Site-Specific Enzymatic Incorporation of an Unnatural Base, N^6-(6-Aminohexyl)isoguanosine, into RNA

Yitzhak Tor and Peter B. Dervan*

Contribution from The Beckman Institute, California Institute of Technology, Pasadena, California 91125

Received December 11, 1992

Abstract: An efficient enzymatic method is described for the sequence-specific incorporation of a functionalizable modified base into RNA molecules. A deoxy-5-methylisocytidine (d^{Me}isoC) in the DNA template directs the T7 RNA polymerase incorporation of N^6-(6-aminohexyl)isoguanosine (6-AH-isoG) into the transcribed RNA product. The misincorporation of isoGTP derivatives opposite T is eliminated in the presence of ATP, and the misincorporation of A opposite d^{Me}isoC is negligible in the presence of isoGTP derivatives. The isolated yield of RNA products using modified templates is approximately 50% that for reactions using natural templates. A post-transcriptional modification of the reactive primary amino group with *N*-hydroxysuccinimide-activated biotin or the dianhydride of ethylenediaminetetraacetic acid affords site-specifically modified RNA sequences suitable for further studies. This method for the generation of RNA molecules containing a primary amine suitable for post-transcriptional modification should be useful for mapping the structure of folded RNA polymers and RNA–protein complexes by affinity cleavage and affinity labeling.

Introduction

RNA plays an important role in cellular processes, including regulation and catalysis. The recent discovery of self-splicing RNA molecules,[1] the involvement of ribosomal RNA in protein synthesis,[2] and the artificial evolution of RNA enzymes[3] have created an increased demand for structural information. Yet, little is known about the three-dimensional structure of RNA biopolymers.[4] Few crystal structures of RNAs have been determined,[5] and only recently has the application of modern high-resolution NMR techniques for the study of small RNA molecules been reported.[6] In the absence of such high-resolution studies, solution methods are needed to study the higher order structure of RNA molecules.[7] Attachment of EDTA·Fe to peptides, proteins, and oligonucleotides is a useful technique for studying the structure of peptide–DNA,[8] protein–DNA,[9] and triple-helical complexes.[10] The EDTA·Fe moiety attached to a discrete location in the ligand–nucleic acid complex generates a diffusible oxidant upon addition of reducing agents such as dithiothreitol (DTT).[11] Cleavage at several nucleotide positions proximal in space to the location of the bound EDTA·Fe affords a set of DNA fragments whose lengths reveal its location, and hence structural information on that position relative to neighboring nucleotide positions. Extending this affinity cleaving approach to the mapping of RNA structures requires a method for the incorporation of EDTA·Fe at discrete nucleotide positions in RNA molecules.

DNA-dependent RNA polymerase of bacteriophage T7 is widely used for the *in vitro* synthesis of relatively large quantities of short[12] as well as long[13] RNAs. Transcription reactions with T7 RNA polymerase in the presence of modified natural bases (*e.g.*, biotin-linked ATP) result in the incorporation of the modified bases into the RNA chain,[14] indicating that this enzyme can tolerate such modifications and may incorporate these modified natural bases into RNA products. In order to site-specifically

(1) Cech, T. *Annu. Rev. Biochem.* **1990**, *59*, 543–568.
(2) Noller, H. F.; Hoffarth, V.; Zimniak, L. *Science* **1992**, *256*, 1416–1419.
(3) Beaudry, A. A.; Joyce, G. F. *Science* **1992**, *257*, 635–641.
(4) (a) Tinoco, I., Jr.; Puglisi, J. D.; Wyatt, J. R. In *Nucleic Acids and Molecular Biology*; Eckstein, F., Lilley, D. M. J., Eds.; Springer-Verlag: Berlin, 1990; Vol. 4, pp 205–226. (b) Draper, D. E. *Acc. Chem. Res.* **1992**, *25*, 201–207.
(5) tRNA crystal structures: (a) Kim, S.-H.; Sussman, J. L.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Wang, A. H. J.; Seeman, N. C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4970. (b) Quigley, G. J.; Rich, A. *Science* **1976**, *194*, 796–806. (c) Westhof, E.; Dumas, P.; Moras, D. *J. Mol. Biol.* **1985**, *184*, 119–145. (d) Brown, R. S.; Dewan, J. C.; Klug, A. *Biochemistry* **1985**, *24*, 4785–4801. RNA dodecamer crystal structure: Holbrook, S. R.; Cheong, C.; Tinoco, I., Jr.; Kim, S.-H. *Nature* **1991**, *353*, 579–581.
(6) (a) Varani, G.; Tinoco, I., Jr. *Q. Rev. Biophys.* **1991**, 24, 479–532 and references cited therein. (b) Nikonowicz, E. P.; Pardi, A. *Nature* **1992**, *355*, 184–186.
(7) For chemical reagents and enzymatic assays for analyzing the structure of RNA in solution, see: (a) Wang, J.-F.; Cech, T. R. *Science* **1992**, *256*, 526–529. (b) Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839–2841. (c) Chow, C. S.; Barton, J. K. *Biochemistry* **1992**, *31*, 5423–5429. (d) Vary, C. P. H.; Vournakis, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6978–6982. (e) Barciszewska, M. Z.; Barciszewski, J.; Erdmann, V. A. *Int. J. Biol. Macromol.* **1992**, *14*, 41–44. (f) Ehresmann, C.; Baudin, F.; Mougel, M.; Romby, P.; Ebel, J.-P.; Ehresmann, B. *Nucleic Acids Res.* **1987**, *15*, 9109–9128.
(8) (a) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861–6863. (b) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457–465. (c) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528. (d) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1986**, *108*, 5008–5009. (e) Wade, W. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 1574–1575. (f) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 6840–6842. (g) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 7564–7566. (h) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783–8794.
(9) (a) Sluka, J. P.; Horvath, S. J.; Bruist, M. F.; Simon, M. I.; Dervan, P. B. *Science* **1987**, *238*, 1129–1132. (b) Sluka, J. P.; Griffin, J. H.; Mack, D. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 6369–6374. (c) Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847–850. (d) Sluka, J. P.; Horvath, S. J.; Glasgow, A. C.; Simon, M. I.; Dervan, P. B. *Biochemistry* **1990**, *29*, 6551–6561. (e) Mack, D. P.; Sluka, J. P.; Shin, J. A.; Griffin, J. H.; Simon, M. I.; Dervan, P. B. *Biochemistry* **1990**, *29*, 6561–6567. (f) Graham, K. S.; Dervan, P. B. *J. Biol. Chem.* **1990**, *265*, 16534–16540. (g) Shin, J. A.; Ebright, R. H.; Dervan, P. B. *Nucleic Acids Res.* **1991**, *19*, 5233–5236.
(10) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645–650. (b) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927–7929. (c) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059–3061. (d) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967–971. (e) Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 7286–7287. (f) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435–2437. (g) Horne, D. A.; Dervan, P. B. *Nucleic Acids Res.* **1991**, *19*, 4963–4965. (h) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360–1363. (i) Kiessling, L. L.; Griffin, L. C.; Dervan, P. B. *Biochemistry* **1992**, *31*, 2829–2834. (j) Beal, P. A.; Dervan, P. B. *Nucleic Acids Res.* **1992**, *20*, 2773–2776. (k) Beal, P. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 4976–4982. (l) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6957–6965.
(11) Dervan, P. B. *Science* **1986**, *232*, 464–471.
(12) (a) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783–8798. (b) Milligan, J. F.; Uhlenbeck, O. C. *Methods Enzymol.* **1989**, *180*, 51–62.
(13) (a) Yisraeli, J. K.; Melton, D. A. *Methods Enzymol.* **1989**, *180*, 42–50. (b) Reyes, V. M.; Abelson, J. N. *Methods Enzymol.* **1989**, *180*, 63–69.

