A Comparison of RNA with DNA in Template-Directed Synthesis

by **Magdalene Zielinski**, **Igor A. Kozlov**[1]), and **Leslie E. Orgel**[2])*

The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, CA 92186, USA

Dedicated to Prof. Dr. *Albert Eschenmoser* on the occasion of his 75th birthday

Nonenzymatic template-directed copying of RNA sequences rich in cytidylic acid using nucleoside 5′-(2-methylimidazol-1-yl phosphates) as substrates is substantially more efficient than the copying of corresponding DNA sequences. However, many sequences cannot be copied, and the prospect of replication in this system is remote, even for RNA. Surprisingly, wobble-pairing leads to much more efficient incorporation of G opposite U on RNA templates than of G opposite T on DNA templates.

Introduction. – The nonenzymatic synthesis of complementary oligonucleotides on DNA or RNA templates is of considerable interest in the context of chemical evolution. Consequently, template-directed synthesis on DNA and RNA templates has received a good deal of attention [1–10]. One particularly favorable system utilizes nucleoside 5′-(2-methylimidazol-1-yl phosphates) (*Fig. 1,a*) as activated monomers, and cytidine-rich oligonucleotides or oligodeoxynucleotides as templates. Most of the early experiments were carried out at a time when sequence-specific RNA synthesis was difficult, so DNA templates and deoxynucleotide-containing primers were used in place of RNA oligomers [11–16], although RNA oligomers are more relevant to chemical evolution. Only in the case of oligo(C)s and oligo(dC)s has it been shown that ribo and deoxy templates behave in a qualitatively similar way [14][17].

Recently, *Göbel* and co-workers showed that, for some template sequences, replacement of deoxynucleotides by ribonucleotides leads to more efficient complementary synthesis [5][6]. They suggested that efficient template-directed synthesis occurs in systems that, like RNA, adopt the nucleic-acid-A structure. Subsequently, work with nucleic acid analogues such as HNA (HNA is a DNA analogue built up from standard nucleobases and a phosphorylated 1,5-anhydrohexitol backbone [18]) and ANA (ANA is an RNA analogue built up from standard nucleobases and a phosphorylated D-altritol backbone [19]), two polynucleotide analogues that are more or less pre-organized to form double-helices with the A structure, has reinforced this conclusion [20–22]. It seemed interesting, therefore, to repeat some of the earlier experiments involving DNA templates and a chimeric RNA-DNA primer $(dG)_3G$, replacing deoxynucleotides by ribonucleotides both in the template and the primer (*Fig. 1,b* and *c*). Experiments of this kind are reported here.

[1]) Present address: The Scripps Research Institute, Department of Chemistry, BCC 338, 10550 North Torrey Pines Road, La Jolla, CA, 92037, USA.

[2]) Tel.: (858) 453-4100; Fax: (858) 558-7359; e-mail: orgel@salk.edu

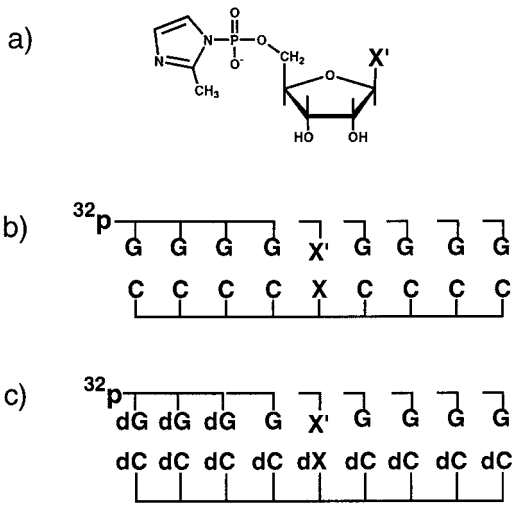

Fig. 1. a) *Structure of activated nucleoside 5′-phosphates (2-MeImpX), where X can be G, C, A or U.* b) *Schematic representation of the primer* $^{32}pG_4$ *extension reaction with 2-MeImpG and 2-MeImpX′* (X′ is C, U, or A) *on a* C_4XC_4 *template* (X = G, U, A, AA, or UU). c) *Schematic representation of the primer* $^{32}p(dG)_3G$ *extension reaction with 2-MeImpG and 2-MeImpX′* (X′ is C, U, or A) *on a* $d(C_4XC_4)$ *template* (X = G, T, or A).

