Efficient Transfer of Information from Hexitol Nucleic Acids to RNA during Nonenzymatic Oligomerization

Igor A. Kozlov,[†] Bart De Bouvere,[#] Arthur Van Aerschot,[#] Piet Herdewijn,[#] and Leslie E. Orgel*,[†]

Contribution from The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, California 92186, and the Laboratory of Medical Chemistry, Rega Institute, Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Received February 11, 1999

Abstract: Hexitol nucleic acids (HNAs) are DNA analogues that contain the standard nucleoside bases attached to a phosphorylated 1,5-anhydrohexitol backbone. We find that HNAs support efficient information transfer in nonenzymatic template-directed reactions. HNA heterosequences appeared to be superior to the corresponding DNA heterosequences in facilitating synthesis of complementary oligonucleotides from nucleoside-5′-phosphoro-2-methyl imidazolides.

Introduction

Information transfer from DNA heterosequences to RNA by nonenzymatic template-directed synthesis with use of activated nucleoside 5′-phosphates has been studied in detail.[1−6] Nonenzymatic information transfer from peptide nucleic acid (PNA) heterosequences to RNA has also been demonstrated, but it is less efficient than that for the corresponding DNA or RNA templates.[7] Hexitol nucleic acids (HNAs) are new DNA analogues containing the standard nucleoside bases and a phosphorylated 1,5-anhydrohexitol backbone (Figure 1a). In contrast with previously described DNA or RNA analogues that have backbones based on a six-membered sugar ring,[8−10] HNA oligomers form A-type duplexes with complementary DNA or RNA oligomers.[11,12] These duplexes are more stable than duplexes formed by the corresponding DNA or RNA oligomers.[11,13] In a preliminary study we found that an HNA decacytidylate oligomer facilitates oligomerization of guanosine 5′-phosphoro-2-methylimidazole (2-MeImpG, Figure 1b) more efficiently than DNA (dC_{10}) or RNA (C_{10}) templates presumably due to the partial preorganization of HNA in an A-type nucleic acid helical structure.[14] This finding led us to investigate information transfer from HNA oligomers to RNA. Here we report experiments on information transfer from HNA heterosequences to RNA by nonenzymatic template-directed synthesis. We carried out the oligomerization reactions on HNA C_4XC_4 templates (X = G, T, or A) and compared them with reactions on the corresponding DNA C_4XC_4 templates (Figure 1). The reaction conditions were chosen to facilitate direct comparison with previously published results.[5−7]

Results

Extension of a ^{32}p(dG)$_3$G Primer on C_4XC_4 (X = G, T, or A) HNA and DNA Templates. The product distributions in the reactions of ^{32}P-labeled p(dG)$_3$G with 2-MeImpG or an equimolar mixture of 2-MeImpG with 2-MeImpC, 2-MeImpA, or 2-MeImpU on C_4XC_4 (X = G, T, or A) HNA and DNA templates are illustrated in Figure 2. Significant extension of the primer p(dG)$_3$G with 2-MeImpG alone does not take place on C_4GC_4 or C_4AC_4 templates in either the DNA or the HNA series (Figure 2, lanes 1−2, 9, and 10). In the presence of a C_4TC_4 HNA or DNA template, conversion of about 50% of the primer to products p(dG)$_3$G(G)$_n$ up to the octamer is observed (Figure 2, lanes 5 and 6). This is in agreement with previously reported data for incorporation of a G residue opposite T on DNA templates, presumably due to the formation of G−T "wobble" pairs.[6]

The extension of the primer p(dG)$_3$G with an equimolar mixture of 2-MeImpG with 2-MeImpC in the presence of a C_4GC_4 HNA template leads to conversion of more than 95% of the primer to p(dG)$_3$GC(G)$_3$ and p(dG)$_3$GC(G)$_4$ products while in the presence of a C_4GC_4 DNA template conversion is less than 40%, and the longest product formed is p(dG)$_3$GC(G)$_3$ (Figure 2, lanes 3 and 4). The extension of the primer p(dG)$_3$G with an equimolar mixture of 2-MeImpG with 2-MeImpA on a C_4TC_4 HNA template leads to conversion of more than 95% of

* To whom correspondence should be addressed. Phone: (619) 453-4100. Fax (619) 558-7359. E-mail: orgel@salk.edu.
[†] The Salk Institute for Biological Studies.
[#] Katholieke Universiteit Leuven.