0002-7863/93/1515-4461$04.00/0 © 1993 American Chemical Society

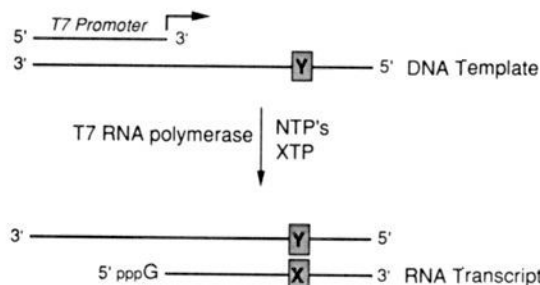

Figure 1. Site-specific enzymatic incorporation of a novel base X into an RNA transcript using a DNA template containing its novel base complement Y, T7 RNA polymerase, XTP, and the natural triphosphates (NTPs). Note that G 5'-triphosphate, which corresponds to the last nucleotide of the 18-mer promoter, is incorporated at the 5'-end of the RNA transcript.[12]

introduce modified bases into the RNA, a unique base-pairing scheme is needed. Such a novel hydrogen-bonding scheme should be distinguishable from Watson–Crick A-T(U) and G-C base pairing in order to retain high fidelity in the transcription reaction.[15] Substituting an unnatural base at a particular position in the DNA template should lead to the site-specific introduction of its complement into the transcribed RNA (Figure 1). Modifying the novel base triphosphate with functionalizable groups (*e.g.*, primary alkylamines) should allow the incorporation of reactive functionalities into the enzymatically synthesized RNAs suitable for post-transcriptional modification.[16]

Three decades ago Rich proposed that an isoC-isoG base-pairing system was potentially suitable for specific incorporation into nucleic acids because it forms a hydrogen-bonding pattern distinct from the natural A-T(U) or G-C base pairs (Figure 2).[17,18] Recent theoretical calculations suggest that the isoC-isoG base pair is comparable in stabiliy to the C-G base-pair.[19] Benner and co-workers demonstrated that isoG can be site-specifically incorporated into RNA using disoC containing DNA templates and T7 RNA polymerase, although with apparently lower efficiency when compared to native DNA templates.[20] The level of misincorporation and fidelity are polymerase dependent. Their studies with DNA polymerase (Klenow fragment) indicated that isoG is misincorporated opposite T in the DNA template. Additional misincorporation of A (*ca.* 15%) into the primer-extended DNA polymerase products was postulated to arise from the hydrolysis of disoC to produce a dU residue.[20] As part of an exploratory effort to incorporate modified bases into RNA

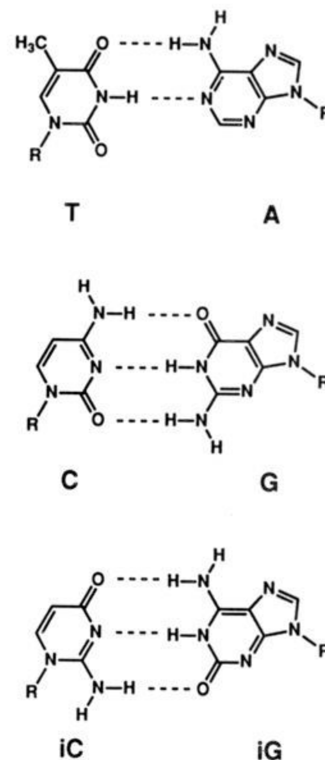

Figure 2. T-A and C-G Watson–Crick base pairing and the putative Watson–Crick base pairing of isoC-isoG (R = β-ribofuranosyl).

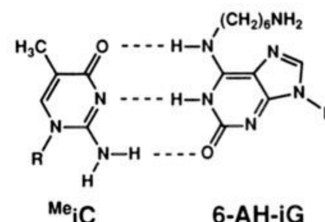

Figure 3. Putative Watson–Crick base pairing of 5-methylisoC and N^6-(6-aminohexyl)isoG (R = β-ribofuranosyl).

polymers, we have investigated the potential of this base-pairing scheme for the site-specific enzymatic incorporation of a functionalizable isoG derivative into RNA.

We replaced deoxyisoC (disoC) by deoxy-5-methylisoC (d^{Me}-isoC) and examined the fidelity of the d^{Me}isoC-isoG system in T7 RNA polymerase transcription reactions (Figure 3). We report that (i) d^{Me}isoC in a DNA template will efficiently direct the T7 RNA polymerase incorporation of isoGTP, (ii) the misincorporation of isoG opposite T is eliminated *in the presence of ATP*, and (iii) a novel synthetic N^6-(6-aminohexyl)isoguanosine triphosphate with a primary amine suitable for post-transcriptional modification can be incorporated into the RNA transcript by T7 RNA polymerase, providing a site-specifically functionalized RNA (Figure 3). The reactive primary amino group in the transcribed RNA can be treated with activated carboxylates (*e.g.*, *N*-hydroxysuccinimide-activated biotin and ethylenediaminetetraacetic acid dianhydride) to provide chemically modified RNAs suitable for structural studies by affinity labeling and affinity cleaving methods.

Results and Discussion

d^{Me}isoC-isoGTP System. A DMT-protected phosphoramidite of 5-methylisocytidine, **5**, suitable for incorporation into oligonucleotides by automated methods was synthesized in five steps from 2,5'-anhydrothymidine (**1**, Figure 4). Treatment of anhydrothymidine (**1**)[21] with saturated methanolic ammonia afforded the free nucleoside 2'-deoxy-5-methylisocytidine (**2**) in

(14) (a) Folsom, V.; Hunkler, M. J.; Haces, A.; Harding, J. D. *Anal. Biochem.* **1989**, *182*, 309–314. (b) Luehresen, K. R.; Baum, M. P. *BioTechniques* **1987**, *5*, 660–670.

(15) Ojha, R. P.; Roychoudhury, M.; Sanyal, N. K. THEOCHEM **1991**, *233*, 247–273.

(16) The sequence-specific modification of RNA has been predominantly accomplished by the chemical synthesis of short RNA fragments. For recent examples of backbone or base-modified short synthetic RNAs see: (a) Slim, G.; Gait, M. J. *Nucleic Acids Res.* **1991**, *19*, 1183–1188. (b) Olson, D. B.; Benseler, F.; Aurup, H.; Dieken, W. A.; Eckstein, F. *Biochemistry* **1991**, *30*, 9735–9741. (c) Perreault, J.-P.; Labuda, D.; Usman, N.; Yang, J.-H.; Cedergren, R. *Biochemistry* **1991**, *30*, 4020–4025. (d) Williams, D. M.; Pieken, W. A.; Eckstein, F. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 918–921. (e) Scaringe, S. A.; Francklyn, C.; Usman, N. *Nucleic Acids Res.* **1990**, *18*, 5433–5441. (f) Green, R.; Szostak, J. W.; Benner, S. A.; Rich, A.; Usman, N. *Nucleic Acids Res.* **1991**, *19*, 4161–4166. For other approaches see: (g) Draper, D. E. *Nucleic Acids Res.* **1984**, *12*, 989–1002. (h) Pitulle, C.; Kleineidam, R. G.; Sproat, B.; Krupp, G. *Gene* **1992**, *112*, 101–105. (i) Bain, J. D.; Switzer, C. *Nucleic Acids Res.* **1992**, *20*, 4372. (j) Moore, J. M.; Sharp, P. A. *Science* **1992**, *256*, 992–997.

(17) Rich, A. In *Horizons in Biochemistry*; Kasha, M., Pullmann, B., Eds.; Academic Press: New York, 1962; pp 103–126.

(18) *Iso*G has been shown to be predominantly in the N1-H tautomeric form shown in Figure 2. See: Sepiol, J.; Kazimierczuk, Z.; Shugar, D. *Z. Naturforsch.* **1976**, *31C*, 361–370.

(19) Leach, A. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1992**, *114*, 3675–3683.

(20) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322–8323. See also: Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537–539. For the enzymatic incorporation of other novel base pairs, see: Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33–37.