Results. – In our first experiments, we compare template-directed extension of a G_4 primer on a C_4XC_4 template (ribo series, *Fig. 1,b*) with the corresponding extension of a chimeric $(dG)_3rG$ primer on a $d(C_4XC_4)$ template (deoxy series, *Fig. 1,c*), where X = G, U(T), or A. It is not practical to use a $(dG)_4$ primer since addition to the isolated OH of a 3′-terminal deoxynucleotide is intrinsically much slower than addition to the *cis*-glycol of a 3′-terminal ribonucleotide [13]. Consequently, comparisons of the rate of extension of a $(dG)_4$ primer with the rate of extension of a rG_4 primer would be dominated by the reactivity of the terminal nucleotide and would not give useful information about the effect of helix conformation on reactivity.

The nature of the products formed in experiments using DNA components (DNA series) is illustrated in lanes *2–4* of *Fig. 2*; corresponding results for experiments involving RNA oligomers (RNA series) are shown in lanes *6–8*. A comparison of the lane *2* of *Fig. 2* with lane *6* shows that, on templates with the sequence C_4GC_4 and $d(C_4GC_4)$, copying is effective in both the ribo and the deoxy series. However, short oligomeric products are less abundant in the ribo series than in the deoxy series, and the ratio of the yield of full-length 9-mer product to that of the 8-mer is substantially greater in the ribo series. Synthesis is efficient, therefore, in both series, but more efficient in the ribo series than in the deoxy series.

A comparison of lane *3* and lane *7* establishes that the overall efficiency of synthesis on C_4UC_4 and $d(C_4TC_4)$ templates is again greater in the ribo series than in the deoxy-series. However, the detailed results are somewhat surprising. The band corresponding to the 8-mers in the deoxy series is accompanied by a 'shadow' (lane *3*), but in the ribo series there are two bands of comparable intensity (lane *7*). In independent experiments, we have confirmed that the slower moving and somewhat more intense component of the double band is G_8 rather than the anticipated G_4AG_3. In a similar

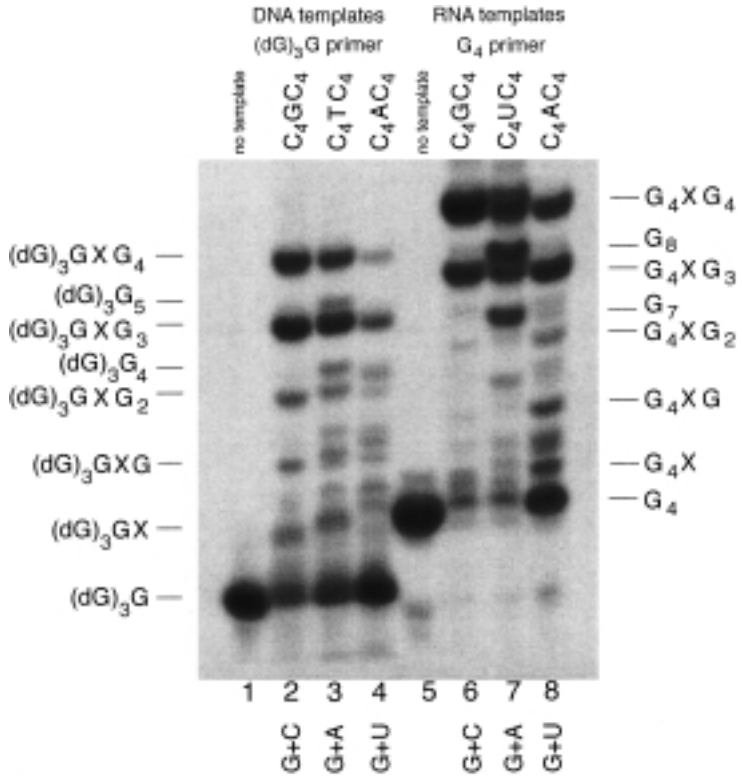

Fig. 2. *Extension of a $^{32}p(dG)_3G$ primer on DNA templates $d(C_4XC_4)$* (X = G, T, or A) *and of a $^{32}pG_4$ primer on RNA templates C_4XC_4* (X = G, U, or A) *after 5 days.* G represents 2-MeImpG, G + C represents an equimolar mixture of 2-MeImpG and 2-MeImpC, G + A represents an equimolar mixture of 2-MeImpG and 2-MeImpA, G + U represents an equimolar mixture of 2-MeImpG and 2-MeImpU. The fastest-moving band in the diagram corresponds to the primer.

way, the single band corresponding to a 7-mer in the ribo-series (lane *7*) has a mobility corresponding to G_7 rather than G_4AG_2; a much weaker band is present at this position in lane *3*. The results imply that wobble-pairing, which leads to the incorporation of G in the product opposite U or T in the template, is much more extensive in the ribo series than in the deoxy series.

