is fluorophore-linked nucleotides, in which fluorescent residues are attached to base or sugar moieties in the nucleotides.[7] The intensity of the fluorescent base analogs in single-stranded DNA and RNA is greatly affected by the stacking of neighboring bases and by duplex formation with the complementary strands, and thus they are useful for local structural analyses of nucleic acids.[8] In contrast, the fluorophore-linked nucleotides are rarely affected by adjacent bases. Thus, they are better suited to function as molecular beacons, and various fluorescent residues with different properties are available.[9] These two classes of fluorescent labeling characteristics permit a wide range of applications, and have expanded the demand for novel fluorescent nucleotide derivatives and new methods for the site-specific fluorescent labeling of nucleic acids, especially large RNA molecules.

An attractive method for site-specific labeling is the expansion of the genetic alphabet by unnatural base pairs, which would facilitate the enzymatic incorporation of fluorescent nucleotide derivatives into DNA and RNA molecules at desired positions by replication and transcription. A quencher-linked unnatural base, which can be site-specifically and enzymatically incorporated into DNA, has been applied to

* Corresponding authors. Tel.: +81 45 503 9644; fax: +81 45 503 9645; e-mail: ihirao@riken.jp

0040–4020/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tet.2006.11.096

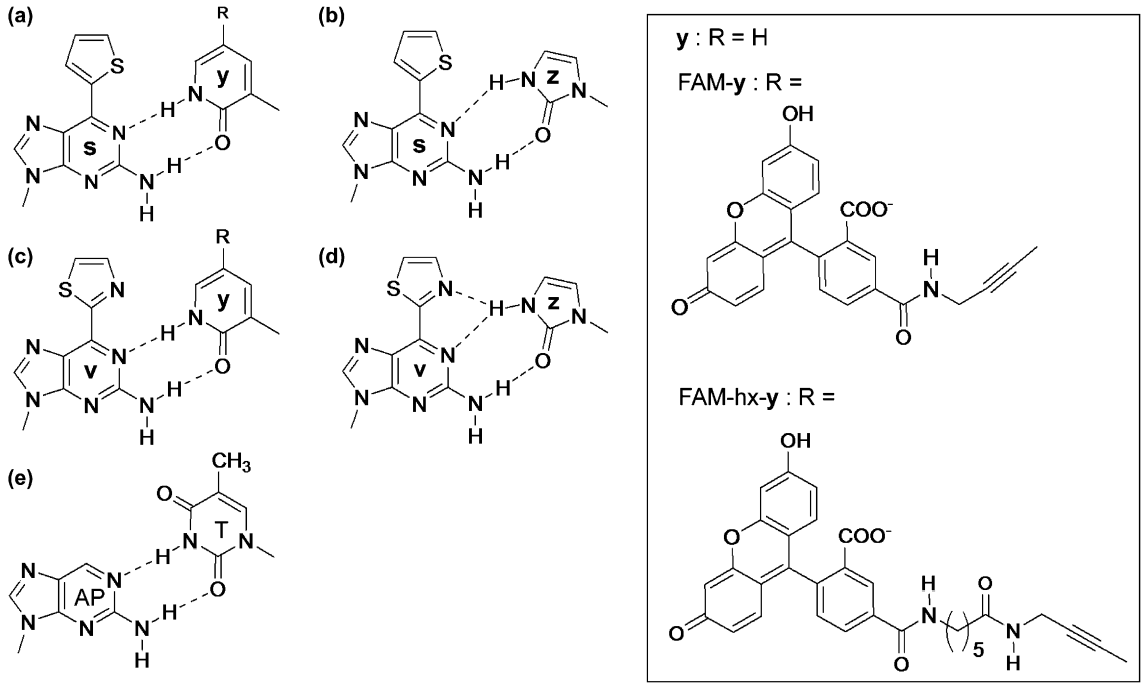

Figure 1. The unnatural base pairs: (a) **s**–**y**, (b) **s**–**z**, (c) **v**–**y**, (d) **v**–**z**, and (e) the AP–T base pair. The structures of the fluorescent groups attached to **y** at position 5 are enclosed.

the simultaneous detection and identification of multiple targets.[10] Recently, we developed several unnatural base pairs, such as 2-amino-6-(2-thienyl)purine (**s**) and 2-oxopyridine (**y**),[11] 2-amino-6-(2-thiazolyl)purine (**v**) and **y**,[12] and **s** and imidazolin-2-one (**z**)[13] (Fig. 1). These unnatural base pairs function in transcription using T7 RNA polymerase. The substrates of **y** and 5-modified **y** bases, such as biotin- and fluorophore-linked **y**, can be site-specifically incorporated into RNA opposite to **s** or **v** in DNA templates by T7 transcription.[14,15] The DNA templates can be prepared by solid-phase synthesis using the phosphoramidite units of the unnatural bases. We chemically synthesized a modified unnatural base substrate of 5-(3-amino-1-propynyl)-**y** linked with 5-carboxyfluorescein (FAM-**y**TP) (Fig. 1), and incorporated FAM-**y** site-specifically into a theophylline-binding RNA aptamer (41-mer). This sensor aptamer could detect concentrations of ~500 nM theophylline, as revealed by an increase in the fluorescence intensity upon the theophylline binding.[15]

In addition to the fluorophore-linked **y** bases, the pairing partners, the **s** and **v** bases, act as fluorescent base analogs themselves. For example, the nucleoside of **s** exhibits a fluorescence emission centered at 430 nm, characterized by two major excitation maxima (299 and 352 nm), and its fluorescence quantum yield was 0.41 at pH 7.0.[13] Thus, the unnatural base pairs between **s** and FAM-**y**, and **v** and FAM-**y**, represent unique base pairs comprising two classes of fluorescent nucleotide analogs—the fluorescent base analogs and the fluorophore-linked bases. Furthermore, the FAM-**y**TP substrate exhibits a fluorescence emission centered at 521 nm, which is characterized by an absorption maximum (493 nm), close to the emission maxima of **s** and **v**, and thus FRET is expected to be observed between **s** and FAM-**y**. Besides the site-specific enzymatic incorporation of FAM-**y** into RNA, the substrate of **s** can also be incorporated into RNA opposite to **z** in templates by T7 transcription mediated by the **s**–**z** pair.[13] Therefore, the fluorescent base pairs between **s** and FAM-**y**, and **v** and FAM-**y**, could be potentially powerful tools for a wide range of applications in nucleic acid-based technology. Here, we report the characterization of the fluorescence properties of **s** and **v** in oligonucleotides. In addition, we also describe the unique FRET systems between **s** and FAM-**y**, and **v** and FAM-**y**.

