A highly fluorescent nucleoside analog based on thieno[3,4-*d*]pyrimidine senses mismatched pairing†

Seergazhi G. Srivatsan, Haim Weizman and Yitzhak Tor*

Received 21st January 2008, Accepted 22nd February 2008
First published as an Advance Article on the web 10th March 2008
DOI: 10.1039/b801054d

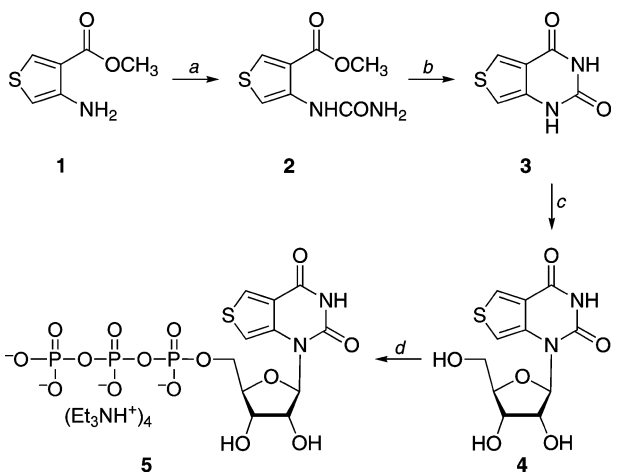
A highly emissive nucleobase analog, based on a thieno[3,4-*d*]pyrimidine core, is enzymatically incorporated into RNA oilgonucleotides that function as base discriminating fluorescent probes.

A fundamentally simple yet practically challenging task in nucleic acid diagnostics is the detection of single nucleotide polymorphism (SNP).[1] Such mutated base pairs, while in many cases associated with benign biological outcomes, could also be linked to certain diseases or reflect susceptibility to specific therapeutics.[2] Whereas several approaches for the detection of SNPs have been developed and commercialized,[3,4] efforts have recently been made to advance base-discriminating fluorescent (BDF) nucleosides, a term coined by Saito *et al.*[5] Upon hybridization to a given target, such emissive modified nucleobases can, in principle, identify the paired base on the opposite strand by eliciting a specific change in their photophysical characteristics.[6] Such fluorescence-based tools can facilitate the future development of enzyme-free and homogenous SNP detection protocols. They present, however, a considerable challenge in nucleobase probe design, as our ability to correlate structure and microenvironment with photophysical properties is currently limited.[7]

Our program aims at the development of simple and minimally perturbing emissive nucleobases. The primary design principle entails maintaining the highest possible structural similarity to the natural nucleobases, while bestowing useful photophysical properties upon the probes. Of particular significance is their ability to respond to and report on changes in their microenvironment. Several useful nucleobase analogs have been obtained by either conjugating or fusing five-membered aromatic heterocyclic rings onto the pyrimidines at their 5 or 5,6 positions, respectively.[8] Conjugating furan to a dU core, for example, has generated an emissive analog that responds to the presence of abasic sites in DNA *via* enhanced emission.[9] The analogous furan containing ribonucleoside has been shown to serve as a useful probe for monitoring RNA–ligand binding.[10] Here we report on a new motif for highly emissive nucleobase analogs based on a thieno[3,4-*d*]pyrimidine core. We report the synthesis of the ribonucleoside, its basic photophysical characteristics, and its conversion into a triphosphate, as well as its enzymatic incorporation into RNA oligonucleotides. We then illustrate the ability of the emissive RNA strand to photophysically distinguish between perfect and mismatched base pairing upon hybridization to DNA oligonucleotides.

Short RNA constructs were selected in this case as fluorescent hybridization probes as such oligonucleotides tend to form stable heteroduplexes with DNA, frequently of higher stability than the corresponding DNA–DNA duplexes.[11] Although solid-phase synthesis is a powerful tool for the site specific modification of nucleic acids, we have chosen transcription reactions as a method of choice for incorporating the fluorescent nucleotide into short RNA oligonucleotides due to the following reasons: (a) the one step preparation of the necessary triphosphate is simple, not requiring extensive protection of the nucleoside, (b) transcription reactions provide the desired RNA product in one step, eliminating the lengthy deprotection steps that follow typical solid-phase oligoribonucleotide synthesis; (c) reasonably large quantities of RNA can be synthesized by transcription reactions with minuscule amounts of a triphosphate; and (d) several modified constructs can easily be made by changing the transcription template.