A comparison of the products of synthesis on $d(C_4AC_4)$ (lane *4*) with the results on C_4AC_4 (lane *8*) shows that very little full-length product is formed in the deoxy series (lane *4*), while the yield is substantially larger in the ribo series (lane *8*). However, even in the ribo series the absolute product yield is much smaller when the heterobase in the template is A than when the heterobase is G or U. This confirms that the ribo system always leads to more efficient synthesis than the deoxy system. In addition, it shows that the difference between the two systems is greatest for the least efficient reaction, that is the incorporation of U opposite A.

To distinguish the effect of the template on the efficiency and fidelity of copying from that of the primer, we carried out experiments with all four combinations of RNA or DNA in the primer and template. The results in *Fig. 3* show that the nature of the

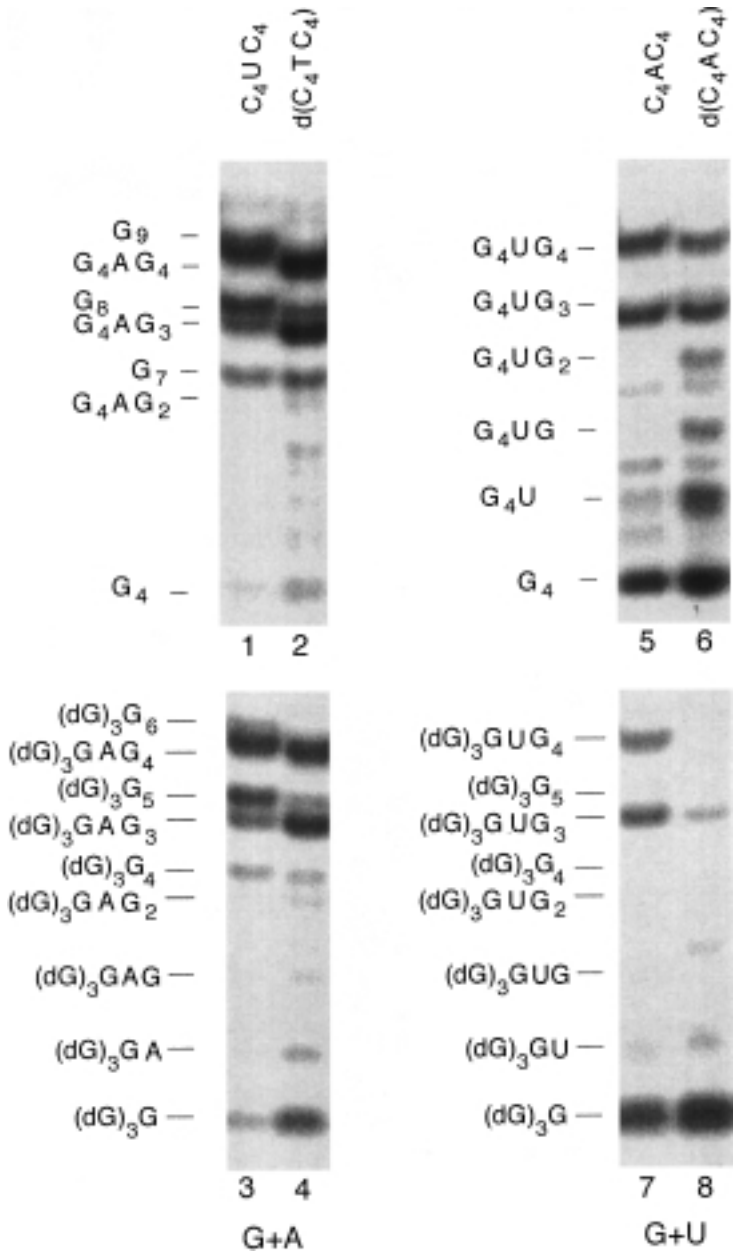

Fig. 3. *Extension of a $^{32}pG_4$ primer* (lanes *1, 2, 5, 6*) *or a $^{32}p(dG)_3G$ primer* (lanes *3, 4, 7, 8*) *on a C_4UC_4 RNA template* (lanes *1, 3*), *a d(C_4TC_4) DNA template* (lane *2, 4*), *a C_4AC_4 RNA template* (lanes *5, 7*), *or a d*(C_4AC_4) *DNA template* (lane *6, 8*) *after 5 days in the presence of an equimolar mixture of 2-MeImpG and 2-MeImpA* (lanes *1–4*), *or an equimolar mixture of 2-MeImpG and 2-MeImpU* (lanes *5–8*). The fastest-moving bands in the diagrams correspond to the primers.

template is dominant in determining the fidelity when U or T is the heterobase. Comparison of lane *1* with lane *3*, or lane *2* with lane *4* shows that the nature of the primer is unimportant, while comparison of lane *1* with lane *2*, or lane *3* with lane *4* demonstrates the importance of the template. Wobble-pairing is controlled by the template.

The situation is more complex when the heterobase is A. Comparison of lanes *6* and *8* shows that the nature of the primer is crucial in determining the efficiency on a DNA template, but comparison of lanes *5* and *7* demonstrate a lesser degree of dependence in the primer when an RNA template is present. However, the nature of the template is also important as seen by comparing lanes *5* and *6* with lanes *7* and *8*. In this case, it is not easy to partition the difference between the RNA and DNA systems between template and primer.

The results described above suggest that C and U are incorporated with high fidelity opposite G and A, respectively, but that mispairing of G with U in the templates is extensive. This is confirmed by the results in *Fig. 4,a*. Comparison of lanes *6* and *7* shows that little G is incorporated opposite G on a C_4GC_4 template even when C is absent, and that the small minicorporation that does occur is suppressed in the presence of C. Similarly, the incorporation of G opposite A is small (lane *4*) and is suppressed by the