2. Results and discussion

2.1. Fluorescent properties of s and v

First, we characterized the fluorescent base analogs, **s** and **v**, by comparing them to the well characterized fluorescent base analog, AP. The UV absorption and fluorescence properties of the deoxyribonucleosides of **s**, **v**, and AP[16] are listed in Table 1. The excitation and emission spectra of these nucleosides in their free forms, as well as in single- and double-stranded DNA fragments, which were chemically synthesized with their phosphoramidite units,[11,14] are shown in Figure 2. Since the **s** and **v** bases correspond to 6-modified

Table 1. Spectral properties of deoxyribonucleosides of **s**, **v**, and AP, and triphosphates of FAM-**y**TP and FAM-hx-**y**TP

Nucleoside or triphosphate	Absorption maxima (nm)	Emission maximum (nm)	Quantum yield
s	297, 348	434	0.41
v	297, 359	461	0.46
AP	304	370	0.68[16]
FAM-**y**TP[15]	493	521	0.67
FAM-hx-**y**TP[15]	493	522	0.55

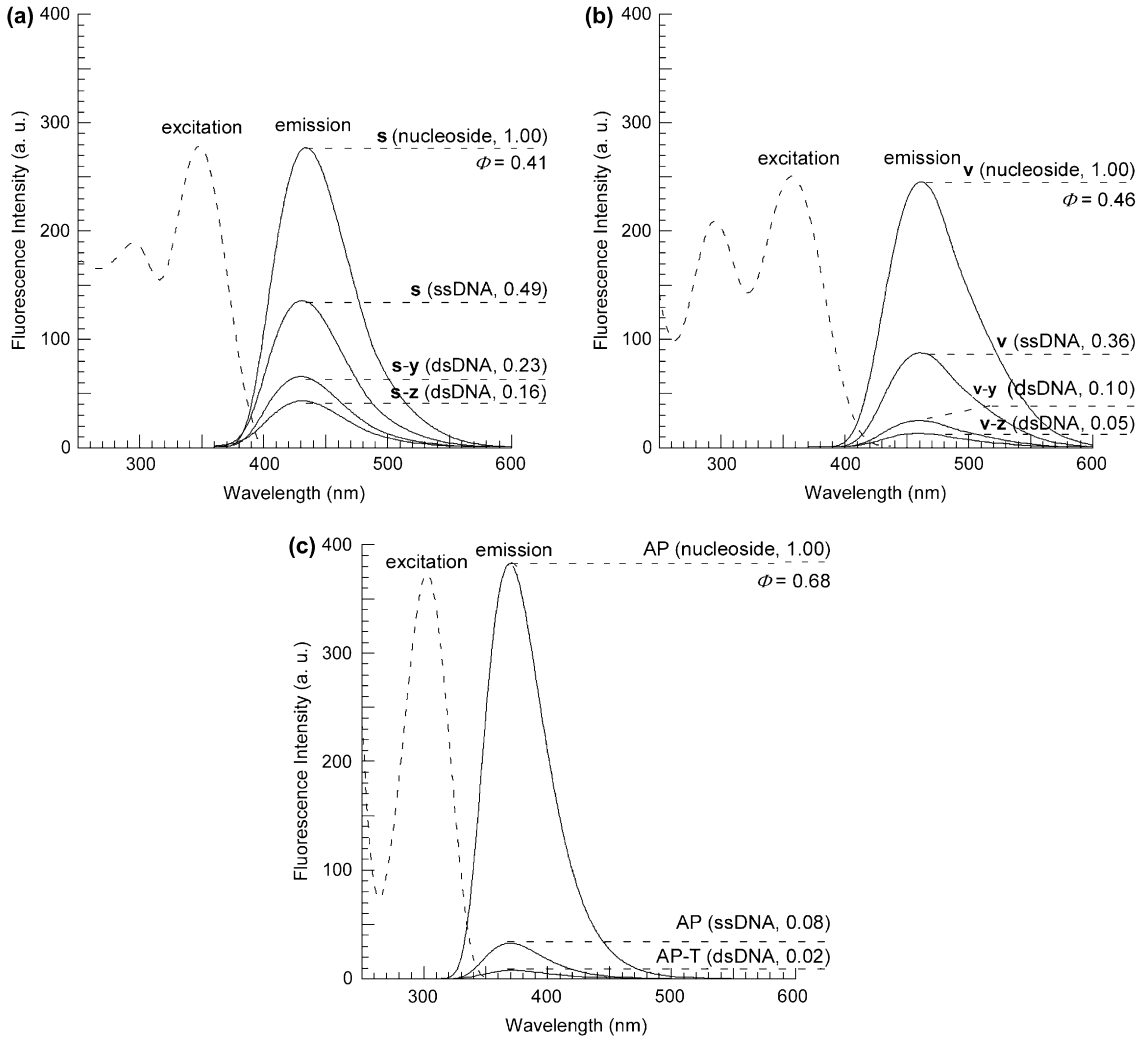

Figure 2. Steady-state fluorescent excitation (dashed line) and emission (solid line) spectra of deoxyribonucleosides and single-stranded or double-stranded DNA fragments (ssDNA or dsDNA) containing **s** (a), **v** (b), and AP (c). For the excitation spectra, the emission wavelength was monitored at 434 nm (**s**), 461 nm (**v**), and 370 nm (AP), while for the emission spectra, the excitation wavelength was set at 352 nm (**s**), 363 nm (**v**), and 303 nm (AP). The profiles were measured at 25 °C, using 5 μM deoxyribonucleosides and DNA fragments. The sequences of the DNA fragments used were 5′-GGTAACNATGCG-3′ (N=**s**, **v**, or AP) and 5′-CGCATN′GTTACC-3′ (N′=**y**, **z** or T).

AP analogs, the heterocyclic fluorophores, such as the thienyl and thiazolyl groups, at position 6 would be expected to endow AP with large Stokes shifts. In fact, both the excitation and emission centers of **s** and **v** were shifted to longer wavelengths (Fig. 2), relative to those of AP, and revealed fluorescence in the visible ranges for both **s** and **v** (Fig. 3). Although the fluorescence quantum yields of the free nucleosides of **s** and **v** were somewhat lower than that of AP, the quenching of their fluorescence intensities in the single- or double-stranded DNA fragments was less than that of AP (Figs. 2 and 3). In the single-stranded DNA fragment 5′-d(GGTAACNATGCG)-3′ (where N=**s**, **v** or AP), the fluorescence intensity of AP was drastically reduced, to 8%, but those of **s** and **v** were reduced to 49% and 36%, respectively. Similarly, the fluorescence intensity of AP, pairing with T in the duplex, was also extremely reduced to 2%. In contrast, **s** and **v**, pairing with **y** in the duplex, still retained their fluorescence with 23% and 10% intensities, respectively. This may be because the protruding thienyl and thiazolyl moieties in the **s** and **v** bases are less likely to stack with the neighboring bases in the strands, relative to the AP moiety. This strong

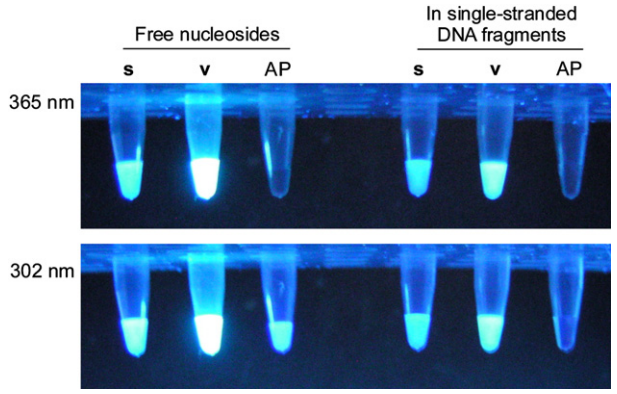

Figure 3. Comparison of the fluorescence for the deoxyribonucleosides of **s**, **v**, and AP (free nucleoside) and the single-stranded DNA fragments (5′-GGTAACNATGCG-3′, N=**s**, **v**, or AP). The samples (10 μM) were dissolved in a buffer containing 10 mM sodium phosphate (pH 7.0), 100 mM NaCl, and 0.1 mM EDTA, and their solutions were illuminated with a 365 or 302 nm transilluminator.

fluorescence in the strands is useful for detecting and analyzing fluorescently-labeled nucleic acids (Fig. 3).