Thieno[3,4-*d*]pyrimidine **3** was synthesized in two steps from commercially available starting materials (Scheme 1). Methyl 3-aminothiophene-4-carboxylate hydrochloride (**1**) was treated with KOCN to yield the corresponding urea, which cyclized to the corresponding pyrimidine **3** upon treatment with sodium methoxide in methanol.[12] The emissive heterocycle **3** was converted into the ribonucleoside **4** using a standard glycosylation



Scheme 1 Synthesis of thieno[3,4-*d*]pyrimidine-based emissive nucleoside and its triphosphate. *Reagents*: (*a*) KOCN, aq. acetic acid, RT, 81%; (*b*) NaOMe, MeOH, RT; (*c*) (i) *N*,*O*-bis(trimethylsilyl) acetamide, TMSOTf, CH_3CN, RT, 57% (steps *b* and *c*); (ii) NH_4OH, dioxane, 60 °C, 84%; (*d*) (i) $POCl_3$, $(MeO)_3PO$, 0–4 °C; (ii) tributylammonium pyrophosphate, Bu_3N, 0–4 °C, 56%.

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, 92093-0358, USA. E-mail: ytor@ucsd.edu; Fax: 858 534 0202; Tel: 858 534 6401

† Electronic supplementary information (ESI) available: Details of synthesis and characterization of compounds, transcription reactions and gel shift experiments, and additional fluorescence and thermal melting data. See DOI: 10.1039/b801054d

procedure[13] employing 1-*O*-acetyl-2,3,5-tri-*O*-Bz-ribofuranoside and TMSOTf, followed by deprotection with aqueous ammonia.

The triphosphate **5**, required for enzymatic RNA synthesis, was then prepared in a reasonable yield by treating **4** with $POCl_3$ in trimethylphosphate followed by tributylammonium pyrophosphate (Scheme 1). Ion exchange chromatography followed by HPLC purification provided the analytically pure triphosphate suitable for transcription reactions.

The sensitivity of the nucleoside's photophysical characteristics to changes in its microenvironment was evaluated prior to its incorporation into RNA oligonucleotides (Table 1). Nucleoside **4** displays two major absorption maxima (260 and 304 nm), which exhibit minimal sensitivity to solvent polarity (Fig. 1). When an aqueous solution of the nucleoside is excited at 304 nm, a strong emission is observed at 412 nm (Fig. 1), with a quantum yield of 0.48 ± 0.05.[12,14] In contrast to the ground state absorption profiles, the emission maximum (λ_{em}) and intensity are considerably affected by solvent polarity (Fig. 1, Table 1). As solvent polarity is decreased from water to methanol, and finally acetonitrile, significant hypsochromic shift (26 nm) and hypochromic effect (>3-fold) are observed. Excellent linear correlation between emission energy and $E_T(30)$, Reichardt's microscopic solvent polarity parameter,[15] is observed (Figure S1†), suggesting that the highly emissive nucleoside is responsive to its microenvironment, a necessary property for a useful nucleobase probe.[16]

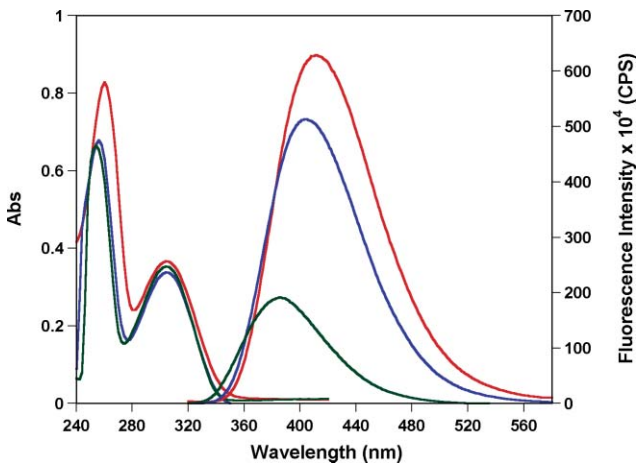


Fig. 1 Absorption (100 μM) and emission (5.0 μM) spectra of nucleoside **4** in water (red), methanol (blue) and acetonitrile (green). Excitation wavelength was 304 nm, and excitation and emission slit widths were 3 and 5 nm, respectively. Stock solutions for absorption and emission spectra contained 10% and 0.5% DMSO, respectively.