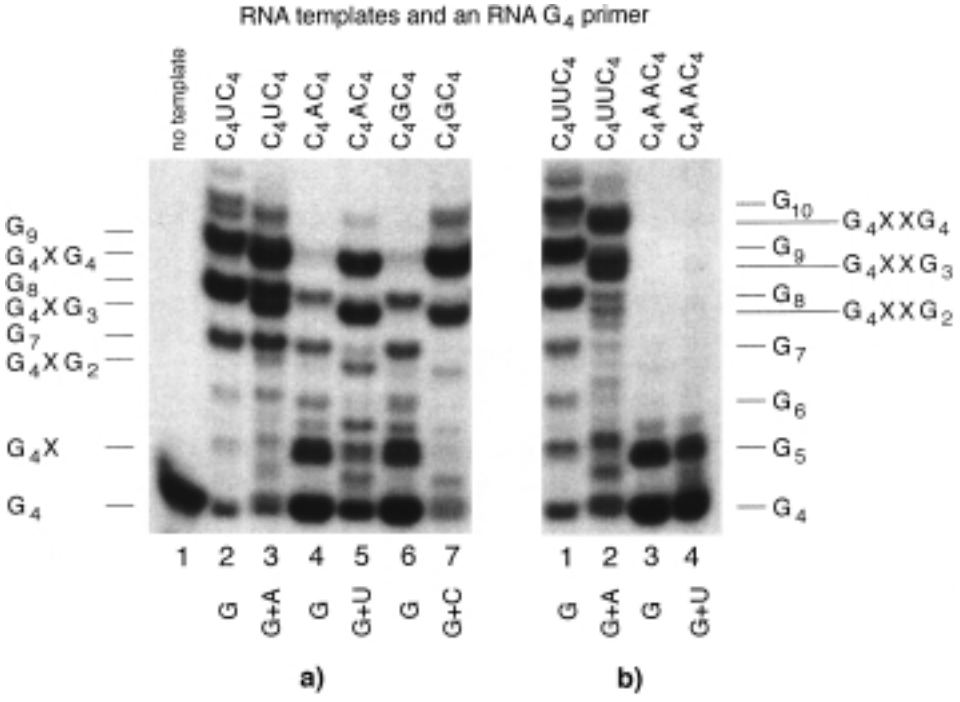

Fig. 4. a) Extension of a $^{32}pG_4$ primer on RNA templates C_4XC_4 (X = G, U or A) after 5 days; G represents 2-MeImpG, G + C represents an equimolar mixture of 2-MeImpG and 2-MeImpC, etc. b) Extension of a $^{32}pG_4$ primer on RNA templates C_4XXC_4 (X = U or A) primer after 5 days; G represents 2-MeImpG, G + U represents an equimolar mixture of 2-MeImpG and 2-MeImpU, etc. The fastest-moving band in the diagrams corresponds to the primer.

presence of U (lane *5*). The situation is completely different when U is the heterobase in the template. Then, G is incorporated very efficiently in the absence of A and, as we have already seen, competes effectively with A when both are present.

We have shown above that synthesis using RNA templates and primers is substantially more efficient than synthesis using DNA templates or deoxynucleotide-containing primers, but that wobble-pairing is more extensive on RNA templates. We decided, therefore, to explore these two aspects of template-directed synthesis further using two additional RNA template sequences, C_4AAC_4 and C_4UUC_4. The results illustrated in *Fig. 4,b*, show clearly that a pair of adjacent A residues almost completely blocks primer extension in the ribo-system just as it does on deoxy-templates [12][16]. No oligomers longer than the pentamer are formed on a C_4AAC_4 template whether G alone or G + U are used as substrates (lanes *3* and *4*). The results illustrated in *Fig. 4,b*, lane *1*, show that synthesis proceeds efficiently past two consecutive U residues on a C_4UUC_4 template even when A is omitted from the substrate mixture, emphasizing again the importance of wobble-pairing on RNA templates.

Discussion. – Our experiments show that complementary copying of RNA sequences in our experimental system is, in general, more efficient than copying of DNA sequences. As first suggested by *Göbel* and co-workers and as later confirmed using a variety of nucleic-acid analogues, the most efficient copying occurs in double-helical systems that adopt the A structure. Since both RNA and DNA/RNA hybrids usually form A structures, our results also imply that the most efficient copying requires that both template and product should have backbones that favor the formation of the A structure.