We next examined the fluorescence alterations of double-stranded DNA fragments containing unnatural bases, by thermal melting. As shown in Figure 4, the changes in the fluorescence intensity of the double-stranded DNA fragments (12-mer) containing the **s**–**y**, **v**–**y**, **s**–**z**, and **v**–**z** pairs in the thermal denaturation from 20 to 60 °C were significantly larger than that containing the AP–T pair. The changes showed a two-state transition, similar to that of their UV melting profiles, and the melting temperature (T_{mf}) could be determined from the fluorescence profiles. Table 2 lists the melting temperatures (T_m) obtained by the UV hyperchromicity and T_{mf} values of the double-stranded 12-mer DNA fragments containing cognate or non-cognate unnatural base pairs, involving **s**, **v**, and AP. Most of the T_{mf} values are similar to the T_m values. However, the T_{mf} values might reflect the local structural changes in the duplexes, while the T_m values correspond to the midpoint of the temperature at which 50% of the duplexes are denatured. Thus, the T_{mf} values of some duplexes slightly differ from their T_m values. For example, the T_{mf} value of the duplex containing the AP–**y** pair was 2.6 °C lower than its T_m value. This difference indicates that the AP–**y** pair is the fragile part of the duplex, which becomes unstable at around 42.0 °C, and then the whole structure denatures at around 44.6 °C. This instability of the AP–**y** pair might be caused by the weaker stacking ability of **y**, as compared to that of T. The **y** base is a T analog, but it lacks the 4-keto and 5-methyl groups of T, and its stacking ability with the neighboring bases is weaker than that of T. Consequently, the stability of the AP–**y** pair in the duplex is also lower than that of the AP–T pair, although both **y** and T are capable of pairing with AP, through two hydrogen bonds. Thus, the lower T_{mf} value of the AP–**y** duplex reflects the weaker stacking ability of **y**. From this viewpoint, the higher T_{mf} value of the **s**–A duplex, relative to its T_m value, indicates that the combination of **s** and A increases the stacking stability at this position, but this non-cognate pair between the large purine bases might cause distortion, and thus did not contribute to the stability of the entire duplex structure.

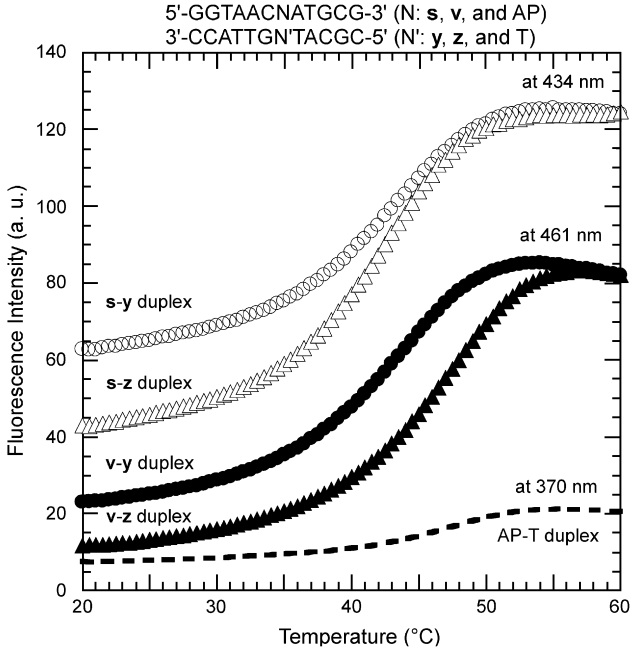

Figure 4. Fluorescence-monitored melting profiles of the DNA duplexes containing an **s**–**y** pair (open circles, at 434 nm), an **s**–**z** pair (open triangles, at 434 nm), a **v**–**y** pair (filled circles, at 461 nm), a **v**–**z** pair (filled triangles, at 461 nm), and an AP–T pair (dashed line, at 370 nm).

Table 2. T_m and T_{mf} values of duplex DNA fragments

5′-GGTAAC**N1**ATGCG-3′, 3′-CCATTG**N2**TACGC-5′			
N1–N2	T_m[a] (°C)	T_{mf}[b] (°C)	$T_{mf}-T_m$ (°C)
s–y	43.1	43.5	0.4
s–z	42.5	43.0	0.5
s–A	43.1	45.5	2.4
s–G	44.7	45.0	0.3
s–C	41.5	41.5	0.0
s–T	40.1	41.0	0.9
v–y	44.3	44.0	−0.3
v–z	48.1	47.5	−0.6
v–A	43.2	43.5	0.2
v–G	44.8	45.0	0.2
v–C	42.5	42.0	−0.5
v–T	40.2	41.5	1.3
AP–**y**	44.6	42.0	−2.6
AP–**z**	40.0	40.5	0.5
AP–A	38.6	38.5	−0.1
AP–G	37.2	37.0	−0.2
AP–C	41.1	41.0	−0.1
AP–T	47.6	47.5	−0.1
A–T	49.3	—	—

[a] The T_m values were determined by the UV melting profiles at 260 nm.
[b] The T_{mf} values were determined by the melting profiles of the fluorescence intensity alterations for **s** (at 434 nm), **v** (at 461 nm), and AP (at 370 nm).

As for the development of unnatural base pairs for expanding the genetic alphabet, the duplex DNA fragment containing the **v**–**z** pair displayed high stability in terms of both the T_m (48.1 °C) and T_{mf} values (47.5 °C), and the T_m value was marginally lower than that of the DNA fragment containing the natural A–T pair at the same position (49.3 °C). In contrast, the **s**–**z** duplex did not show such high stability. These results suggest that the nitrogen in the thiazolyl group of **v** electrostatically interacts with the N–H moiety of **z**. This high stability indicates that the **v**–**z** pair could be a potential extra base pair candidate.