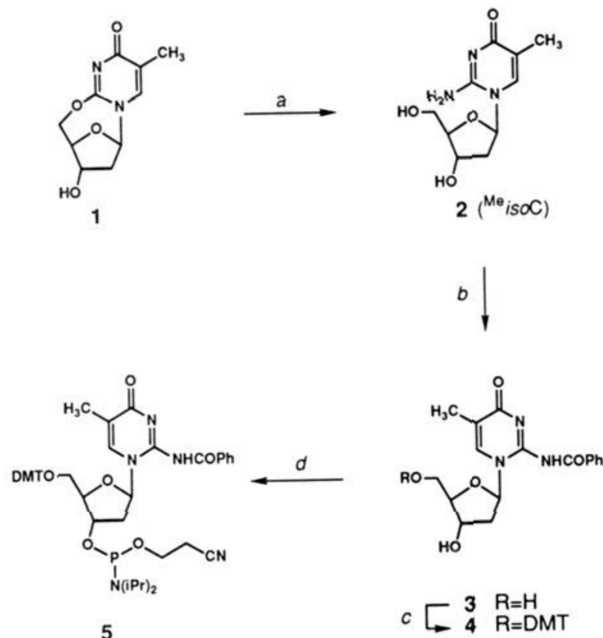

Figure 4. Synthesis of deoxy-5-methylisocytidine phosphoramidite **5** used in the automated DNA synthesis of the modified DNA templates: (a) NH_3/MeOH, 72%; (b) (i) PhCOCl, pyridine; (ii) NaOH/MeOH/H_2O/pyridine, 63%; (c) DMT-Cl, pyridine, Et_3N, DMAP, 61%; (d) 2-cyanoethyl *N,N*-diisopropylphosphoramidochloridite, iPr_2NEt, CH_2Cl_2, 73%.

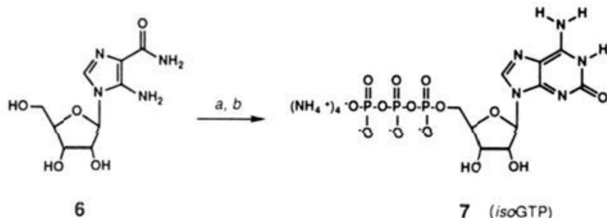

Figure 5. Two-step synthesis of isoGTP (**7**) from commercially available 5-amino-4-(aminocarbonyl)imidazolyl β-D-ribofuranoside (AICA-riboside, **6**): (a) (i) PhCONCS, DMF; (ii) DCC; (iii) MeOH, NH_4OH, 55% (b) (i) $POCl_3$, $(MeO)_3PO$; (ii) tributylammonium pyrophosphate, 57%.

72% yield. Benzoylation of the aromatic amino group was achieved in 63% yield by the peracylation procedure[22] followed by selective hydrolysis of the benzoate esters. Treatment of the resulting nucleoside **3** with 4,4′-dimethoxytrityl chloride furnished the 5′-protected nucleoside **4** in 61% yield. Activation using 2-cyanoethyl *N,N*-diisopropylphosphoramidochloridite afforded phoshoramidite **5** in 73% yield, which was then incorporated into oligodeoxyribonucleotides using standard automated solid-phase 2-cyanoethyl phosphoramidite chemistry.[22] Isoguanosine triphosphate (**7**, isoGTP) was prepared from 5-amino-4-(aminocarbonyl)imidazolyl β-D-ribofuranoside (**6**) according to known synthetic procedures (Figure 5).

Three DNA templates, **8–10**, 27 nucleotides in length which varied at one position (Y = dC, d^{Me}isoC, or dT) were synthesized. In addition, a shorter oligodeoxyribonucleotide (18-mer) complementary to the 3′-end of **8–10** was synthesized to create a T7 RNA polymerase promoter upon annealing (Figure 6). Template **8** (Y = dC) served as a control for the efficiency of the transcription reactions and for the product distribution. Template **9** (Y = d^{Me}isoC) contains the novel base d^{Me}isoC and was used to study the incorporation of isoG derivatives and the production of uniquely modified RNAs.[23] Template **10** (Y = dT) was used to

(21) Watanabe, K. A.; Reichman, U.; Chu, C. K.; Fox, J. J. In *Nucleic Acid Chemistry*; Tipson, R. S., Townsend, C. B., Eds.; John Wiley and Sons: New York, 1978; Part 1, pp 273–277.

(22) *Oligonucleotide Synthesis*; Gait, M. J., Ed.; IRL Press: Oxford, U.K., 1984.

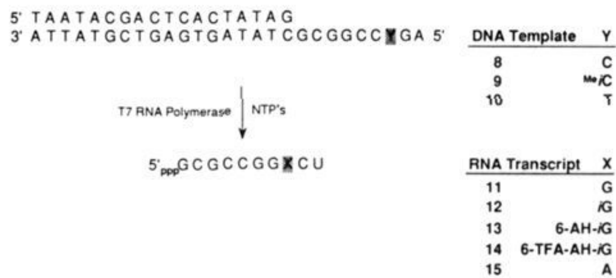

Figure 6. DNA templates **8–10** used in this work (annealed to the consensus 18-mer promoter) and the expected T7 RNA polymerase transcription products **11–15**.

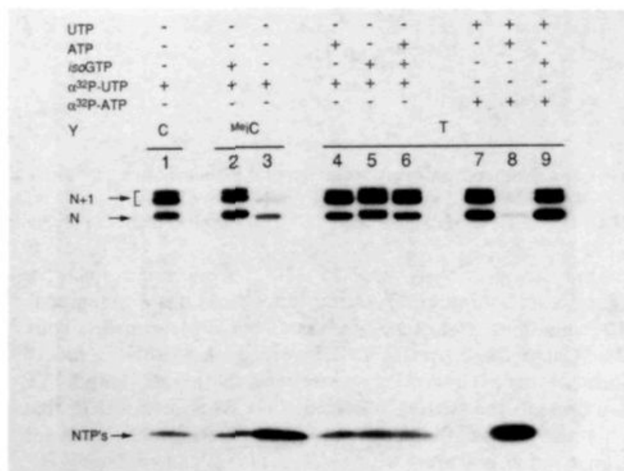

Figure 7. Autoradiogram of a 20% polyacrylamide gel used to separate the runoff transcription reaction products obtained with DNA templates **8–10**. The reactions were carried out as described in the Experimental Section with the triphosphates indicated. All reaction mixtures contained CTP and GTP, which are required for the production of full-length (N = 10) transcripts. Note that most of the radiolabeled UTP (lanes 1, 2, 4–6) and all the radiolabeled ATP (lanes 7 and 9) is consumed. Under these electrophoresis conditions, sequences differing by a single base are well resolved. The formation of a full-length product (N) together with a heterogeneous mixture of $N + 1$ products is a common feature of transcription reactions with T7 RNA polymerase.[12]

examine the misincorporation of isoG opposite dT. The last base in the DNA templates is a unique dA residue which allows for the detection of full-length RNA products by using labeled α-^{32}P-UTP in the transcription reaction.[24]

The synthetic DNA templates (**8–10**) were annealed to the 18-mer T7 consensus oligodeoxyribonucleotide promoter and were incubated with T7 RNA polymerase under standard *in vitro* transcription reaction conditions[12] in the presence of the various triphosphates. The transcription products obtained were separated by high-resolution gel electrophoresis (Figure 7). The full-length RNA transcription product 10-mer (lower band – N) and a heterogeneous mixture of $N + 1$ products, resulting from the random incorporation of an uncoded extra base,[12] were obtained with template **8** (Figure 7, lane 1).[25] A comparable yield and similar pattern of transcription products is obtained with the d^{Me}isoC template **9** in the presence of isoGTP (Figure 7, lane 2). When isoGTP is eliminated from the reaction mixture, very little full-length product is formed (Figure 7, lane 3).[26] In order to

(23) Enzymatic digestion of the oligonucleotides confirmed the presence of the d^{Me}isoC in the expected amount (±10%).

(24) Note that the use of α-^{32}P-UTP allows the observation of full-length (and longer) products and does not reveal any abortive sequences.

(25) Note that under these conditions, products differing by a single base are well resolved. See: Frank, R.; Koster, H. *Nucleic Acids Res.* **1979**, *6*, 2069–2987.

(26) The weak band which presumably corresponds to a full-length transcription product may result from misincorporation of one of the natural bases due to the limited fidelity of the RNA polymerase itself. See Coleman, J. E.; Martin, C. T.; Muller, D. K. *Biochemistry* **1988**, *27*, 3966–3974.