Diverse interactions with neighboring bases can affect the photophysical characteristics of an emissive nucleoside upon incorporation into oligonucleotides.[17] To preliminarily evaluate the impact of the native nucleobases on the excited state of the new emissive nucleoside, nucleoside monophosphates (AMP, TMP, GMP and CMP) were titrated into a buffered solution of nucleoside **4** and emission spectra were taken.[12] Stern–Volmer plots illustrate that GMP exhibits greater quenching than TMP and AMP with quenching constants (K_{sv}) of 0.3, 0.1 and 0.02 mM^{-1}, respectively (Fig. 2). CMP displays a nonlinear behavior, but significant quenching, comparable to that of GMP, is observed at high concentrations.

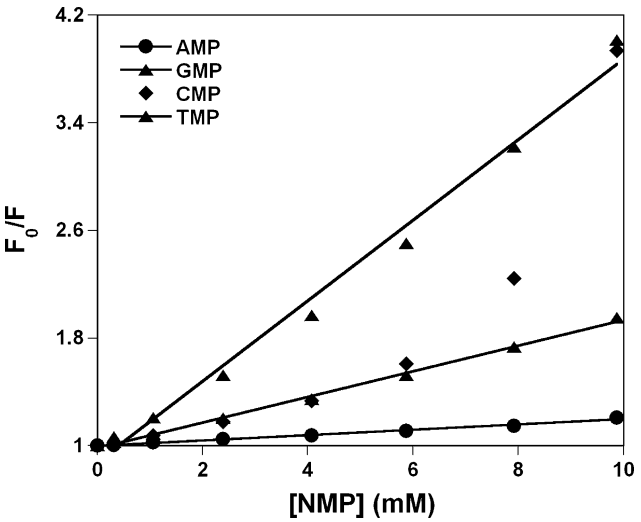


Fig. 2 Steady-state Stern–Volmer plot for the titration of nucleoside **4** with AMP, GMP, CMP and TMP.

As noted above,[10,18] enzymatic incorporation of modified nucleosides could be a powerful approach for the synthesis of useful unnatural RNA constructs. To explore if the modified nucleoside is suitable for enzymatic incorporation, transcription reactions with T7 RNA polymerase and triphosphate **5**, were performed. A duplex DNA template was assembled by annealing the 18-mer T7 consensus promoter to a synthetic DNA template **6** (Fig. 3).[19] This template contains a unique dA residue to direct the incorporation of the modified U analog. It also contains a unique T residue at the 5′-end. When performed in the presence of α-^{32}P ATP, a successful transcription reaction would result in the 3′-end labeling of the RNA transcript, which could be resolved on a denaturing sequencing polyacrylamide gel and visualized. Truncated RNA transcripts would remain undetected in failed transcription reactions. A phosphorimage of the PAGE resolved transcription reaction with template **6**, showed the formation of full length 10-mer RNA product **7** (Fig. 4, lane 2), corresponding to an efficiency of 51 ± 6% relative to an unmodified RNA

Table 1 Photophysical data of thiophene-modified uridine **4**[a]

Solvent	λ_{max}[b]/nm	λ_{em}/nm	I_{rel}[c]	$E_T(30)$/Kcal mol^{-1}	Energy/cm^{-1}
Water	304	412	3.3	63.10	2.43×10^4
Methanol	304	404	2.9	55.40	2.48×10^4
Acetonitrile	304	386	1.0	45.60	2.59×10^4

[a] Conditions for absorption and emission spectra: 100 and 5.0 μM, respectively. [b] The lowest energy maximum is given. [c] Relative emission intensity with respect to intensity in acetonitrile.