2.2. Incorporation of s into a thermally stable, mini-hairpin DNA

To examine the potential of **s** as a fluorescent probe, we chemically synthesized a thermally stable DNA hairpin, a GNA mini-hairpin, containing **s**. Short DNA fragments with GCGNAGC sequences, where N is A, G, C or T, form an extraordinarily stable mini-hairpin structure, consisting of a two G–C paired stem and a small GNA loop[17,18] (Fig. 5a). For example, the T_m value of the d(GCGAAGC) fragment is 76 °C in a 0.1 M NaCl solution. NMR structural analysis of the d(GCGNAGC) fragments revealed that a shared G–A pair forms in the GNA loop, and that the fragment is folded back between the fourth N and the fifth A.[17,18] In addition, the NMR studies suggested that the fourth N is stacked with the third G. The stacking interaction might affect the stability of the mini-hairpin, and thus the thermal stability of the GAA-loop mini-hairpin is higher than that

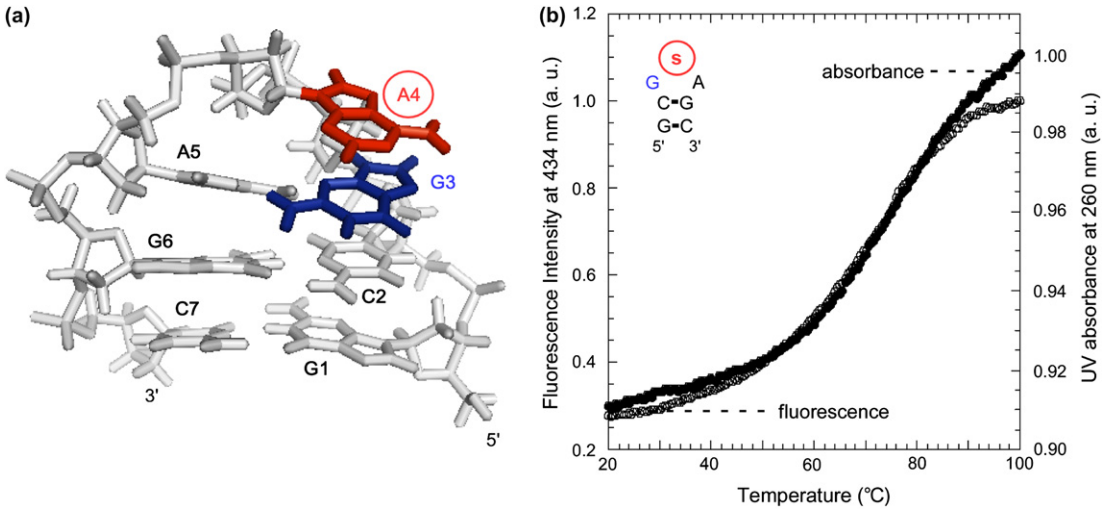

Figure 5. The three-dimensional structure of d(GCGAAGC)[17] (a) and the melting curves obtained by the changes in the fluorescence intensity at 434 nm (open circles) and the UV absorbance at 260 nm (filled circles) of d(GCG**s**AGC) (b). The A4 position, which was substituted by **s**, is circled and shown in red. The stacked base, G3, is shown in blue.

of the GTA-loop mini-hairpin.[18] However, very little is known about the dynamics of the stacking between the fourth N and the third G. Therefore, to study the stacking interactions and the mobility of the fourth N in the loop, we introduced the fluorescent **s** base into the fourth N position, and analyzed the fluorescence alteration of the d(GCG**s**AGC) fragment upon heat denaturation. The stacking between the third G and the fourth **s** could be monitored by the quenching of the fluorescence.

The thermodynamic melting profiles of the fluorescence emission at 434 nm of **s** (excited at 352 nm) and the UV absorbance at 260 nm of the d(GCG**s**AGC) fragment were measured (Fig. 5b). In both profiles, single transition points were observed at around 75 °C, indicating that the d(GCG**s**AGC) fragment also formed the mini-hairpin structure, with a stability similar to that of the d(GCGAAGC) mini-hairpin. The fluorescence intensity was efficiently quenched by the stacking with the third G in the hairpin structure, and the intensity increased with temperature, as a result of the unstacking. The T_{mf} value, which was as high as the T_m value, revealed that the fourth **s** tightly stacks with the third G in the G**s**A loop structure, and that the stability of the stacking is closely linked to the stability of the entire mini-hairpin structure.

2.3. Fluorescence resonance energy transfer (FRET) in the unnatural base pairs

Next, we examined the possibility of the existence of a FRET system between the fluorescent base analog, **s** or **v**, and the fluorophore-linked **y**, in DNA/RNA duplexes. As shown in the fluorescence profiles of the DNA fragments containing **s** or **v** (Figs. 2 and 6a and Table 1), their emission spectra overlapped with the absorption spectra of the nucleotide derivatives of 5-(3-amino-1-propynyl)-**y** linked with FAM (FAM-**y**TP) or 5-[3-(6-aminohexanamido)-1-propynyl]-**y** linked with FAM (FAM-hx-**y**TP). Thus, **s** or **v** in DNA fragments acts as a fluorescent donor, and FAM-**y** or FAM-hx-**y** in the complementary RNA fragments acts as a fluorescent acceptor. RNA fragments (17-mer) containing FAM-**y** or FAM-hx-**y** at a specific position were prepared by transcription mediated by the **v**–**y** pair, using T7 RNA polymerase, a **v**-containing DNA template (DNA35**v**1), and the FAM-**y**TP or FAM-hx-**y**TP substrate[15] (Fig. 6b). The transcripts were hybridized with DNA fragments containing one or two **s** or **v** bases (DNA35**s**1, 35**s**2a, 35**s**2b, 35**v**1, 35**v**2a, and 35**v**2b), and the emission spectra of the DNA/RNA duplexes were measured at excitation wavelengths of 350 nm for **s** and 360 nm for **v**, at different temperatures (20 and 70 °C) (Fig. 6b).

As expected, the resonance energy transfer was observed between a low-fluorescence state of **s** or **v** and a high-fluorescence state of FAM-**y** or FAM-hx-**y**. As shown in Figure 7a and b, FRET occurred when DNA35**s**1 or DNA35**v**1 formed a duplex with transcripts containing FAM-**y** or FAM-hx-**y** at 20 °C, and FRET was significantly reduced when the duplex was denatured at 70 °C. In addition, FRET in the duplexes of DNA35**s**2a and DNA35**v**2a, with transcripts containing FAM-**y** or FAM-hx-**y** also occurred effectively (Fig. 7c and d). These phenomena are very interesting, because the fluorescence intensity of the single-stranded DNA35**s**2a or DNA35**v**2a, containing two adjacent unnatural bases, was considerably reduced (Fig. 6c and d). Since no significant reduction of the fluorescence intensity was observed with the single-stranded DNA35**s**2b and DNA35**v**2b fragments, which contain two unnatural bases but separated by one T (Fig. 6c and d), the reduction of the fluorescence intensity of the single-stranded DNA35**s**2a and DNA35**v**2a was caused by self-quenching between the two adjacent unnatural bases. Even in this case, the strong FAM fluorescence of the transcripts was observed by the duplex formation with DNA35**s**2a or DNA35**v**2a (Fig. 7c and d). This FRET system, using the self-quenched probes, would facilitate the use of excess amounts of the probes relative to the target strands. When 10 mol equiv of DNA35**s**1 and DNA35**v**1 relative to the transcripts were used, the resonance energy transfer was completely hidden in the background fluorescence of DNA35**s**1