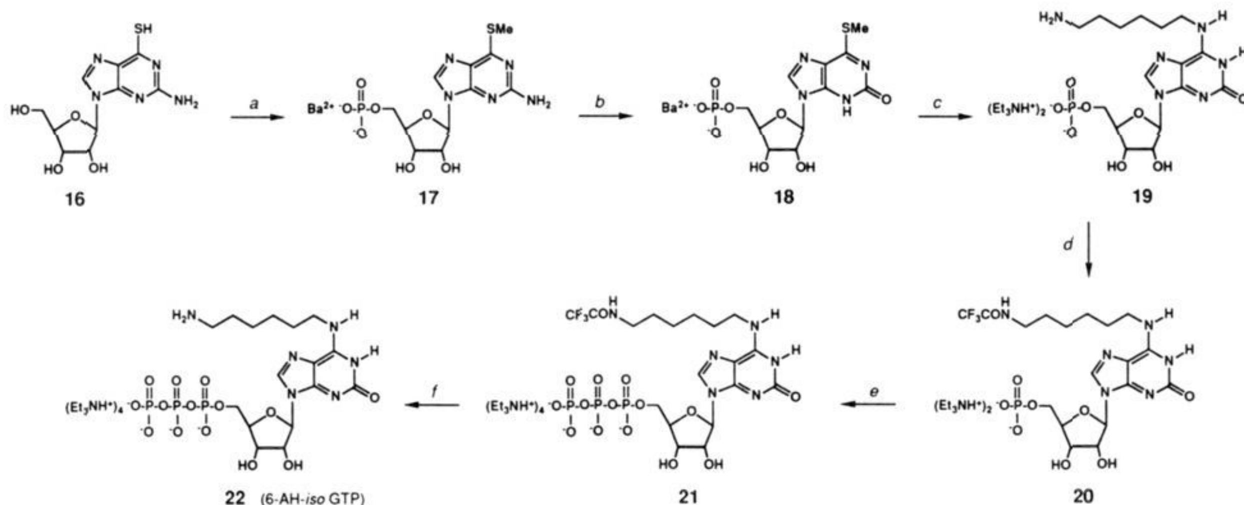

Figure 8. Synthetic route to N^6-(6-aminohexyl)isoguanosine triphosphate (**22**): (a) (i) $POCl_3$, $(MeO)_3PO$; (ii) LiOH, H_2O, $(MeO)_3PO$; (iii) $Ba(OAc)_2$, 94%; (b) (i) $NaNO_2$, $AcOH/H_2O$; (ii) $Ba(OAc)_2$, 85%; (c) (i) Dowex, 50X8-200; (ii) $H_2N(CH_2)_6NH_2$, MeOH, 65 °C, 75%; (d) CF_3COSEt, LiOH/H_2O; (e) (i) 1,1'-carbonyldiimidazole; (ii) tributylammonium pyrophosphate, DMF, 68%; (f) NaOH, pH 11–11.5, 75%.

examine the previously reported DNA polymerase misincorporation of isoG opposite dT, transcription reactions with template **10** (containing dT instead of d^{Me}isoC) were performed (Figure 7). Comparing lanes 4 and 5 reveals that isoG is indeed incorporated opposite dT; however, when both isoGTP and ATP are present, the pattern obtained (lane 6) is identical to that obtained with ATP alone (lane 4). Furthermore, running the transcription reactions with α-^{32}P-ATP instead of α-^{32}P-UTP (Figure 7, lanes 7–9) shows that while 1 mM unlabeled ATP completely displaces the radiolabeled adenosine (lane 8), 1 mM isoGTP does not compete with the labeled adenosine (lane 9), despite a *ca.* 10^4-fold excess in the reaction mixture. We conclude that while isoG is capable of incorporation opposite dT (either as the enol tautomer or by the formation of a wobble base pair), *it does not compete with A for available dT sites in the presence of both triphosphates.* Therefore the d^{Me}isoC-isoG system seems suitable for the purpose of site-specific incorporation of modified bases into RNA.

Synthesis of N^6-(6-Aminohexyl)isoGTP. 6-Aminohexyl-substituted isoGTP **22** (6-AH-isoGTP) was synthesized in six steps from 6-mercaptoguanosine (**16**, Figure 8). 6-Mercaptoguanosine (**16**) was phosphorylated at the 5'-hydroxyl group using phosphoryl chloride and subsequently S-methylated with trimethylphosphate according to known procedures.[27] Deamination of 6-methylthioguanosine 5'-monophosphate (**17**) was achieved in 85% yield by treatment with sodium nitrite in dilute acetic acid, affording 6-methylthioxanthosine 5'-monophosphate (**18**) as the Ba^{2+} salt. Treatment with a large excess of 1,6-diaminohexane afforded the N^6-(6-aminohexyl)isoguanosine 5'-monophosphate (**19**) in 75% yield. The primary amine on nucleotide **19** was protected as the trifluoroacetamide, and the 5'-monophosphate **20** was activated as the imidazolate and condensed with tributylammonium pyrophophate[28] to yield 6-TFA-AH-isoGTP (**21**; 68% yield from **20**). The TFA group was removed by mild basic hydrolysis to afford the desired triphosphate 6-AH-isoGTP (**22**) in 75% yield.[29]

T7 RNA Polymerase Incorporation of 6-AH-isoGTP. IsoGTP (**7**), 6-AH-isoGTP (**22**), and its precursor 6-TFA-AH-isoGTP (**21**) are substrates for T7 RNA polymerase, as is evident from the formation of full-length products in all cases (Figure 9, lanes 1, 4, and 6, respectively). Subtle changes in the pattern of the

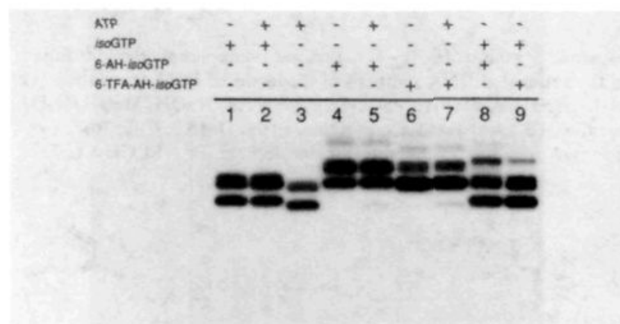

Figure 9. Autoradiogram of a 20% polyacrylamide gel used to resolve the transcription reaction products obtained with the d^{Me}isoC template (**9**) and isoGTP (**7**), 6-AH-isoGTP (**22**), and 6-TFA-AH-isoGTP (**21**). All reaction mixtures contained 1 mM CTP, 1 mM GTP, and α-^{32}P-UTP, in addition to 1 mM of the triphosphates indicated.

$N + 1$ bands may indicate slightly different interactions of the polymerase with the modified RNA chains. All three triphosphates compete favorably against misincorporation of ATP. While some ATP gets incorporated against the d^{Me}isoC in template **9** in the absence of any isoG triphosphate (Figure 9, lane 3), their presence suppresses this misincorporation (Figure 9, lanes 2, 5, and 7). The negligible incorporation of A into the RNA product in the presence of isoGTP derivatives suggests that d^{Me}isoC is less susceptible to hydrolysis as apparently occurs with disoC.[20] Short RNA transcripts **13** and **14** containing one N^6-substituted isoG have very different gel mobility than RNA transcript **12** containing a single isoG (compare lanes 4 and 6 to lane 1 in Figure 9). A competition experiment between the parent isoGTP and the N^6-substituted isoGTPs reveals that the unsubstituted isoG is incorporated somewhat more efficiently than the substituted derivatives (Figure 9, lanes 8 and 9). This result is in agreement with previous studies showing lower incorporation efficiencies of modified natural bases compared to their unsubstituted counterparts.[14] It is interesting to note that in both cases (Figure 9, lanes 8 and 9) the incorporation of the N^6-substituted isoG as the extra base in an $N + 1$ product is evident from the slowly migrating extra band. It indicates that T7 RNA polymerase accepts and tolerates these modified derivatives as part of the triphosphate pool.

Large-Scale RNA Synthesis. The *in vitro* transcription reactions were scaled up 12.5-fold using unlabeled triphosphates, while keeping the reaction conditions essentially the same.[30] Gel electrophoresis purification of the products afforded approximately 1.5–2.5 nmol of isolated isoG or 6-AH-isoG modified

(27) Perini, F.; Hampton, A. *J. Heterocycl. Chem.* **1970**, 969–971.
(28) Hoard, D. E.; Ott, D. G. *J. Am. Chem. Soc.* **1965**, *87*, 1785–1788.
(29) Comparing the UV spectra of the N^6-alkyl-substituted isoG derivatives in water to that of the parent isoG shows very similar patterns (λ_{max} and relative intensities of bands), which qualitatively indicates that these modified derivatives are also predominantly in the indicated N1-H form. See ref 18 for a discussion concerning the tautomeric forms of isoG dervatives.