(1) Acevedo, O. L.; Orgel, L. E. *J. Mol. Biol.* **1987**, *197*, 187−193.
(2) Hill, A. R., Jr.; Orgel, L. E.; Wu, T. *Origins Life Evol. Biosphere* **1993**, *23*, 285−290.
(3) Joyce, G. F.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1984**, *176*, 279−306.
(4) Joyce, G. F. *Nonenzymatic template-directed synthesis of informational macromolecules*; Joyce, G. F., Ed.; Cold Spring Harbor Press: Cold Spring Harbor, NY, 1987; Vol. LII, pp 41−51.
(5) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 5496−5501.
(6) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 7964−7969.
(7) Schmidt, J. G.; Nielsen, P. E.; Orgel, L. E. *Nucleic Acid Res.* **1997**, *25*, 4797−4802.
(8) Eschenmoser, A. *Origins Life* **1994**, *24*, 389−423.
(9) Eschenmoser, A. *Origins Life* **1997**, *27*, 535−553.
(10) Beier, M.; Reck, F.; Wagner, T.; Krishnamurthy, R.; Eschenmoser, A. *Science* **1999**, *283*, 699−703.
(11) Hendrix, C.; Rosemeyer, H.; Verheggen, I.; Seela, F.; Van Aerschot, A.; Herdewijn, P. *Chem. Eur. J.* **1997**, *3*, 110−120.
(12) Winter, H. D.; Lescrinier, E.; Van Aerschot, A.; Herdewijn, P. *J. Am. Chem. Soc.* **1998**, *120*, 5381−5394.
(13) Hendrix, C.; Rosemeyer, H.; De Bouvere, B.; Van Aerschot, A.; Seela, F.; Herdewijn, P. *Chem. Eur. J.* **1997**, *3*, 1513−1520.
(14) Kozlov, I. A.; Politis, P. K.; Van Aerschot, A.; Busson, R.; Herdewijn, P.; Orgel, L. E. *J. Am. Chem. Soc.* **1999**, *121*, 2033−2036.