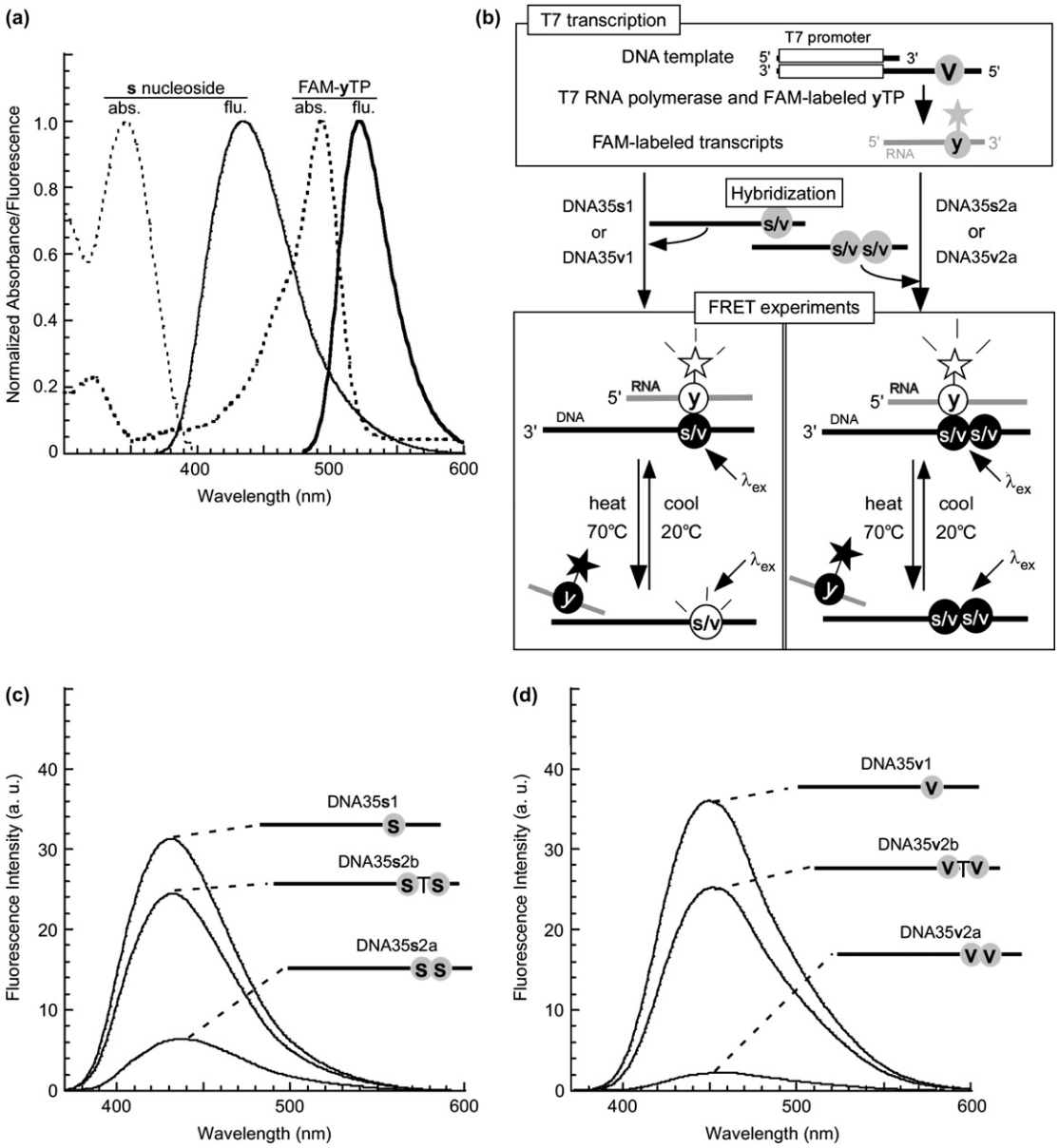

Figure 6. A FRET system using the unnatural base pairs. (a) Normalized fluorescence (flu, solid line) and absorbance (abs, dashed line) of the **s** deoxyribonucleoside (thin lines) and FAM-**y**TP (thick lines) in phosphate buffer (pH 7.0). (b) Schemes of the FRET experiments. RNA fragments (transcripts) containing a FAM-linked **y** at a specific position (FAM-labeled transcripts) were prepared by T7 transcription mediated by the **v**–**y** pair. The FAM-labeled transcripts were hybridized with their complementary DNA fragments containing one **s** or **v** (DNA35**s**1 or DNA35**v**1) and two adjacent **s** or **v** bases (DNA35**s**2a or DNA35**v**2a). The unnatural bases are shown with filled circles (black) in a low-fluorescence state and with open circles (white) in a high-fluorescence state, when excited for **s** or **v** (λ_{ex}). (c and d) Steady-state fluorescent emission spectra of the single-stranded DNA fragments containing **s** bases (c: DNA35**s**1, DNA35**s**2a, and DNA35**s**2b) and containing **v** bases (d: DNA35**v**1, DNA35**v**2a, and DNA35**v**2b). The DNA35**s**2b fragment contains an **sTs** sequence and DNA35**v**2b contains a **vTv** sequence. The sequences of the fragments are shown in Section 4. The excitation wavelengths for **s**- and **v**-containing DNA fragments are 350 nm (c) and 360 nm (d), respectively. The measurements were carried out with a 100 nM concentration of each DNA fragment at 20 °C.

and DNA35**v**1, respectively (Fig. 8a and b). In contrast, with the use of DNA35**s**2a and DNA35**v**2a with 10 mol equiv relative to the transcripts, FRET between the respective **ss**- and **vv**-containing DNA fragments and the FAM-containing transcripts was clearly observed (Fig. 8c and d). Therefore, large amounts of the self-quenched DNA probes can increase the FRET efficiency by facilitating the duplex formation with the complementary fluorescent acceptor strands.

The FRET efficiency between FAM-linked **y** and the fluorescent base analog, **s** or **v**, depended on the difference in the linker moieties of FAM-linked **y**. When hybridized with DNA35**s**1 or DNA35**v**1, the FRET intensity of the FAM-**y** in the transcript was larger than that of the FAM-hx-**y** (Fig. 7a and b). In contrast, the intensity of the FAM-hx-**y** in transcript hybridized with DNA35**s**2a, DNA35**s**2b, or DNA35**v**2b was larger than that of the FAM-**y** (Fig. 7c, e, and f). Among these DNA fragments, the combination of FAM-hx-**y** and DNA35**s**2a showed the highest FRET efficiency (Figs. 7c and 8c). This suggests that the flexible hexanamido-linker is suited for putting the FAM residue of **y** close to the second, neighboring **s** in the complementary strand. For more detailed discussion, the structural analysis of the duplexes is in progress.