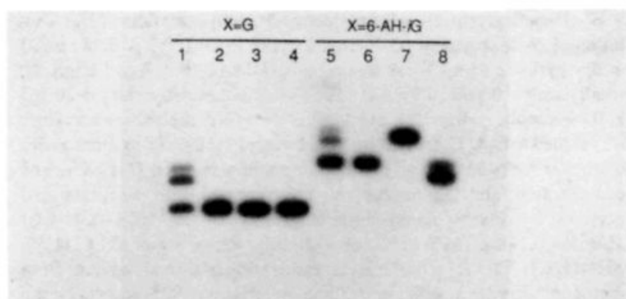

Figure 10. Autoradiogram of a 20% polyacrylamide gel used to separate the crude small-scale transcription reaction mixtures performed with α-^{32}P-UTP and DNA templates **8** (lane 1) and **9** (lane 5); the 5′-labeled purified large-scale transcription products **11** and **13** obtained with the same DNA templates (lanes 2 and 6, respectively) and the post-transcriptional chemical modifications of these labeled purified transcripts with *N*-hydroxysuccinimidobiotin (lanes 3 and 7 for **11** and **13**, respectively) or EDTA dianhydride (lanes 4 and 8 for **11** and **13**, respectively). The purified control RNA (RNA transcript **11**) and the 6-AH-isoG containing RNA (RNA transcript **13**) had been treated with alkaline phosphatase in order to remove their 5′-triphosphate groups and then 5′-labeled using T4 polynucleotide kinase and γ-^{32}P-ATP. The α-^{32}P-U-labeled crude transcription reaction mixture had been similarly treated (alkaline phosphatase followed by kinase 5′-phosphorylation using unlabeled ATP).

RNAs (compared to *ca.* 5 nmol of unmodified RNA obtained under similar conditions). Despite the fact that abortive sequences have also been observed on the purification gel, the amount obtained indicates up to a 40-fold amplification over the amount of the DNA template. The purified large-scale RNA products correspond to the full-length (*N*) product in the crude transcription reaction mixtures (compare lane 1 to 2, and lane 5 to 6 in Figure 10). Enzymatic digestion of the RNA products followed by an HPLC analysis confirmed the presence of 6-AH-isoG in the transcribed product **13** (Figure 11). It is important to note that adenosine is not present among the nucleosides in trace c (in agreement with the results described above). It demonstrates again that d^{Me}isoC is likely to be more stable than disoC toward hydrolysis[31] and that T7 RNA polymerase shows satisfying fidelity with the d^{Me}isoC-isoGTP system. One possible explanation for the improved transcriptional behavior of d^{Me}isoC vs disoC is that the presence of 5-methyl groups allows better contact between the template and the polymerase.[32] Note that the reverse base-pairing system, disoG-^{Me}isoCTP, has not yet been investigated and might present different fidelity during *in vitro* transcription reactions.

Post-Transcriptional Modification. Introducing a functionalized, nonstandard base into RNA can be accomplished by two routes: by enzymatic incorporation of the fully modified unnatural base triphosphate or by a post-transcriptional modification of the isolated RNA containing the alkylamine modified base. While the results reported above demonstrate that both approaches are potentially feasible using the d^{Me}isoC-isoG system (as the enzyme tolerates both the TFA-protected triphosphate **21** and the unprotected amino-linker triphosphate **22**), the latter approach appears more attractive, primarily due to the ability to modify a common "reactive" RNA moiety with a variety of activated chemical reagents. Furthermore, the post-transcriptional modification approach avoids possible repulsive steric interactions between the polymerase and large groups on the triphosphate as well as the need to optimize transcription reaction conditions for each modified triphosphate.[14]

(30) Note that all the triphosphates (including ATP) were present in the large-scale transcription reactions of the d^{Me}isoC-containing DNA template (see Experimental Section).

(31) Since we have demonstrated that isoGTP derivatives do not compete with ATP over T sites (Figure 7, lanes 4–9), Figure 9 clearly shows that when both triphosphates are present only isoGTP derivatives get incorporated (lanes 2, 5, and 7) and that the level of A incorporation is very low (<1%). Furthermore, HPLC analysis shows the absence of A in the isolated large-scale RNA transcript **13** (Figure 11, trace c). Taken together, we conclude that the presence of the dT residue at the unique position must be very low and, therefore, that d^{Me}isoC does not undergo significant hydrolysis to dT under our experimental conditions.

(32) Jaworski, M. D.; Martin, C. T. *Biophys. J.* **1990**, *57*, 64a.

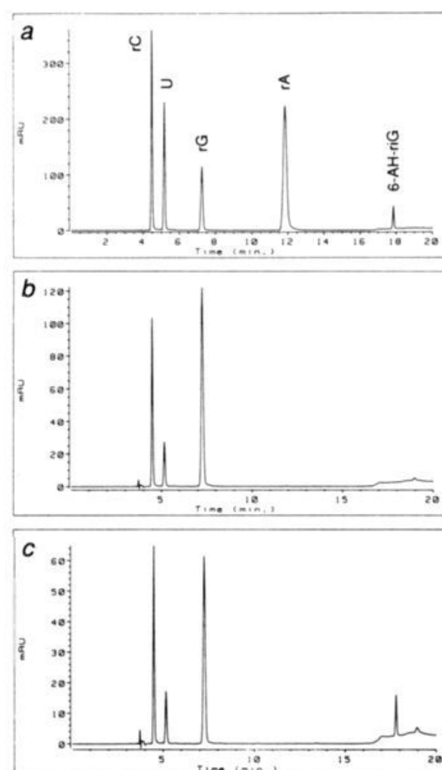

Figure 11. HPLC analysis of the enzymatic digestion of the control RNA transcript **11** (trace b) and the 6-AH-isoG-containing RNA transcript **13** (trace c) compared to the chromatogram of authentic nucleosides (trace a). Enzymatic digestion has been carried out using phosphodiesterase, alkaline phosphatase, and RNase A, T_1, and T_2 (see Experimental Section) and analyzed by reversed-phase HPLC using gradient B. The chromatograms shown were monitored at 260 nm.

The isolated functionalizable RNA transcript **13** is easily modified using simple activated carboxylates (Figure 10). The 6-AH-isoG containing RNA (**13**) is smoothly biotinylated (lane 7) and reacts almost quantitatively with EDTA dianhydride (lane 8), while the control RNA transcript **11**, containing only natural bases, shows no reaction at all (Figure 10, lanes 3 and 4, respectively). These results clearly demonstrate that a post-transcription chemical modification of the aminohexyl-modified RNA is readily achieved.

Implications. An *in vitro* enzymatic method for the sequence-specific incorporation of a nonstandard base bearing an alkylamine side chain into RNA has been presented. T7 RNA polymerase accepts N^6-substituted isoGTPs as substrates and incorporates them into unique sites within RNA transcripts using DNA templates containing deoxy-5-methylisocytidine. The reactive amino groups on the RNA transcripts obtained using N^6-(6-aminohexyl)isoGTP can be post-transcriptionally modified using activated carboxylate derivatives, affording site-specifically modified RNAs. The relatively straightforward syntheses of the N^6-substituted isoGTP derivatives presented here should allow for the preparation of other modified triphosphates bearing different functional groups and variable linker sizes. Enzymatic incorporation of such derivatives should provide a variety of site-specifically modified RNA molecules suitable for structure–activity studies. The introduction of reactive moieties such as EDTA·Fe and 2-bromoacetamide at single nucleotide positions within large biologically relevant RNA polymers should facilitate the mapping of their three-dimensional shapes by affinity cleaving and cross-linking methods.