10.1021/ja990440u CCC: $18.00
Published on Web 06/05/1999

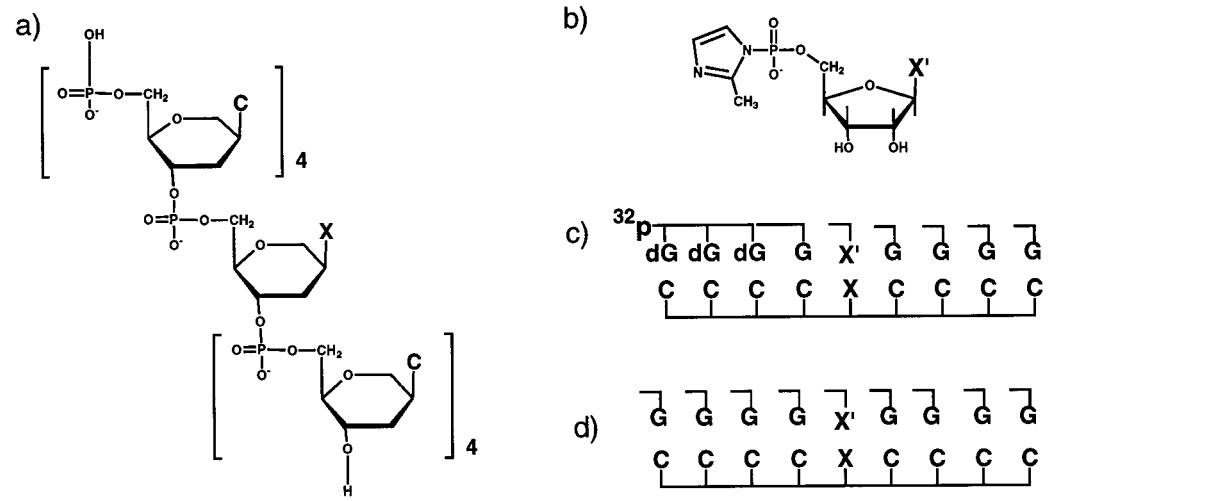

Figure 1. Structure of (a) hexitol nucleic acid (HNA) templates (X = G, T, or A) and (b) activated nucleoside 5′-phosphates (2-MeImpX′), where X′ can be A, G, C, or U. (c) Schematic representation of the primer $^{32}p(dG)_3G$ extension reaction with 2-MeImpG and 2-MeImpX′ on HNA or DNA C_4XC_4 templates, where X = G, T, or A and X′ is the complement of X. (d) Schematic representation of the template-directed oligomerization of 2-MeImpG and 2-MeImpX′ on DNA or HNA templates which have the structure C_4XC_4, where X = G, T, or A and X′ is the complement of X.

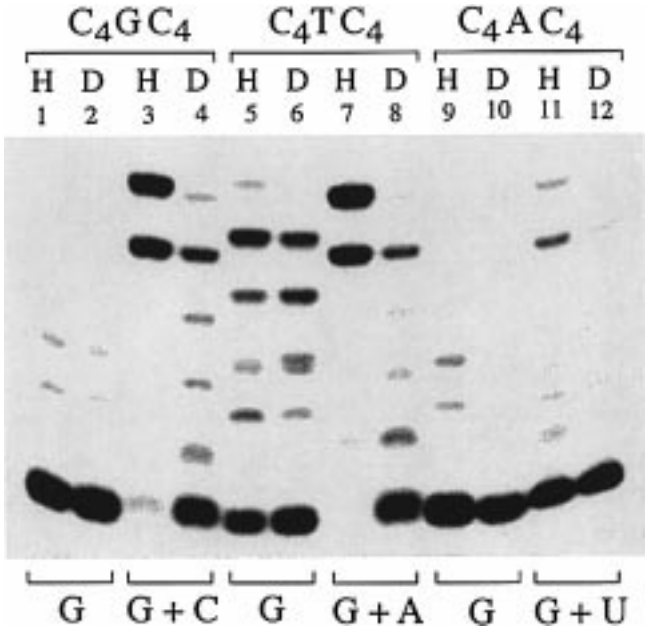

Figure 2. Extension of a $^{32}p(dG)_3G$ primer on HNA and DNA C_4XC_4 templates (X = G, T, or A). H represents an HNA template, D represents a DNA template, G represents 2-MeImpG, G + C represents a mixture of 2-MeImpG and 2-MeImpC, G + A represents a mixture of 2-MeImpG and 2-MeImpA, and G + U represents a mixture of 2-MeImpG and 2-MeImpU. The lowest band in the picture corresponds to the $^{32}p(dG)_3G$ primer.

the primer to a mixture of $p(dG)_3GA(G)_3$ and $p(dG)_3GA(G)_4$ while in the presence of a C_4TC_4 DNA template conversion is less than 40%, and the longest product formed is $p(dG)_3GA(G)_3$ (Figure 2, lanes 7 and 8). Strikingly, $p(dG)_3G(G)_n$ (n = 1−4) which are obtained with 2-MeImpG alone on C_4TC_4 HNA or DNA templates are not formed when a mixture of 2-MeImpG with 2-MeImpA is used. The extension of the primer $p(dG)_3G$ with an equimolar mixture of 2-MeImpG and 2-MeImpU on a C_4AC_4 HNA template leads to conversion of about 15% of the primer to $p(dG)_3GU(G)_3$ and $p(dG)_3GU(G)_4$, while a C_4AC_4 DNA template catalyzes conversion of less than 2% of the primer into $p(dG)_3GU(G)_n$ products no longer than the octamer (Figure 2, lanes 11 and 12).