The most surprising difference is the degree to which RNA supports copying by wobble-pairing. Incorporation of G opposite U is appreciable when DNA is used as template, but this form of mispairing can compete effectively with normal pairing only on RNA. If wobble-pairing is a general feature of non-enzymatic synthesis of RNA and not a special problem associated with the use of phosphoroimidazolides as substrates, it is hard to see how a genetic system involving four bases could have evolved *de novo*. Instead, one would have to consider an initial system involving only one purine and one pyrimidine base, or postulate the existence in a pre-RNA world of enzyme-like catalysts that suppressed wobble-pairing. It is important to determine whether the difference in the extent of wobble-pairing on DNA and RNA is entirely due to differences in the backbones, or if the Me group on T helps to suppress wobble-pairing.

Since synthesis on RNA is more efficient than on DNA, our previous quantitative estimates of the rate of copying of RNA, which were based on deoxynucleotide-containing oligomers, are too low. However, our qualitative conclusions are unchanged. RNA Sequences involving A residues are copied less efficiently than other templates, and consecutive A residues are found to be a complete barrier to copying RNA just as they are for copying DNA. RNA Sequences very rich in C and G form secondary structures similar to those formed by DNA sequences, so G,C-rich sequences will not act as efficient templates [23]. Thus, although copying may be somewhat more efficient on RNA than on DNA templates, exponential replication, that is repeated rounds of copying, using nucleoside 5′-(2-methylimidazol-1-yl phosphates) as substrates is still not possible.

We thank *Aubrey R. Hill, Jr.* for technical assistance and *Bernice Walker* for manuscript preparation. This work was supported by grants from *NASA*, NAG5-8512 and NAG5-4546 NSCORT/Exobiology.

Experimental Part

Unless otherwise noted, all chemicals were reagent grade, were purchased from commercial sources, and were used without further purification. Nucleoside 5′-(2-methylimidazol-1-yl phosphates) (2-MeImpX, X = G, C, A, U) were obtained according to a published method [24] in at least 95% yield. Oligodeoxyribonucleotides were synthesized and purified as described in [25]. Oligoribonucleotides were purchased from *Dharmacon Research, Inc.*, and deprotected according to the manufacturer's protocols.