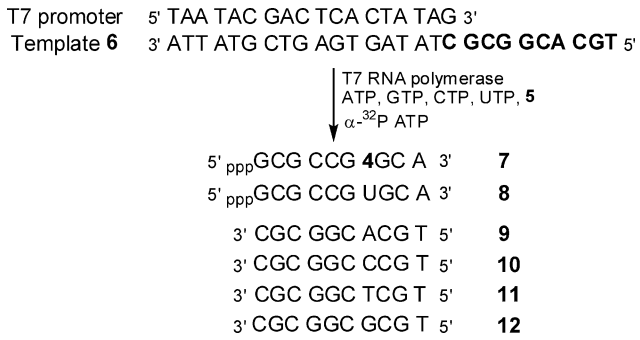


Fig. 3 Enzymatic incorporation of thiophene-modified ribonucleoside triphosphate **5**. DNA template **6** annealed to 18mer consensus T7 promoter and corresponding RNA transcripts (**7** and **8**). Also shown are synthetic complementary and mismatch deoxyribooligonucleotides (**9**–**12**) used in this study.

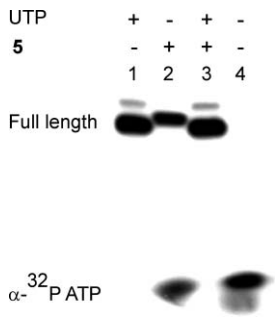


Fig. 4 Transcription reaction to study the incorporation efficiency of thiophene-modified triphosphate **5**.

transcript **8**, formed in the presence of natural NTPs only (Fig. 4, compare lanes 1 and 2). Slower migration of the modified RNA transcript **7** compared to the unmodified transcript **8** is indicative of the incorporation of the higher molecular weight modified nucleoside (Fig. 4, compare lanes 1 and 2). When the transcription reaction was performed with this template in the presence of equimolar concentrations of UTP and **5**, T7 RNA polymerase exhibited a preference for the natural UTP (Fig. 4, lane 3). A control reaction in the absence of **5** did not result in the formation of any full-length product (Fig. 4, lane 4). Taken together, these observations clearly indicate that the enzyme accepts and incorporates the unnatural triphosphate **5** into the RNA strand with reasonable efficiency when not in competition with UTP.[20]

Larger quantities of oligoribonucleotide **7** were made by large-scale transcription reactions using template **6**. MALDI-TOF MS measurement of the PAGE purified transcript established the integrity of the modified full-length RNA (Figure S4†).[12] To further ascertain the presence of the intact thiophene-modified nucleobase in the RNA transcript, enzymatic digestions of transcript **7** with snake venom phosphodiesterase I, calf intestine alkaline phosphatase, RNase A, and RNase T1 were performed.[12] HPLC analysis of the resulting ribonucleoside mixture unambiguously revealed the presence of the modified nucleoside **4** and verified the expected stoichiometry (Fig. 5).[21]

While the RNA transcript **7** showed an emission at 412 nm, similar to the emission profile of the free nucleoside **4** in water, its intensity was ~5.5-fold lower than the free nucleoside (Figure S5†). Further quenching was observed upon hybridization to a perfect

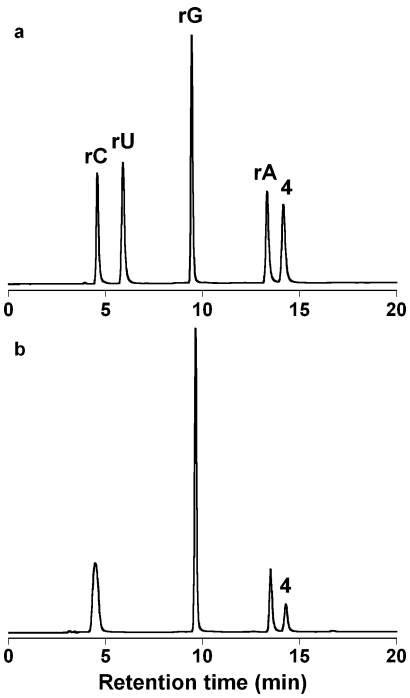


Fig. 5 HPLC profile of enzymatic digestion of transcript **7** at 260 nm. (a) Authentic nucleoside samples and modified nucleoside **4**. (b) Digested RNA transcript. See ESI for details.†

DNA complement **9** (Fig. 6). This quenching of fluorescence could be attributed to partial stacking of the chromophore and alteration of its microenvironment, or by neighboring bases *via* a photoinduced electron transfer mechanism, or both.[22] Similar reduction in emission has also been seen with other fluorescent nucleoside analogs (*e.g.*, 2-aminopurine, pyrroloC).[23] Remarkably, the emission of the duplex (**7•10**) containing the fluorescent nucleoside **4** opposite to C, was enhanced over 7-fold as compared to the perfect duplex **7•9** (Fig. 6), where **4** is found opposite to A. Mismatched duplexes **7•11** and **7•12**, placing T and G opposite