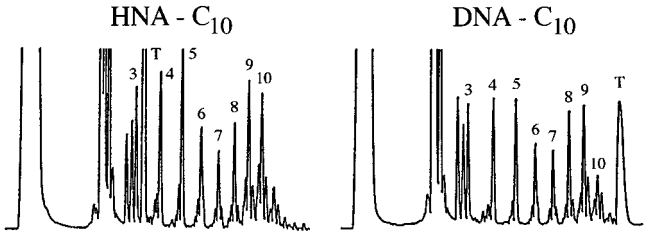

Figure 3. Elution profiles after HPLC on an RPC5 of the products from the oligomerization of 2-MeImpG on an HNA C_{10} template or a DNA C_{10} template after 14 days. The numbers above the peaks indicate the length of the all 3′-5′-linked oligo$(G)_n$ products, and T indicates the template.

The presence of predominantly 3′-5′ internucleotide bonds upstream and downstream of the X residue in $p(dG)_3GX(G)_n$ primer extension products (X = C, U or A) was confirmed by RNase hydrolysis. The CpG and UpG internucleotide bonds in $p(dG)_3GC(G)_n$ and $p(dG)_3GU(G)_n$ products were cleaved with more than 90% efficiency by RNase A, an enzyme that cleaves 3′-5′ internucleotide bonds after C and U residues. The ApG internucleotide bond in $p(dG)_3GA(G)_n$ oligomers was cleaved with more than 90% efficiency with RNase U_2, an enzyme which cleaves 3′-5′ internucleotide bonds after A residues. The GpX internucleotide bonds in $p(dG)_3GX(G)_n$ products were cleaved with RNase T_1, an enzyme that cleaves 3′-5′ internucleotide bonds after G residues, with more than 90% efficiency, for X = U, C, or A.

Oligomerization of Activated Mononucleotides on C_4XC_4 (X = G, T, or A) HNA and DNA Templates. In the absence of a template, 0.1 M aqueous solutions of 2-MeImpG yield as products only dimers and smaller amounts of trimers. The presence of HNA C_{10} or DNA C_{10} templates leads to efficient oligomerization of 2-MeImpG (Figure 3). The major peaks on the HPLC profiles (Figure 3) correspond to all 3′-5′-linked oligo$(G)_n$ ranging in length from the dimer to 10-mer on an HNA template and from the dimer to 9-mer on a DNA template.[15] The presence of a C_4GC_4 or a C_4AC_4 HNA template leads to oligomerization of 2-MeImpG alone to give large amounts of G_4 and small amounts of G_5 (Figure 4) while the presence of a

(15) Kozlov, I. A.; Politis, P. K.; Pitsch, S.; Herdewijn, P.; Orgel, L. E. *J. Am. Chem. Soc.* **1999**, *121*, 1108−1109.