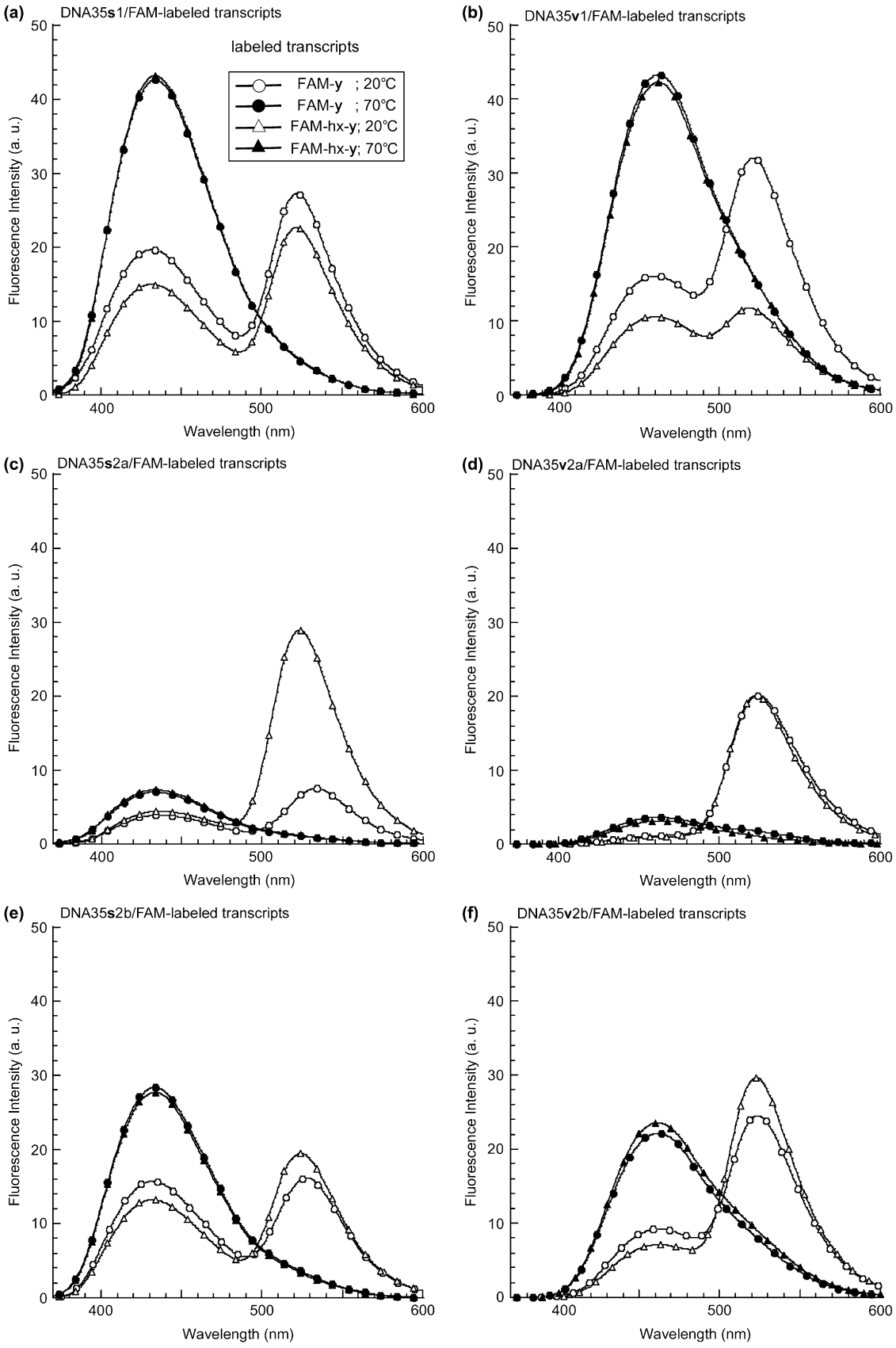

Figure 7. Steady-state fluorescent emission spectra of DNA/RNA duplexes. Each 35-mer DNA fragment (a, DNA35**s**1; b, DNA35**v**1; c, DNA35**s**2a; d, DNA35**v**2a; e, DNA35**s**2b; f, DNA35**v**2b) was hybridized with the 17-mer transcripts (RNA) containing FAM-**y** (open and filled circles) or FAM-hx-**y** (open and filled triangles). The fluorescent emission spectra of the duplexes (100 nM) were recorded with an excitation wavelength of 350 nm (a, c, and e) or 360 nm (b, d, and f) at 20 °C (open circles and open triangles) or 70 °C (filled circles and filled triangles). The sequences of the transcripts are shown in Section 4. To obtain clear FRET profiles, the background spectra of the transcripts were subtracted from those of the duplexes.