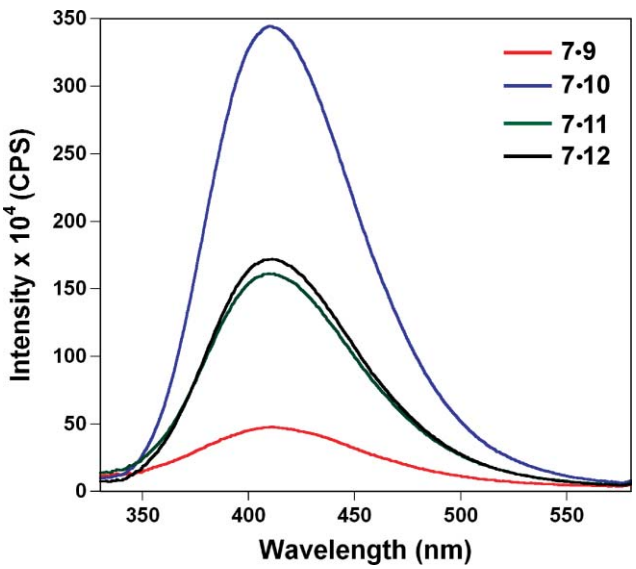


Fig. 6 Emission spectra of duplexes (1 μM) in 20 mM cacodylate buffer (500 mM NaCl, 0.5 mM EDTA, pH 7.0) at 25 °C. See ESI for details.†

to **4**, respectively, while exhibiting stronger emission compared to the perfect duplex **7•9**, were still significantly weaker than observed for a **4**•C mismatch in **7•10**. Nucleoside **4** therefore positively responds to the presence of a C nucleoside in a complementary oligodeoxynucleotide and can be classified as a base discriminating fluorescent nucleoside.

Several factors could be responsible for the enhanced emission observed for the mismatched duplex **7•10**. Possibly the simplest one would be severe destabilization caused by the putative **4**-C interaction leading to ineffective hybridization and, consequently, high abundance of the more emissive single stranded RNA **7**. To preemptively address such a possibility, we have conducted the hybridization and fluorescence experiments at slightly elevated ionic strengths (500 mM NaCl). Importantly, thermal denaturation studies with all **4**-containing duplexes and their corresponding control unmodified duplexes indicated that the incorporation of the thiophene modification had minimal effect on duplex stability (Table 2). While the perfectly-matched and least emissive duplex **7•9** suffered a ~6 °C of destabilization, all modified duplexes were as stable as or more stable than their unmodified counterparts (Table 2, Figures S6, S7†). Additional native gel retardation experiments with ^{32}P-labeled transcript RNA **7** with ssRNA as the control (Figure S8†) showed that all duplexes were completely intact under the condition utilized for the emission measurements.[12] Taken together, these results suggest that the modified nucleobase is likely to reside inside the helix and that the drastic emission enhancement observed for a mismatched duplex where **4** is found opposite C, when compared to a perfect duplex with **4** opposite A, is likely to result from differences in the microenvironment of the fluorescent nucleobase.

Numerous mechanisms, operating either in the ground or excited states, may account for the different impact of the opposite base upon the photophysical characteristics of an emissive nucleobase. Bulging out of a duplex, differential hydration of a non-canonical base pair, or base-pairing mediated tautomerization, constitutes a few ground state processes. Photo-induced electron or proton transfer represents plausible excited state events. Based on the Stern–Volmer titrations reported above it is clear that quenching by specific nucleobases is not directly correlated with the observed enhanced emission upon mispairing of **4** with C, although it is likely to impact the quantum yield of the emissive nucleobase. Additional experiments are required to further ascertain the fundamental causes of the observed photophysical behavior of **4** when incorporated into oligonucleotides.