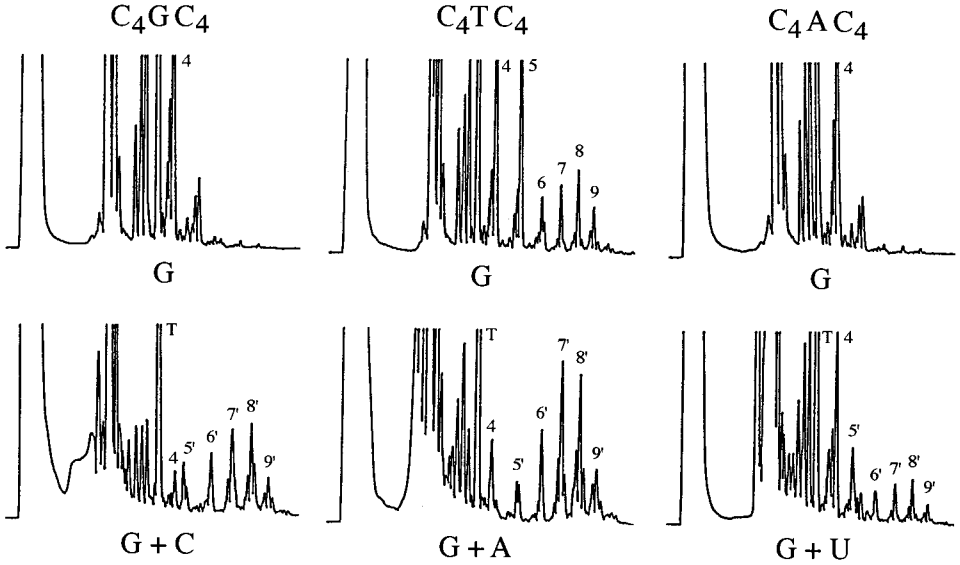

Figure 4. Elution profiles after HPLC of the products from the oligomerization on HNA C_4XC_4 templates (X = G, T, or A) of 2-MeImpG (G), a mixture of 2-MeImpG and 2-MeImpC (G + C), a mixture of 2-MeImpG and 2-MeImpA (G + A), or a mixture of 2-MeImpG and 2-MeImpU (G + U) after 14 days. The numbers above the peaks indicate the length of the all 3′-5′-linked oligoribonucleotide products, and T indicates the template.

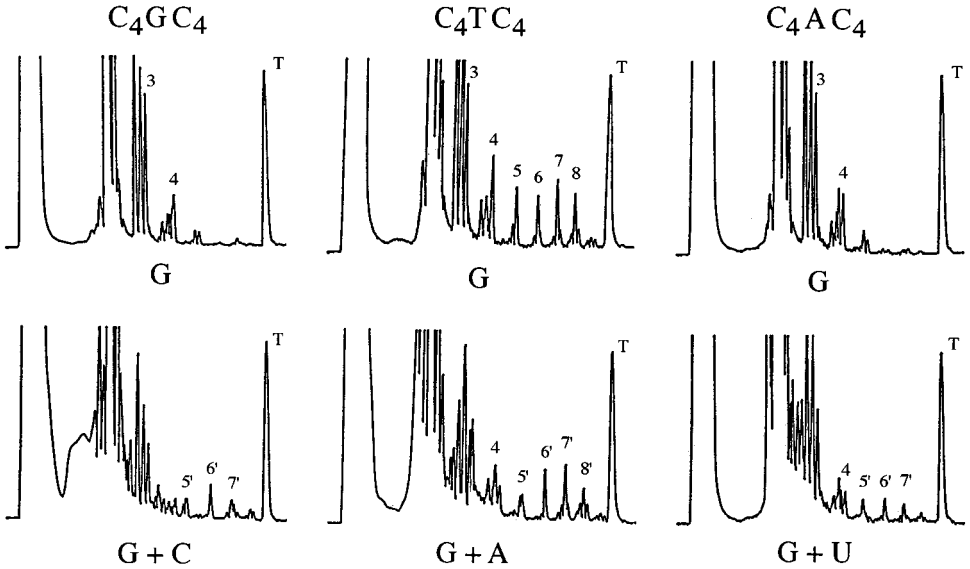

Figure 5. Elution profiles after HPLC of the products from the oligomerization on DNA C_4XC_4 templates (X = G, T, or A) of 2-MeImpG (G), a mixture of 2-MeImpG and 2-MeImpC (G + C), a mixture of 2-MeImpG and 2-MeImpA (G + A), and a mixture of 2-MeImpG and 2-MeImpU (G + U) after 14 days. The numbers above the peaks indicate the length of the all 3′-5′-linked oligoribonucleotide products, and T indicates the template.

C_4GC_4 or a C_4AC_4 DNA template leads to oligomerization of 2-MeImpG to give products only up to G_3 with small amounts of G_4 (Figure 5). The presence of a C_4TC_4 HNA or DNA template leads to oligomerization of 2-MeImpG to give products up to G_9 (Figures 4 and 5).