Experimental Section

NMR spectra were recorded on a GE 300 instrument operating at 300 MHz (^{1}H) and 75 MHz (^{13}C). Chemical shifts are reported in ppm relative to the solvent residual signal. UV spectra were measured on a Hewlett-Packard Model 8452A diode array spectrophotometer. High-resolution FAB mass spectra were recorded at the Mass Spectrometry Laboratory at the University of California, Riverside. HPLC analysis was performed on an HP 1090 Series II analytical HPLC using a Vydac C18 reversed-phase column (0.46 × 25 cm, 5 μm TP silica) with 0.25 M triethylammonium acetate pH 6.0 and acetonitrile as eluents at a flow rate of 0.8 mL/min. Two different gradients were used: gradient A 0–20% acetonitrile in 20 min; gradient B 0–10 min, 0% acetonitrile, 10–25 min, 0–20% acetonitrile, 25–30 min, 20% acetonitrile. Chromatograms were monitored at 260, 294, and 320 nm. Flash column chromatography[33] was carried out using silica gel 60 (230–400 mesh, Merck). Thin-layer chromatography (TLC) was performed on silica gel 60 F_{254} precoated plates (Merck). Reagent-grade chemicals were used as received unless otherwise stated. Acetonitrile, dichloromethane, and diisopropylethylamine were distilled from calcium hydride. Anhydrous *N,N*-dimethylformamide and pyridine, stored over 4-Å molecular sieves, were purchased from Fluka. *N*-hydroxysuccinimidobiotin and tributylammonium pyrophosphate were obtained from Sigma. EDTA dianhydride was purchased from Aldrich. 2,5′-Anhydrothymidine was synthesized according to Watanabe.[21] Isoguanosine was synthesized from AICA-riboside (Sigma) according to Chern et al.[34] and phosphorylated using the reported $POCl_3/(MeO)_3PO$ procedure followed by *in situ* treatment with tributylammonium pyrophosphate.[35] 6-Methylthioguanosine monophosphate barium salt was prepared from the commercially available 6-mercaptoguanosine (Sigma) according to Perini and Hampton.[27]

2′-Deoxy-5-methylisocytidine (2). A suspension of 2,5′-anhydrothymidine (**1**)[21] (0.255 g, 1.1 mmol) in methanol (15 mL) at 0 °C was saturated with ammonia. The suspension was heated to 100 °C for 3 h in a sealed tube. The clear solution obtained was concentrated to *ca.* 1 mL, and acetonitrile (20 mL) was added. The precipitate was collected by filtration, washed with acetonitrile, and dried to afford 0.19 g (72% yield) of a white powder, mp 167–168 °C (lit.[36] 180–182 °C). TLC (10% H_2O/CH_3CN) $R_f = 0.15$; HPLC (gradient A) $R_t = 7.0$ min. UV (H_2O) λ_{max} 260 nm (ϵ = 6300). ^{1}H NMR (D_2O) δ 7.58 (s, 1H, H4), 5.94 (t, J = 6.6 Hz, 1H, H1′), 4.51 (m, 1H, H3′), 4.07 (m, 1H, H4′), 3.83 (sp ABq, 2H, H5′), 2.48 (m, 2H, H2′), 1.88 (s, 3H, Me). ^{13}C NMR (10% CD_3OD/D_2O) δ 175.05, 155.93, 137.36, 115.87, 89.62, 87.86, 71.02, 61.67, 39.08, 13.68. High-resolution MS (positive-ion FAB) calcd for $C_{10}H_{16}N_3O_4$ [MH]$^+$ 242.1141, found 242.1142.

***N*-Benzoyldeoxy-5-methylisocytidine (3).** A cold suspension of d^{Me}-isoC (**2**) (0.148 g, 0.61 mmol) in dry pyridine (6.2 mL) was treated with benzoyl chloride (0.75 mL, 6 mmol, 10 equiv). The reaction mixture was allowed to warm up to room temperature, and stirring was continued for 3 h. The reaction mixture was chilled again to 0 °C and treated with 1 M sodium bicarbonate (15 mL). After 1 h the reaction mixture was extracted with dichloromethane (300 mL). The organic phase was washed with phosphate buffer (pH 7.0), dried (Na_2SO_4), and concentrated. The residue was diluted with pyridine/methanol/water (7.5 mL, 6.3:3:0.5), chilled to 0 °C, and treated with 2 M sodium hydroxide (5 mL in methanol/water, 1:1) for 30 min. Sodium hydrogen phosphate (30 mL, 1 M solution) was added and the mixture extracted with ethyl acetate (2 × 300 mL). The organic phase was dried (Na_2SO_4), evaporated, and coevaporated with *n*-heptane. The residue was purified by flash column chromatography (3–5% $MeOH/CH_2Cl_2$), followed by elution through a short column of basic alumina (7.5% $MeOH/CH_2Cl_2$) to remove residual benzoic acid, afferding 0.132 g (63% yield) of **3** as a white crystalline material, mp 173–174 °C dec. TLC (5% $MeOH/CH_2Cl_2$) $R_f = 0.26$. ^{1}H NMR (CD_3OD) δ 8.24 (unresolved q, J = 1.2 Hz, 1H, H6), 8.19 (m, 2H, Ph H2), 7.51 (m, 1H, Ph H4), 7.42 (m, 2H, Ph H3), 6.82 (t, J = 6.35 Hz, 1H, H1′), 4.43 (m, 1H, H3′), 4.03 (m, 1H, H4′), 3.88, 3.78 (m, ^2J = 12.1 Hz, ^3J = 3.0, 3.45 Hz, 2H, H5′), 2.53, 2.28 (m, 2H, H2′), 1.96 (d, J = 1.2 Hz, Me). ^{13}C NMR (CD_3OD) δ 178.92, 154.08, 138.35, 138.26, 133.33, 130.51, 129.16, 115.24, 89.49, 88.10, 71.67, 62.37, 42.24, 13.02. High-resolution MS (positive-ion FAB) calcd for $C_{17}H_{20}N_3O_5$ [MH]$^+$ 346.1403, found 346.1411.

5′-(Dimethoxytrityl)-*N*-benzoyldeoxy-5-methylisocytidine (4). A solution of *N*-Benzoyldeoxy-5-methylisocytidine (**3**) (0.13 g, 0.38 mmol) in dry pyridine (4 mL) was treated with DMAP (0.012 g, 0.1 mmol), triethylamine (0.1 mL, 0.7 mmol), and 4,4′-dimethoxytrityl chloride (0.165 g, 0.49 mmol). After 12 h at room temperature, the reaction mixture was chilled to 0–4 °C, treated with cold water (5 mL), and extracted with diethyl ether (450 mL). The organic phase was dried (Na_2SO_4) and concentrated, and the residue was coevaporated with *n*-heptane and acetonitrile. Flash column chromatography (0–1% $MeOH/CH_2Cl_2$) afforded 0.148 g (61% yield) of a slightly yellow foam. TLC (1.5% $MeOH/CH_2Cl_2$) $R_f = 0.23$. Recrystallization of a small amount from methanol afforded a white crystalline material (ca. 90% recovery), mp 182–185 °C dec. ^{1}H NMR (CD_3OD) δ 8.21 (m, 2H, Bz H2), 8.04 (unresolved q, J = 1.1 Hz, H6), 7.53–7.23 (m, 12H, DMT, Bz), 6.87 (m, 4H, DMT), 6.83 (t, J = 6.3 Hz, 1H, H1′), 4.57 (m, 1H, H3′), 3.76 (s, 6H, OMe), 3.89 and 3.39 (m, ^2J = 10.7 Hz, ^3J = 2.7, 3.3 Hz, 2H, H5′), 2.60 and 2.45 (m, 2H, H2′), 1.43 (d, J = 1.1 Hz, 3H, Me). High-resolution MS (positive-ion FAB) calcd for $C_{38}H_{38}N_3O_7$ [MH]$^+$ 648.2710, found 648.2715.

Phosphoramidite 5. A solution of 5′-DMT-*N*-benzoyldeoxy-5-methylisocytidine (**4**) (0.080 g, 0.124 mmol) in dry dichloromethane (2.5 mL) was treated with freshly distilled diisopropylethylamine (0.12 mL, 0.69 mmol), followed by 2-cyanoethyl *N,N*-diisopropylphosphoramidochloridite (0.040 mL, 0.17 mmol). After 1 h at room temperature, the reaction mixture was diluted with ethyl acetate (300 mL containing 5% Et_3N), washed with 1 M sodium bicarbonate (2 × 50 mL) and brine (2 × 50 mL), dried (Na_2SO_4), and concentrated. Flash column chromatography (25% hexane/5% Et_3N in CH_2Cl_2) afforded the diastereomeric phosphoramidites as a colorless foam (0.077 g, 73% yield), after concentration from dry acetonitrile. TLC (1% MeOH/0.5% Et_3N in CH_2Cl_2, pretreated plates) $R_f = 0.38$. ^{1}H NMR (CD_3CN) δ 8.25–8.21 (m, 2H, Bz H2), 7.91, 7.87 (two br s, 1H, H6), 7.58–7.28 (m, 12H, DMT, Bz), 6.94–6.90 (m, 4H, DMT), 6.84 (m, 1H, H1′), 4.70 (br m, 1H), 4.25–4.23 (m, 1H), 3.90–3.37 (m, 12H), 2.71–2.46 (m, 4H), 1.57 (br s, 3H), 1.27–1.09 (m, 12H). High-resolution MS (positive-ion FAB) calcd for $C_{47}H_{55}N_5O_8P$ [MH]$^+$ 848.3788, found 848.3821.