Table 2 Thermal melting of duplexes derived from modified (**7**) and unmodified (**8**) RNA transcripts[a]

Tm/°C		Tm/°C	
Modified duplex		Unmodified duplex	
7•9	53.3 ± 0.7	**8•9**	59.4 ± 0.1
7•10	52.8 ± 0.1	**8•10**	48.3 ± 0.7
7•11	54.6 ± 0.5	**8•11**	50.3 ± 0.7
7•12	50.3 ± 0.7	**8•12**	49.8 ± 1.4

[a] Duplexes were formed by annealing a 1 : 1 mixture of the oligonucleotides in 20 mM cacodylate buffer (pH 7.0, 500 mM NaCl, 0.5 mM EDTA). Concentration of each duplex was 1 μM.[12]

In summary, the thieno[3,4-*d*]pyrimidine core represents a new motif for emissive nucleobase analogs, with emission in the visible range and respectable quantum yield. The corresponding ribonucleoside triphosphate can be enzymatically incorporated into RNA oligonucleotides using T7 RNA polymerase and the resulting fluorescent constructs can be employed as hybridization probes, positively reporting the presence of a C mismatch by enhanced emission. Nucleobase **4** therefore expands the repertoire of emissive isomorphic nucleoside analogs and opens up new opportunities to explore photophysical properties of modified nucleic acids.

Acknowledgements

We thank the National Institutes of Health (GM 069773) for support. We are grateful to Yun Xie and Mary Noe.

Notes and references

1 (*a*) The International HapMap Consortium, *Nature*, 2005, vol. 437, p. 1299; (*b*) H. Haga, Y. Yamada, Y. Ohnishi, Y. Nakamura and T. Tanaka, *J. Hum. Genet.*, 2002, **47**, 605; (*c*) A. J. Brookes, *Gene*, 1999, **234**, 177; (*d*) A. Chakravarti, *Nat. Genet. Suppl.*, 1999, **21**, 56.
2 (*a*) T. Pastinen and T. J. Hudson, *Science*, 2004, **306**, 647; (*b*) J. J. McCarthy and R. Hilfiker, *Nat. Biotechnol.*, 2000, **18**, 505.
3 (*a*) Reviews: Y. Suh and C. Cantor, *Mutat. Res.*, 2005, **573**, 1; (*b*) B. Sobrino, M. Brion and A. Carracedo, *Forensic Sci. Int.*, 2005, **154**, 181; (*c*) R. M. Twyman, *Curr. Top. Med. Chem.*, 2004, **4**, 1423; (*d*) K. Nakatani, *ChemBioChem*, 2004, **5**, 1623; (*e*) C. S. Carlson, T. L. Newman and D. A. Nickerson, *Curr. Opin. Chem. Biol.*, 2001, **5**, 78; (*f*) P. Y. Kwok, *Annu. Rev. Genomics Hum. Genet.*, 2001, **2**, 235.
4 (*a*) A. Friedrich, J. D. Hoheisel, N. Marmé and J.-P. Knemeyer, *FEBS Lett.*, 2007, **581**, 1644; (*b*) T. S. Kumar, J. Wengel and P. J. Hrdlicka, *ChemBioChem*, 2007, **8**, 1122; (*c*) E. Ergen, M. Weber, J. Jacob, A. Herrmann and K. Müllen, *Chem.–Eur. J.*, 2006, **12**, 3707; (*d*) U. Asseline, M. Chassignol, Y. Aubert and V. Roig, *Org. Biomol. Chem.*, 2006, **4**, 1949; (*e*) L. Valis, N. Amann and H.-A. Wagenknecht, *Org. Biomol. Chem.*, 2005, **3**, 36; (*f*) K. Yamana, Y. Fukunaga, Y. Ohtani, S. Sato, M. Nakamura, W. J. Kim, T. Akaike and A. Maruyama, *Chem. Commun.*, 2005, 2509; (*g*) A. O. Crockett and C. T. Wittwer, *Anal. Biochem.*, 2001, **290**, 89; (*h*) P. L. Paris, J. M. Langenhan and E. T. Kool, *Nucleic Acids Res.*, 1998, **26**, 3789.
5 A. Okamoto, Y. Saito and I. Saito, *Photochem. Photobiol. C: Photochem. Rev.*, 2005, **6**, 108.
6 (*a*) K. Tainaka, K. Tanaka, S. Ikeda, K.-I. Nishiza, T. Unzai, Y. Fujiwara, I. Saito and A. Okamoto, *J. Am. Chem. Soc.*, 2007, **129**, 4776; (*b*) Y. Saito, E. Mizuno, S. S. Bag, I. Suzuka and I. Saito, *Chem. Commun.*, 2007, 4492; (*c*) F. Takei, H. Suda, M. Hagihara, J. Zhang, A. Kobori and K. Nakatani, *Chem.–Eur. J.*, 2007, **13**, 4452; (*d*) R. H. E. Hudson and A. Ghorbani-Choghamarani, *Org. Biomol. Chem.*, 2007, **5**, 1845; (*e*) J. H. Ryu, Y. J. Seo, T. Hwang, J. Y. Lee and B. H. Kim, *Tetrahedron*, 2007, **63**, 3538; (*f*) Q. Xiao, R. T. Ranasinghe, A. M. P. Tang and T. Brown, *Tetrahedron*, 2007, **63**, 3483; (*g*) N. Moran, D. M. Bassani, J.-P. Desvergne, S. Keiper, P. A. S. Lowden, J. S. Vyle and J. H. R. Tucker, *Chem. Commun.*, 2006, 5003; (*h*) E. Mayer-Enthart and H.-A. Wegenknecht, *Angew. Chem., Int. Ed.*, 2006, **45**, 3372; (*i*) D. J. Hurley, S. E. Seaman, J. C. Mazura and Y. Tor, *Org. Lett.*, 2002, **4**, 2305.
7 Y. Tor, *Tetrahedron*, 2007, **63**, 3425.
8 (*a*) N. J. Greco and Y. Tor, *Tetrahedron*, 2007, **63**, 3515; (*b*) Y. Tor, S. Del Valle, D. Jaramillo, S. G. Srivatsan, A. Rios and H. Weizman, *Tetrahedron*, 2007, **63**, 3608.
9 N. J. Greco and Y. Tor, *J. Am. Chem. Soc.*, 2005, **127**, 10784.
10 (*a*) S. G. Srivatsan and Y. Tor, *J. Am. Chem. Soc.*, 2007, **129**, 2044; (*b*) S. G. Srivatsan and Y. Tor, *Tetrahedron*, 2007, **63**, 3601.
11 (*a*) S. Wang and E. T. Kool, *Biochemistry*, 1995, **34**, 4125; (*b*) N. Sugimoto, S. Nakano, M. Yoneyama and K. Honda, *Nucleic Acids Res.*, 1996, **24**, 4501; (*c*) F. Barone, L. Cellai, M. Matzeu, F. Mazzei and F. Pedone, *Biophys. Chem.*, 2000, **86**, 37; (*d*) H. Binder, *J. Phys.: Condens. Matter*, 2006, **18**, S491.