An equimolar mixture of 2-MeImpG with 2-MeImpC in the presence of a C_4GC_4 HNA template leads to generation of G_4CG_n products up to a nonamer (Figure 4) while a C_4GC_4 DNA template catalyzes much less efficient formation of G_4CG_n products not longer than a heptamer (Figure 5). An equimolar mixture of 2-MeImpG with 2-MeImpA in the presence of a C_4TC_4 HNA template leads to very efficient generation of G_4AG_n products up to a nonamer (Figure 4) but again a C_4TC_4 DNA template catalyzes much less efficient formation of G_4AG_n products not longer than an octamer (Figure 5). As in the case of the extension of a primer, products G_n which were obtained in the presence of 2-MeImpG alone on C_4TC_4 HNA or DNA templates are not present when a mixture of 2-MeImpG with 2-MeImpA is used. An equimolar mixture of 2-MeImpG with 2-MeImpU in the presence of a C_4AC_4 HNA template leads to formation of G_4UG_n products up to a nonamer (Figure 4) but a C_4AC_4 DNA template facilitates much less efficient formation of G_4UG_n products not longer than a heptamer (Figure 5).

To determine the nature of internucleotide bonds after the X residue in major products, we purified the material corresponding to the numbered 8-mer peaks in the lower panels of Figure 4. These products were then hydrolyzed by appropriate RNAses and were analyzed by HPLC (data not shown). The CpG and UpG internucleotide bonds in G_4CG_3 and G_4UG_3 oligomers were cleaved with more than 90% efficiency by RNase A. The ApG internucleotide bond in G_4AG_3 oligomers was cleaved with more than 90% efficiency by RNase U_2. These experiments show the presence of 3′-5′-linked internucleotide bonds down-

stream of the X residue in G_4XG_3 oligomers. However, we did not analyze the material corresponding to side peaks in the lower chromatograms of Figure 4. They must be pyrophosphate-capped oligomers and oligomers terminated by a 2′-5′-linked residue.[1−7]

Discussion

In previous publications we have reported that the synthesis of 3′-5′-linked oligo G's is more efficient[14] and more enantioselective[15] on an HNA C_{10} template than on a corresponding RNA or DNA template. The present report extends these results to templates of the type HNA C_4XC_4, where X = A, T, C. We find that each of these three templates facilitates the synthesis of its 3′-5′-linked complement, and that in each case synthesis on an HNA oligomer is more efficient than synthesis on the corresponding DNA oligomer. It appears that HNA templates are, in general, superior to RNA or DNA templates in facilitating synthesis of complementary oligonucleotides from nucleoside-5′-phosphoro-2-methyl imidazolides. We believe that this reflects the partial preorganization of HNA oligomers into the A structure of the double-helical nucleic acids.[14] The importance of the A structure in template-directed reactions was first pointed out by Gobel and co-workers.[22,23]

Another indication of the superiority of HNA templates is the nature of the products formed when C_4AC_4 and C_4GC_4 HNA or DNA templates are incubated with 2-MeImpG in the absence of another activated nucleotide. When HNA is used as the template, synthesis proceeds efficiently to produce the 3′-5′-linked G_4 product. When DNA is used as the template, the reaction is not regiospecific and yields little product longer than the trinucleotide. Again, we believe that more extensive preorganization of HNA than of DNA in the nucleic acid A structure is responsible for the difference.

We do not suggest that HNAs were prebiotic molecules, since we do not see any easy synthetic route to the monomers. Nonetheless, our results, together with those reported in numerous other publications on nucleic acid analogues with modified backbones,[8,9,16−20] emphasize an important question raised by Eschenmoser and co-workers.[10] Why were RNA and DNA chosen as the genetic materials for all life on the Earth? Do the standard nucleic acids have some intrinsic advantage that we do not as yet recognize or is the choice of RNA a "frozen accident" that reflects the availability of nucleotides rather than their analogues at the dawn of the RNA world?