6-Methylthioxanthosine 5′-Monophosphate Barium Salt (18). 6-Methylthioguanosine 5′-monophosphate barium salt (**17**)[27] (0.3 g, 0.57 mmol) was dissolved in acetic acid/water (1:1, 10 mL) and the solution was chilled down to 0–4 °C. Sodium nitrite (40 mg, 0.58 mmol) in water (3.8 mL) was added dropwise (syringe pump) over a period of 30 min. The solution was stirred at 0–4 °C for an additional 30 min and then was treated with triethylamine, adjusting the pH to *ca.* 6. The solution was warmed to room temperature and then treated with barium acetate (0.37 g, 1.4 mmol). After 1 h at room temperature the solution was slowly added to vigorously stirred ethanol (120 mL). The white precipitate was separated by centrifugation, resuspended in water, and reprecipitated with ethanol. It was finally washed with ethanol and acetone. Drying afforded a slightly yellowish powder (0.257 g, 85% yield). HPLC (gradient A) $R_t = 9.0$ min. UV (0.25 M TEAA, pH 6.0) λ_{max} 246, 328 nm. ^{1}H NMR (D_2O) δ 8.33 (s, 1H, H8), 5.99 (d, J = 5.9 Hz, H1′), 4.73 (t, J = 5.5 Hz, 1H, H2′), 4.47 (m, 1H, H3′), 4.32 (br m, 1H, H4′), 3.99 (m, 2H, H5′), 2.67 (s, 3H, SCH_3). High-resolution MS (negative-ion FAB) calcd for $C_{11}H_{14}N_4O_8PS$ [M^{2-} + H^+]$^-$ 393.0270, found 393.0285.

Alkaline phosphatase removal of the 5′-phosphate group under standard conditions gave a product which was identical to an authentic sample of 6-methylthioxanthosine chemically synthesized from 5-amino-4-(aminocarbonyl)imidazolyl β-D-ribofuranoside (**6**) using multistep procedure reported by Yamazaki et al.[37]

***N*^{6}-(6-Aminohexyl)isoguanosine 5′-Monophosphate Triethylammonium Salt (19).** 6-Methylthioxanthosine 5′-monophosphate triethylammonium salt (28 mg, 0.0568 mmol, obtained by passing an aqueous solution of the Ba^{2+} salt **18** through a Dowex 50X8-200 column and collecting the eluent over 50 mM triethylammonium bicarbonate buffer (TEAB), pH 7.5, followed by evaporation and coevaporation with acetonitrile) was dissolved in methanol (1.5 mL) and treated with 1,6-diaminohexane (0.22 g, 1.89 mmol, 33 equiv) in methanol (3 mL). The solution was refluxed under Ar for 12 h. The methanol was evaporated, and the residue was dissolved in water (100 mL). The pH was adjusted to *ca.* 7.2 with carbon dioxide, and the solution was loaded on a Sephadex A-25 column equilibrated with 50 mM TEAB, pH 7.5. The column was eluted with a gradient of 50–500 mM TEAB, pH 7.5. The pure fractions were

(33) Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923–2925.
(34) Chern, J.-W.; Lin, G.-S.; Chen, C.-S.; Townsend, L. B. *J. Org. Chem.* **1991**, *56*, 4213–4218.
(35) Ludwig, J. *Acta Biochim. Biophys. Acad. Sci. Hung.* **1981**, *16*, 131–133.
(36) Kowollik, G.; Langen, P. *J. Prakt. Chem.* **1968**, *37*, 311–318.
(37) Yamazaki, A.; Kumashiro, I.; Takenishi, T.; Ikehera, M. *Chem. Pharm. Bull.* **1968**, *16*, 2172–2181.

combined, evaporated to dryness, and coevaporated several times with methanol to afford a colorless, viscous material (24 mg, 75% yield). HPLC (gradient A) R_t = 8.2 min. UV (H_2O) λ_{max} 250, 296 nm. ^1H NMR (D_2O) δ 8.24 (br s, 1H, H8), 5.89 (d, *J* = 5.72 Hz, 1H, H1′), 4.68 (t, *J* = 5.3 Hz, 1H, H2′), 4.47 (br m, 1H, H3′), 4.32 (br m, 1H, H4′), 4.01 (br m, 2H, H5′), 2.95 (m, 4H, CH_2N), 1.63 (m, 4H, NCH_2CH_2), 1.37 (m, 4H, $NCH_2CH_2CH_2$).

N^6-(6-((Trifluoroacetyl)amino)hexyl)isoguanosine 5′-Monophosphate Et_3NH^+ Salt (20). Trifluoroacetylation was performed according to a procedure reported for 6-aminohexyl-dAMP.[38] Purification on a Sephadex A-25 column (50–500 mM TEAB, pH 7.5) afforded a white solid after coevaporation with acetonitrile. HPLC (gradient A) R_t = 17.7 min. ^1H NMR (CD_3OD/D_2O) δ 8.12 (br s, 1H, H8), 5.86 (d, *J* = 6.0 Hz, 1H, H1′), 4.51 (m, 1H, H2′), 4.37 (m, 1H, H3′), 4.23 (br m, 1H, H4′), 4.10 (br m, 2H, H5′), 3.30–2.99 (m, 10H, NCH_2 including Et_3N), 1.69 (br m, 2H, TFA-$NHCH_2CH_2$), 1.57 (br m, 2H, NCH_2CH_2), 1.41 (br m, 4H, $NCH_2CH_2CH_2$), 1.32–1.25 (m, 9H, CH_3[Et_3N]). UV (H_2O) λ_{max} 250, 294 nm. High-resolution MS (negative-ion FAB) calcd for $C_{18}F_3H_{25}N_6O_9P$ [M^{2-} + H^+]$^-$ 557.1373, found 557.1378.

N^6-(6-((Trifluoroacetyl)amino)hexyl)isoguanosine 5′-Triphosphate Et_3NH^+ Salt (21). N^6-(6-((Trifluoroacetyl)amino)hexyl)isoguanosine 5′-monophosphate **20** (20 mg, 0.03 mmol) was dissolved in dry DMF (0.7 mL) and treated with 1,1′-carbonyldiimidazole (25 mg, 0.15 mmol) in DMF (0.35 mL). After 2.5 h at room temperature, methanol (10 μL, 0.243 mmol) was added and stirring continued for 30 min. Tributylammonium pyrophosphate (0.1 g, 0.2 mmol) in DMF (0.3 mL) was added. After 36 h at room temperature the reaction mixture was filtered and concentrated. The residue was dissolved in 100 mM triethylammonium bicarbonate buffer, pH 7.5, (10 mL) and loaded onto a Sephadex A-25 column equilibrated with the same buffer at 4 °C. The column was eluted with a gradient from 100 mM to 1.0 M triethylammonium bicarbonate, pH 7.5. The pure fractions were combined and evaporated with ethanol several times (to remove residual buffer), affording 23 mg (68% yield) of a viscous, colorless material. HPLC (gradient A) R_t = 17.9 min. UV (H_2O) λ_{max} 250, 296 nm. ^1H NMR (CD_3OD) δ 8.17 (s, 1H, H8), 5.84 (d, *J* = 5.4 Hz, 1H, H1′), 4.55 (br m, 1H), 4.48 (m, 1H), 4.28–4.18 (m, 3H), 3.5–3.0 (m, CH_2N including Et_3N), 1.68 (br m, 2H), 1.57 (br m, 2H), 1.41 (br, 4H), 1.30 (br, CH_3 [Et_3N]). High-resolution MS (negative-ion FAB) calcd for $C_{18}F_3H_{27}N_6O_{15}P_3$ [M^{4-} + $3H^+$]$^-$ 717.0699, found 717.0684.