Reaction conditions for [^{32}P]-labeled p$(dG)_3$G and pG_4 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 [15][22]. One set of reactions was incubated for 5 days at 0° in 0.2M 2,6-lutidine buffer (pH 7.9 at 25°) containing 1.2M NaCl, 0.2M $MgCl_2$, 20 μM of a template, 20 nM of the primer, and 50 mM 2-MeImpG. In another set of reactions, the soln. contained 50 mM 2-MeImpG and 50 mM of the activated nucleotide complementary to X in the template (C_4XC_4 or C_4XXC_4). The primers $(dG)_3$G and G_4 were [^{32}P]-labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase [26]. The reaction mixtures were analyzed by electrophoresis on 20% polyacrylamide gels containing 8M urea, as previously described [15][22].

REFERENCES

[1] G. F. Joyce, ' Cold Spring Harbor Symposia on Quantitative Biology', Vol. LII, Cold Spring Harbor Press, Cold Spring Harbor, New York, 1987, p. 41.
[2] A. Eschenmoser, *Origins Life* **1994**, *24*, 389.
[3] A. Eschenmoser, *Origins Life* **1997**, *27*, 535.
[4] T. Inoue, L. E. Orgel, *Science* **1983**, *219*, 859.
[5] M. Kurz, K. Göbel, C. Hartel, M. W. Göbel, *Angew. Chem., Int. Ed.* **1997**, *36*, 842.
[6] M. Kurz, K. Göbel, C. Hartel, M. W. Göbel, *Helv. Chim. Acta* **1998**, *81*, 1156.
[7] A. Kanavarioti, D. H. White, *Orig. Life Evol. Biosph.* **1987**, *17*, 333.
[8] A. Kanavarioti, C. F. Bernasconi, *J. Mol. Evol.* **1990**, *31*, 470.
[9] A. Kanavarioti, M. T. Rosenbach, T. B. Hurley, *Orig. Life Evol. Biosph.* **1991**, *21*, 199.
[10] A. Kanavarioti, *J. Org. Chem.* **1998**, *63*, 6830.
[11] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 5496.
[12] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 7964.
[13] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 317.
[14] C.-H. B. Chen, T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1985**, *181*, 271.
[15] J. G. Schmidt, P. E. Nielsen, L. E. Orgel, *Nucleic Acids Res.* **1997**, *25*, 4797.
[16] A. R. Hill Jr., T. Wu, L. E. Orgel, *Orig. Life Evol. Biosphere* **1993**, *23*, 285.
[17] I. A. Kozlov, P. K. Politis, S. Pitsch, P. Herdewijn, L. E. Orgel, *J. Am. Chem. Soc.* **1999**, *121*, 1108.
[18] C. Hendrix, H. Rosemeyer, I. Verheggen, F. Seela, A. Van Aerschot, P. Herdewijn, *Chem. Eur. J.* **1997**, *3*, 110.
[19] B. Allart, K. Khan, H. Rosemeyer, G. Schepers, C. Hendrix, K. Rothenbacher, F. Seela, A. van Aerschot, P. Herdewijn, *Chemistry. Eur. J.* **1999**, *5*, 2424.
[20] I. A. Kozlov, M. Zielinski, B. Allart, L. Kerremans, A. Van Aerschot, R. Busson, P. Herdewijn, L. E. Orgel, *Chem. Eur. J.* **2000**, *6*, 151.
[21] I. A. Kozlov, P. K. Politis, A. Van Aerschot, R. Busson, P. Herdewijn, L. E. Orgel, *J. Am. Chem. Soc.* **1999**, *121*, 2033.
[22] I. A. Kozlov, B. De Bouvere, A. Van Aerschot, P. Herdewijn, L. E. Orgel, *J. Am. Chem. Soc.* **1999**, *121*, 5856.
[23] G. F. Joyce, L. E. Orgel, *J. Mol. Biol.* **1986**, *188*, 433.
[24] G. F. Joyce, T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1984**, *176*, 279.
[25] I. A. Kozlov, S. Pitsch, L. E. Orgel, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13448.
[26] T. Maniatis, E. F. Fritsch, J. Sambrook, 'Molecular Cloning: A Laboratory Manual', Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1982.

Received March 27, 2000