Experimental Section

Unless otherwise noted, all chemicals were reagent grade, purchased from commercial sources, and were used without further purification. Nucleotide 5′-phosphoro-2-methylimidazoles (2-MeImpX′, X′ = G, C, A, U) were obtained by a published method[3] in at least 95% yield. The hexitol nucleic acid templates were synthesized and purified as described.[11] The oligodeoxyribonucleotides were synthesized and purified as described.[21]

Reaction conditions for the oligomerization of 2-MeImpG (or its mixture with an equal amount of 2-MeImpC, 2-MeImpA, or 2-MeImpU) on various templates were chosen to permit comparison with earlier published work.[7,21] The reactions were run for 14 days at 0 °C in 0.2 M 2,6-lutidine−HCl buffer (pH 7.9 at 25 °C) containing 1.2 M NaCl, 0.2 M $MgCl_2$, 0.5 mM of a template, and 0.1 M of 2-MeImpG. The reaction mixtures were analyzed by HPLC on an RPC5 column as previously described.[21] HPLC fractions corresponding to each of the oligomerization products containing eight nucleotide residues were collected and dialyzed against water (Spectrum dialysis membranes) for 16 h at 4 °C, and then used in experiments on RNase digestion.

Reaction conditions for $p(dG)_3G$ primer extension with 2-MeImpG (or its mixture with an equal amount of 2-MeImpC, 2-MeImpA, or 2-MeImpU) on different templates were again chosen to permit comparison with earlier published work.[7,21] The reaction mixtures were incubated for 5 days at 0 °C in 0.2 M 2,6-lutidine buffer (pH 7.9 at 25 °C) containing 1.2 M NaCl, 0.2 M $MgCl_2$, 20 μM template, 20 nM primer, and 50 mM 2-MeImpG. The reaction mixtures were analyzed by electrophoresis in 20% PAG containing 8 M urea as previously described.[7,21] The reaction mixtures were desalted on a Nensorb column (Nen DuPont) prior to RNase digestion.

Digestion with RNases T_1 and U_2 (10U; Pharmacia) and RNase A (5U; Sigma) was carried out in 12 μL of 25 mM Na−Citrate buffer containing 1 mM EDTA, 6 M urea, and about 1 nmol of RNA 8-mers or about 0.01 pmol of total primer extension products at pH 5.0 (RNases T_1 and A) or pH 3.5 (RNase U_2) at 50 °C for 20 min. In the case of RNase T_1 digestion, the reaction mixture was first incubated for 45 min with RNases T_1, after which an additional portion of RNases T_1 (10U; Sigma) was added and the reaction mixture was incubated for an additional 45 min. The reaction mixtures were analyzed by HPLC on an RPC5 column[21] or by electrophoresis in 20% PAG containing 8 M urea.[7,21]

Acknowledgment. We thank Aubrey R. Hill, Jr., and G. Schepers for technical assistance and Bernice Walker for manuscript preparation. This work was supported by Grants NAG5-4118 and NAG5-4546 NSCORT/Exobiology from NASA and Grant GOA 97/11 from KU Leuven.

JA990440U

(16) De Mesmaeker, A.; Altmann, K. H.; Waldner, A.; Wendeborn, S. *Curr. Opin. Struct. Biol.* **1995**, *5*, 343−355.
(17) Nielsen, P.; Egholm, M.; Buchardt, O. *Bioconjug. Chem.* **1994**, *5*, 3−7.
(18) Nielsen, P. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167−183.
(19) Kuwahara, M.; Arimitsu, M.; Sisido, M. *J. Am. Chem. Soc.* **1999**, *121*, 256−257.
(20) Arya, D. P.; Bruice, T. C. *J. Am. Chem. Soc.* **1998**, *120*, 12419−12427.
(21) Kozlov, I. A.; Pitsch, S.; Orgel, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13448−13452.
(22) Kurz, M.; Gobel, K.; Hartel, C.; Gobel, M. W. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 842−845.
(23) Kurz, M.; Gobel, K.; Hartel, C.; Gobel, M. W. *Helv. Chim. Acta* **1998**, *81*, 1156−1180.