12 See ESI†.
13 F. Seela and X. Peng, *J. Org. Chem.*, 2006, **71**, 81.
14 Determined relative to 2-aminopurine.
15 C. Reichardt, *Chem. Rev.*, 1994, **94**, 2319.
16 R. W. Sinkeldam and Y. Tor, *Org. Biomol. Chem.*, 2007, **5**, 2523.
17 (*a*) E. L. Rachofsky, R. Osman and J. B. A. Ross, *Biochemistry*, 2001, **40**, 946; (*b*) D.-G. Xu and T. M. Nordlund, *Biophys. J.*, 2000, **78**, 1042.
18 S. G. Srivatsan and Y. Tor, *Nature Protocols*, 2007, **2**, 1547.
19 Y. Tor and P. B. Dervan, *J. Am. Chem. Soc.*, 1993, **115**, 4461.
20 Experiments with additional templates (see ESI, Figure S2 and S3†) illustrate that T7 RNA polymerase cannot introduce the modified nucleoside at positions that are closer to the promoter.
21 Mass spectral analysis of individual HPLC fractions further confirmed the authenticity of the native and modified nucleosides present in the enzymatic digest.
22 M. Torimura, S. Kurata, K. Yamada, T. Yokomaku, Y. Kamagata, T. Kanagawa and R. Kurane, *Anal. Sci.*, 2001, **17**, 155.
23 (*a*) J. M. Jean and K. B. Hall, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 37; (*b*) C. Liu and C. T. Martin, *J. Mol. Biol.*, 2001, **308**, 465.