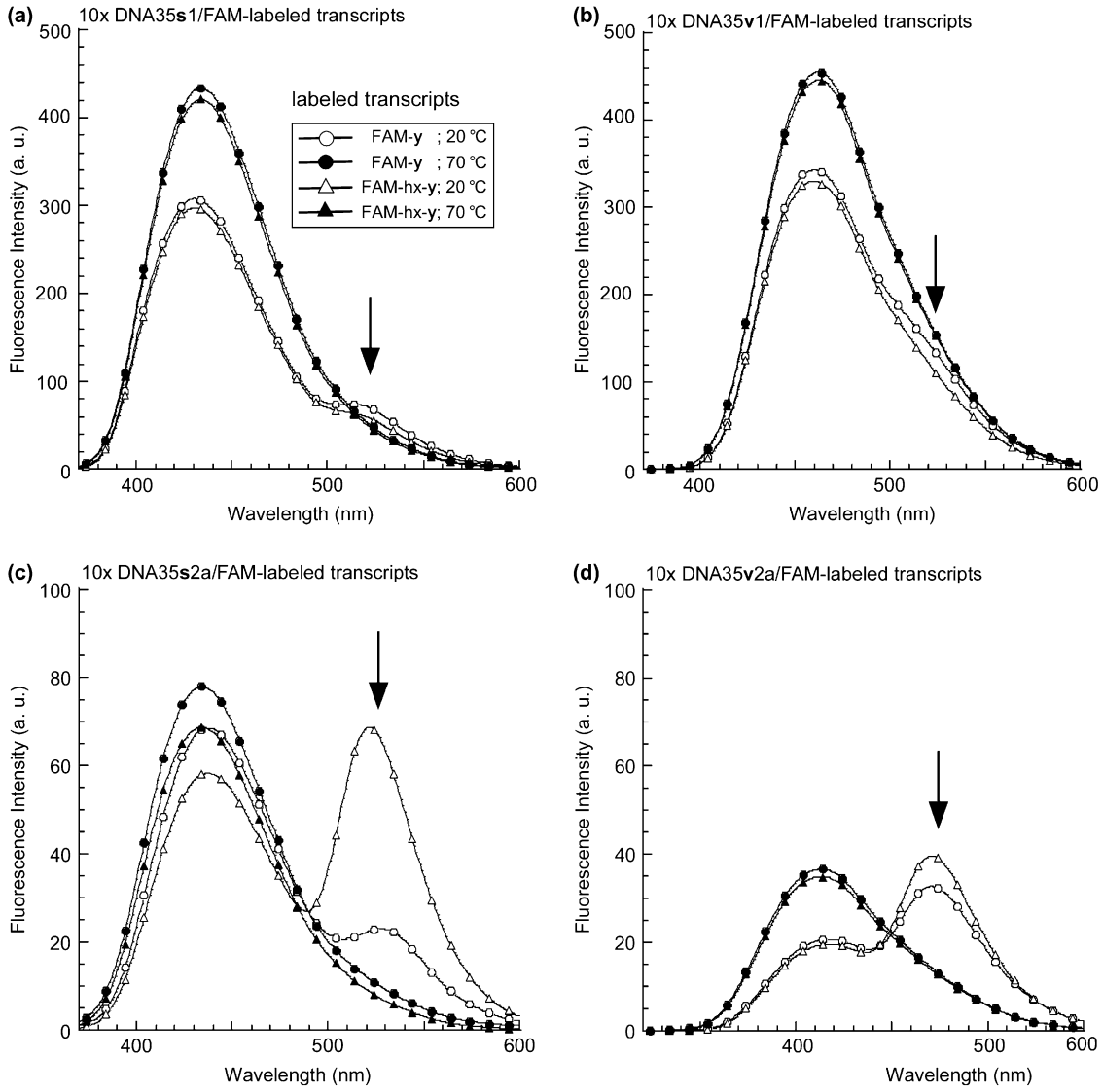

Figure 8. Steady-state fluorescent emission spectra of the mixture of the FAM-labeled transcripts (100 nM) (FAM, open and filled circles; FAM-hx, open and filled triangles) with the DNA fragments (1 μM) (a, DNA35**s**1; b, DNA35**v**1; c, DNA35**s**2a; and d, DNA35**v**2a) at 20 °C (open circles and open triangles) or 70 °C (filled circles and filled triangles). To obtain clear FRET profiles, the background spectra of the transcripts were subtracted from those of the mixture.

3. Conclusion

We examined the fluorescent properties of **s** and **v** to evaluate their potentials as fluorescent probes for nucleic acids. The quenching of the **s** and **v** fluorescence in DNA fragments is smaller than that of AP, and the fluorescence intensity is sensitively altered, depending on the environment, such as stacking with neighboring bases. Thus, the **s** and **v** bases are useful as the fluorescent probes for local structural analyses of nucleic acids and their complexes with other molecules. In addition, both **s** and **v** form base pairs with fluorophore-linked **y** bases in double-stranded DNA/RNA fragments, and FRET can be observed between the unnatural bases in the pair. In particular, the unique FRET system, using the self-quenched probes containing two adjacent **s** or **v** bases, can reduce the background of the fluorescent donor and allow detection of the complementary strands containing the fluorescent acceptor, even in the presence of a large amount of the donor probe. Furthermore, the unnatural base pair systems using the **s**–**y**, **v**–**y**, and **s**–**z** pairs enable the site-specific enzymatic incorporation of **s** and fluorophore-linked **y** bases into RNA by T7 transcription.[11–15] Thus, the fluorescent unnatural base pair system could be applied to large RNA molecules and represents a powerful tool for a wide range of applications in DNA/RNA-based technology.

4. Experimental

4.1. General

The deoxyribonucleosides and their phosphoramidites of **s**, **v**, **y**, and **z** were prepared as previously described.[11–13] The deoxyribonucleoside and phosphoramidite of AP were purchased from standard suppliers (Berry & Associates, Inc. and Glen Research, respectively). DNA fragments were synthesized with an Applied Biosystems 392 DNA synthesizer. The short DNA fragments were purified with a Gilson HPLC

system, using an analytical column (Synchropac RPP, 250× 4.6 mm, Eichrom Technologies). The fluorescent and UV–vis spectra were obtained on a Jasco FP-6500 spectrofluorometer equipped with a temperature controller (ETC-273T) and on a Beckmann DU-530 spectrometer, respectively. UV-monitored thermal denaturation experiments were performed on a Shimadzu UV-2450 spectrometer equipped with a temperature controller (TMSPC-8).

4.2. Preparation of oligodeoxynucleotides

The sequences of the chemically synthesized DNA fragments were 5′-GGTAACNATGCG-3′ (12-mer; N=A, **s**, **v**, and AP), 5′-CGCATN′GTTACC-3′ (12-mer; N′=A, T, G, C, **y**, and **z**), 5′-GCG**s**AGC-3′ (7-mer), 5′-CAN$_1$N$_2$N$_3$CTCGGGA TTCCCTATAGTGAGTCGTATTAT-3′ (35-mer; N$_1$N$_2$N$_3$= CT**s** or CT**v** for DNA35**s**1 or DNA35**v**1, C**ss** or C**vv** for DNA35**s**2a or DNA35**v**2a, and **s**T**s** or **v**T**v** for DNA35**s**2b or DNA35**v**2b), and 5′-ATAATACGACTCACTATAG GG-3′ (21-mer). The DNA fragments were purified by HPLC (the 12- and 7-mers) or by gel electrophoresis (the 35- and 21-mers).

4.3. T7 transcription for 17-mer RNA fragments containing a fluorophore-linked y at a specific position

RNA fragments (17-mer, 5′-GGGAAUCCCGAGN′AGUG-3′), containing 5-(3-amino-1-propynyl)- or 5-[3-(6-aminohexamido)-1-propynyl]-**y** linked with 5-carboxyfluorescein (FAM-**y** or FAM-hx-**y**) at the N′ position, were prepared by an unnatural base pair transcription system,[15] by using T7 RNA polymerase and the substrate of FAM-**y** or FAM-hx-**y** (FAM-**y**TP or FAM-hx-**y**TP). The template (10 μM of DNA35**v**1 and the 21-mer fragment) was annealed in a buffer containing 10 mM Tris–HCl (pH 7.6) and 10 mM NaCl, by heating at 95 °C and slow cooling to 4 °C. Transcription was performed in a buffer containing 40 mM Tris–HCl (pH 8.0), 24 mM $MgCl_2$, 2 mM spermidine, 5 mM DTT, and 0.01% Triton X-100, in the presence of 1 mM natural NTPs, 1 mM FAM-**y**TP or FAM-hx-**y**TP, 2 μM template, and 2.5 U/μL T7 RNA polymerase (Takara). After incubation at 37 °C for 3 h, the reaction was quenched by adding an equivalent volume of a dye solution, containing 10 M urea and 0.05% BPB. The reaction mixtures were heated at 75 °C for 3 min, and the 17-mer transcripts were purified by gel electrophoresis.

4.4. Steady-state spectroscopy of the deoxyribonucleosides and the oligonucleotides

Fluorescent and UV spectra were measured in a buffer containing 10 mM sodium phosphate (pH 7.0), 100 mM NaCl, and 0.1 mM EDTA, or as otherwise indicated. The emission spectra were recorded, with excitation wavelengths of 352 nm for **s** (350 nm in FRET experiments), 363 nm for **v** (360 nm in FRET experiments), and 303 nm for AP, and the excitation spectra were recorded with emission wavelengths of 434 nm for **s**, 461 nm for **v**, and 370 nm for AP. The spectral bandpass for both excitation and emission was 3 nm (5 nm in FRET experiments). Fluorescent and UV melting profiles of the oligonucleotides were recorded at a heating rate of 0.5 °C min^{-1}. For baseline correction when calculating the melting temperature values from the fluorescent intensity, we independently determined the temperature dependences of the fluorescence for the 12-mer single-stranded DNA fragments containing **s**, **v**, and AP (for the duplex DNA fragments) and the deoxyribonucleoside of **s** (for GCG**s**AGC). The T_m values were calculated by the first derivative of the melting curves, using the IgorPro software (WaveMetrics Inc.). The molar extinction coefficients for the deoxyribonucleosides were ε_{254} (8600 M^{-1} cm^{-1}), ε_{297} (9900 M^{-1} cm^{-1}), and ε_{347} (14,000 M^{-1} cm^{-1}) for **s**, ε_{298} (8800 M^{-1} cm^{-1}) and ε_{355} (9600 M^{-1} cm^{-1}) for **v**, and ε_{303} (7100 M^{-1} cm^{-1}) for AP.