N^6-(6-Aminohexyl)isoguanosine Triphosphate Triethylammonium Salt (22). The TFA-protected triphosphate **21** (12 mg, *ca.* 10 μmol) was dissolved in water (0.75 mL), and the pH was adjusted to 11–11.5 with 1 N NaOH. The reaction progress was monitored using HPLC by quenching 1-μL aliquots with 250 mM triethylammonium acetate, pH 6.0 (20 μL). After 12 h at room temperature (75% conversion), water (3 mL) was added and the pH was adjusted to *ca.* 7.5 using carbon dioxide. The solution was then diluted to 15 mL and loaded onto a Sephadex A-25 column equilibrated with 100 mM triethylammonium bicarbonate, pH 7.5. The column was eluted with a 100 mM–1.0 M triethylammonium bicarbonate, pH 7.5, gradient. The pure fractions were combined and evaporated. Coevaporation with water/methanol afforded 6 mg (75% yield based on 75% conversion) of a colorless material. UV (H_2O) λ_{max} 250, 296 nm. HPLC (gradient A) R_t = 8.7 min. ^1H NMR (CD_3OD/D_2O) δ 8.30 (br, 1H, H8), 5.88 (br, 1H, H1′), 4.58 (m, 1H), 4.50 (m, 1H), 4.25 (br m, 3H), 3.55–2.95 (m, CH_2 including Et_3N), 1.66 (br m, 4H), 1.42 (br m, 4H), 1.35–1.25 (m, CH_3 [Et_3N]). High-resolution MS (negative-ion FAB) calcd for $C_{16}H_{28}N_6O_{14}P_3$ [M^{4-} + $3H^+$]$^-$ 621.0876, found 621. 0892. Stock solutions of the trisphosphate (40 and 20 mM) were made by dissolving the product in 50 mM Tris-HCl, pH 7.5, and adjusting the pH back to 7.5 using aliquots of 1.0 M Tris-HCl, pH 8.4. The concentration was estimated by measuring the OD at 294 nm (ε = 9800 M^{-1} cm^{-1}).

DNA and RNA Reagents and Materials. Sterilized 0.1% DEPC-treated water[39] was used for all RNA manipulations. Polyacrylamide gel electrophoresis was performed in 1X TBE buffer.[39] Autoradiography was carried out using Amersham Hyperfilm MP or Kodak X-Omat film. T7 RNA polymerase and BSA (RNase/DNase free) were obtained from Pharmacia. Phosphodiesterase I type VIII and ribonucleases A, T_1, and T_2 were purchased from Sigma. Calf intestine alkaline phosphatase, RNAse inhibitor, and glycogen were purchased from Boehringer Mannheim. T4 polynucleotide kinase was obtained from New England Biolabs. α-^{32}P-UTP (800 or 3000 Ci/mmol), α-^{32}P-ATP (3000 Ci/mmol), and γ-^{32}P-ATP (5000 Ci/mmol) were obtained from Amersham.

Synthesis and Purification of Oligonucleotides. Oligonucleotides were prepared on an Applied Biosystem Model 394B DNA synthesizer using 2-cyanoethyl phosphoramidites. Deoxy-5-methylisoC phosphoramidite **5** was prepared as described and utilized in automated synthesis as a 0.1 M solution in acetonitrile. Oligonucleotides were purified on preparative 15–20% polyacrylamide gels and isolated by crushing and soaking the UV-shadowed bands in 1X Tris–Borate–EDTA buffer,[39] followed by desalting (SepPack-Waters). The concentration of single-stranded oligonucleotides was determined at 260 nm using the following molar extinction coefficients for each base: 15 400 (A), 11 700 (G), 7300 (C), 8800 (T), 10 100 (U), 6300 (^{Me}isoC), 4600 (isoG and N^6-substituted isoG derivatives) M^{-1} cm^{-1}. The synthetic oligonucleotides were enzymatically digested using snake venom phosphodiesterase I, treated with calf intestine alkaline phosphatase, and analyzed by HPLC in order to confirm the presence of the unnatural base in the templates. Additionally, the purified oligonucleotides were ethanol precipitated before annealing.[12b]

Transcription Reactions. *In vitro* small-scale transcription reactions were performed according to the general procedure given by Milligan and Uhlenbeck.[12] The 27-mer single-stranded DNA templates were annealed to the 18-mer promoter in 100 mM NaCl–TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.8) by heating a 1:1 mixture to 90 °C for 2 min and cooling the solution slowly to room temperature. The solutions were diluted to 200 μL with TE buffer (giving a final template concentration of 5 μM) and stored at −20 °C. Each transcription reaction contained 250 nM template and 120 units of T7 RNA polymerase (Pharmacia) in 40 mM Tris-HCl, pH 8.1 (at 37 °C), 20 mM $MgCl_2$, 5 mM DTT, 1 mM spermidine hydrochloride, 0.01% Triton X-100, 50 μg/mL BSA, 4 mg/mL PEG 8000, RNase inhibitor (40 units), 1 mM NTPs, and 10 μCi α-^{32}P-UTP or α-^{32}P-ATP, in a total volume of 20 μL. After 3 h at 37 °C the reactions were quenched upon addition of 55 μL of loading buffer (7 M urea in 10 mM Tris-HCl, 1 mM EDTA, pH 7.8, 0.02% xylene cyanol FF, and bromophenol blue), heated to 75 °C for 3 min, and loaded onto an analytical 20% polyacrylamide gel. In other cases, the RNA was ethanol precipitated. Large-scale transcription experiments have been performed under similar conditions in a total volume of 250 μL. For these reactions, additional DTT (5 mM) was added and the concentration of NTPs was 1 mM (CTP, GTP, and 6-AH-isoGTP) and 0.8 mM (ATP and UTP). After incubation with T7 RNA polymerase (1500 units) for 3 h at 37 °C, the magnesium pyrophosphate precipitate was removed by centrifugation, and the RNA was precipitated by adding ammonium acetate (1.6 M final concentration) and ethanol (three volumes) using glycogen (300 μg) as a carrier. The RNA was resuspended in 7 M urea loading buffer, heated up to 75 °C for 3 min, and loaded onto a preparative 1.5 mm thick gel. The gel was UV shadowed; the bands were excised, cut into small pieces, and extracted with 0.5 M ammonium acetate for 1 h at 37 °C. The RNA was precipitated, dried, resuspended in 0.1% DEPC/water (100 μL), and kept at −80 °C. 5′-Labeling of the large-scale transcription products was performed by alkaline phosphatase removal of the 5′-triphosphate followed by 5′-phosphorylation using T4 polynucleotide kinase and γ-^{32}P-ATP according to general protocols.[39]

RNA Digestion. Enzymatic digestion of the large-scale transcription products was carried out according to a procedure reported by Buck et al.[40] Approximately 1 nmol of RNA was digested with snake venom phosphodiesterase I, calf intestine alkaline phosphatase, and RNase A in 80 mM ammonium formate buffer, pH 7.6, for 12 h at 37 °C, followed by additional treatment with RNase T_1 and RNase T_2 for 4 h at 37 °C. The ribonucleoside mixture obtained was analyzed by reversed-phase HPLC using gradient B.

Post-Transcriptional Modification. The labeled RNA transcripts (in 5 μL of 0.1% DEPC/water) were diluted with 0.4 M sodium bicarbonate (5 μL, pH 8.17) and treated with *N*-hydroxysuccinimidobiotin or EDTA dianhydride (2 μL of 1% w/v DMF solutions) for 90 min at room temperature. The reactions were quenched by adding 100 mM Tris-HCl buffer, pH 7.2, (5 μL) and DEPC/water (33 μL), followed by precipitation using glycogen as a carrier, as described above.

Acknowledgment. We are grateful for financial support from the National Institutes of Health (Grant GM-27681) and a Chaim Weizmann Postdoctoral Fellowship to Y.T.

(38) Gebeyehu, G.; Rao, P. Y.; Soochan, P.; Simms, D. A.; Klevan, L. *Nucleic Acids Res.* **1987**, *15*, 4513–4534.
(39) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*, 2nd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.
(40) Buck, M.; Connic, M.; Ames, B. N. *Anal. Biochem.* **1983**, *129*, 1–13.