Acknowledgements

This work was funded by the RIKEN RSGI, the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan, by a Grant-in-Aid for Scientific Research (KAKENHI 17750152 and 15350097), and by the Program of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NIBIO).

References and notes

1. Walter, N. G.; Burke, J. M. *Methods Enzymol.* **2000**, *317*, 409–440; Klostermeier, D.; Millar, D. P. *Methods* **2001**, *23*, 240–254; DeJong, E. S.; Luy, B.; Marino, J. P. *Curr. Top. Med. Chem.* **2002**, *2*, 289–302; Rist, M. J.; Marino, J. P. *Curr. Org. Chem.* **2002**, *6*, 775–793; Epstein, J. R.; Biran, I.; Walt, D. R. *Anal. Chim. Acta* **2002**, *469*, 3–36; Kool, E. T. *Acc. Chem. Res.* **2002**, *35*, 936–943; Mollova, E. T. *Curr. Opin. Chem. Biol.* **2002**, *6*, 823–828; Rajendran, M.; Ellington, A. D. *Nucleic Acids Res.* **2003**, *31*, 5700–5713; Johansson, M. K.; Cook, R. M. *Chem.—Eur. J.* **2003**, *9*, 3466–3471; Tan, W.; Wang, K.; Drake, T. J. *Curr. Opin. Chem. Biol.* **2004**, *8*, 547–553; Dirks, R. W.; Tanke, H. J. *BioTechniques* **2006**, *40*, 489–496.
2. Patel, N.; Berglund, H.; Nilsson, L.; Rigler, R.; McLaughlin, L. W.; Graslund, A. *Eur. J. Biochem.* **1992**, *203*, 361–366.
3. Bradrick, T. D.; Marino, J. P. *RNA* **2004**, *10*, 1459–1468.
4. Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996**, *14*, 303–308; Zhang, P.; Beck, T.; Tan, W. *Angew. Chem., Int. Ed.* **2001**, *40*, 402–405; Silverman, A. P.; Kool, E. T. *Trends Biotechnol.* **2005**, *23*, 225–230.
5. Blanchard, S. C.; Gonzalez, R. L.; Kim, H. D.; Chu, S.; Puglisi, J. D. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1008–1014; Blanchard, S. C.; Kim, H. D.; Gonzalez, R. L.; Puglisi, J. D.; Chu, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12893–12898; Dorywalska, M.; Blanchard, S. C.; Gonzalez, R. L.; Kim, H. D.; Chu, S.; Puglisi, J. D. *Nucleic Acids Res.* **2005**, *33*, 182–189.
6. Driscoll, S. L.; Hawkins, M. E.; Balis, F. M.; Pfleiderer, W.; Laws, W. R. *Biophys. J.* **1997**, *73*, 3277–3286; Hawkins, M. E.; Pfleiderer, W.; Jungmann, O.; Balis, F. M. *Anal. Biochem.* **2001**, *298*, 231–240; Cuppoletti, A.; Cho, Y.; Park, J.-S.; Strässler, C.; Kool, E. T. *Bioconjugate Chem.* **2005**, *16*, 528–534; Gaied, N. B.; Glasser, N.; Ramalanjaona, N.; Beltz, H.; Wolff, P.; Marquet, R.; Burger, A.; Mély, Y. *Nucleic Acids Res.* **2005**, *33*, 1031–1039; Kim, S. J.; Kool, E. T. *J. Am. Chem. Soc.* **2006**, *128*, 6164–6171; Tinsley, R. A.; Walter, N. G. *RNA* **2006**, *12*, 522–529.

7. Walter, N. G.; Harris, D. A.; Pereira, M. J. B.; Rueda, D. *Biopolymers* **2002**, *61*, 224–242.
8. Holz, B.; Klimasauskas, S.; Serva, S.; Weinhold, E. *Nucleic Acids Res.* **1998**, *26*, 1076–1083; Stivers, J. T. *Nucleic Acids Res.* **1998**, *26*, 3837–3844; Jean, J. M.; Hall, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 37–41; Liu, C.; Martin, C. T. *J. Mol. Biol.* **2001**, *308*, 465–475; Purohit, V.; Grindley, N. D. F.; Joyce, C. M. *Biochemistry* **2003**, *42*, 10200–10211.
9. Yu, H.; Chao, J.; Patek, D.; Mujumdar, R.; Mujumdar, S.; Waggoner, A. S. *Nucleic Acids Res.* **1994**, *22*, 3226–3232; Glazer, A. N.; Mathies, R. A. *Curr. Opin. Biotechnol.* **1997**, *8*, 94–102; Cremo, C. R. *Methods Enzymol.* **2003**, *360*, 128–177; Crey-Desbiolles, C.; Ahn, D.-R.; Leumann, C. J. *Nucleic Acids Res.* **2005**, *33*, e77.
10. Sherrill, C. B.; Marshall, D. J.; Moser, M. J.; Larsen, C. A.; Daudé-Snow, L.; Jurczyk, S.; Shapiro, G.; Prudent, J. R. *J. Am. Chem. Soc.* **2004**, *126*, 4550–4556; Prudent, J. R. *Expert Rev. Mol. Diagn.* **2006**, *6*, 245–252.
11. Fujiwara, T.; Kimoto, M.; Sugiyama, H.; Hirao, I.; Yokoyama, S. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2221–2223; Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nat. Biotechnol.* **2002**, *20*, 177–182.
12. Mitsui, T.; Kimoto, M.; Harada, Y.; Yokoyama, S.; Hirao, I. *J. Am. Chem. Soc.* **2005**, *127*, 8652–8658.
13. Hirao, I.; Harada, Y.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Yokoyama, S. *J. Am. Chem. Soc.* **2004**, *126*, 13298–13305.
14. Kimoto, M.; Endo, M.; Mitsui, T.; Okuni, T.; Hirao, I.; Yokoyama, S. *Chem. Biol.* **2004**, *11*, 47–55; Endo, M.; Mitsui, T.; Okuni, T.; Kimoto, M.; Hirao, I.; Yokoyama, S. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2593–2596; Moriyama, K.; Kimoto, M.; Mitsui, T.; Yokoyama, S.; Hirao, I. *Nucleic Acids Res.* **2005**, *33*, e129.
15. Kawai, R.; Kimoto, M.; Ikeda, S.; Mitsui, T.; Endo, M.; Yokoyama, S.; Hirao, I. *J. Am. Chem. Soc.* **2005**, *127*, 17286–17295.
16. Ward, D. C.; Reich, E.; Stryer, L. *J. Biol. Chem.* **1969**, *244*, 1228–1237.
17. Hirao, I.; Kawai, G.; Yoshizawa, S.; Nishimura, Y.; Ishido, Y.; Watanabe, K.; Miura, K. *Nucleic Acids Res.* **1994**, *22*, 576–582.
18. Yoshizawa, S.; Kawai, G.; Watanabe, K.; Miura, K.; Hirao, I. *Biochemistry* **1997**, *36*, 4761–4767.