Oligomerization of activated D- and L-guanosine mononucleotides on templates containing D- and L-deoxycytidylate residues

(self-replicating systemy**enantiomeric cross-inhibition**y**L-nucleotides**y**alternating D-L-oligonucleotides)**

IGOR A. KOZLOV*, STEFAN PITSCH†, AND LESLIE E. ORGEL*‡

*The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, CA 92186; and †Swiss Federal Institute of Technology, Universitätstrasse 16, CH-8092, Zurich, Switzerland

Contributed by Leslie E. Orgel, September 22, 1998

ABSTRACT The oligomerization of activated D- and Land racemic guanosine-5***-phosphoro-2-methylimidazole on short templates containing D- and L-deoxycytidylate has been studied. Results obtained with D-oligo(dC)s as templates are similar to those previously reported for experiments with a poly(C) template. When one L-dC or two consecutive L-dCs are introduced into a D-template, regiospecific synthesis of 3*****-5*** **oligo(G)s proceeds to the end of the template, but three consecutive L-dCs block synthesis. Alternating D-,L-oligomers do not facilitate oligomerization of the D-, L-, and racemic 2-guanosine-5*****-phosphoro-2-methylimidazole. We suggest that once a ''predominately D-metabolism'' existed, occasional L-residues in a template would not have led to the termination of self-replication.**

All living organisms use the D-enantiomers of ribo- and deoxynucleotides as building blocks for their nucleic acids. Any abiotic synthesis of a nucleotide would yield equal amounts of the D- and L- isomers. It is not clear why only D-nucleotides were chosen during evolution (1–4). In the present study, we explore the way in which L-nucleotides in a prebiotic soup would effect a self-replicating system of D-nucleotides early in the development of nucleic acid replication. As a model system, we used the highly efficient template-directed oligomerization of activated guanosine mononucleotides on oligo(dC) templates (5). We report here on the oligomerization of activated D- and L-guanosine mononucleotides on templates containing D- and L-deoxycytidylate residues (Fig. 1).

MATERIALS AND METHODS

Unless otherwise noted, all chemicals were reagent grade, were purchased from commercial sources, and were used without further purification. D-Guanosine 5'-monophosphate was obtained from Sigma. L-Guanosine 5'-monophosphate was synthesized from N^3 -isobutyryl-9-(2',3'-di-*O*-benzoyl- β -Lribofuranosyl)guanosine (6) by phosphitylation with bis(2 cyanoethyl)-*N,N*-diisopropyl phosphoramidite, followed by oxidation and aminolysis (85% yield). The enantiomers of guanosine 5'-phosphoro-2-methylimidazolide (2-MeImpG) were synthesized by a published procedure (7) with yields of 95%. b-L-Deoxycytidine(N-bz)b-cyanoethyl *N*,*N*-diisopropyl phosphoramidite was purchased from Chem-Genes (Waltham, MA). Oligodeoxyribonucleotide templates were synthesized on a 391A DNA synthesizer (Applied Biosystems), were deprotected in concentrated ammonia at 55°C, were purified by 20% PAGE, and were desalted on a Nensorb column (DuPont/NEN). The $(dG)_{3}G$ primer is an oligodeoxynucleotide terminated by a single ribonucleotide residue at the 3'

terminus. It was obtained as described above but by using a ribo controlled pore glass G^{Ac} column (Glen Research, Sterling, VA) to introduce the $3'$ -ribonucleotide. Nuclease P1 was obtained from Pharmacia. The primer was labeled with $[\gamma^{32}P]$ -ATP and T4 polynucleotide kinase (New England Biolabs) as described in ref. 8 and was purified on a Nensorb column. To determine the position of $Gpp(dG)_{3}G$ in PAGE it was synthesized as described in ref. 9.

All PAGE separations were run on a denaturing (8 M urea) 20% polyacrylamide gel. The elution buffer was 50 mM Tris borate (pH 8.3) containing 1 mM EDTA. Loading buffer was prepared by mixing 900 μ I of deionized formamide, 25 μ I of xylene cyanol (2%) , 25 μ l of bromophenol blue (2%) , and 50 μ l of 10 \times Tris borate, EDTA buffer.

Reaction conditions for the polymerization of 2-MeImpG on various templates were chosen to permit comparison with earlier published work (5, 10, 11). The reactions were run for 7 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 MgCl₂, 0.5 mM template, and 0.1 M 2-MeImpG. When a mixture of D- and L- isomers was used, the concentration of each isomer was 0.1 M. The volume of the reaction mixture was 3μ . To prepare the reaction mixture, the appropriate amounts of stock solutions of NaCl, MgCl₂, and a template were mixed in an Eppendorf tube and were evaporated to dryness. The residue was redissolved in 3 μ l of fresh, prepared solution of 2-MeImpG in 0.2 M 2,6lutidine buffer to start the reaction. HPLC analyses of the reaction mixtures were performed on an RPC5 column as described (7). Reaction products were eluted with a linear gradient of NaClO₄ (pH 12, 0-0.06 M in 60 min) and were monitored by UV absorption at 254 nm.

Reaction conditions for $p(dG)_{3}G$ primer extension with 2-MeImpG on different templates were chosen again to permit comparison with earlier published work (12, 13). The reactions were run 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 MgCl₂, 20 μ M template, 20 nM primer, and 50 mM 2-MeImpG. When a mixture of Dand L-isomers was used, concentration of each isomer was 25 mM. The volume of the reaction mixture was 6μ l. To prepare the reaction mixture, appropriate amounts of stock solutions of NaCl, MgCl₂, primer, and a template were mixed in an Eppendorf tube and were evaporated to dryness. The residue was redissolved in 3 μ l of 0.2 M 2,6-lutidine buffer. The solution was heated to 95°C and was cooled to 25°C for 1 hour, then kept at 0°C for 20 min. Then, fresh prepared 2-MeImpG solution in 0.2 M 2,6-lutidine buffer $(3 \mu l)$ was added to start the reaction. Reaction mixtures were analyzed by 20% PAGE.

RESULTS

Oligomerization of the Enantiomers of 2-MeImpG on a $d(C_{10})$ **Template.** In the absence of a template, oligomerization

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked ''*advertisement*'' in accordance with 18 U.S.C. §1734 solely to indicate this fact.

^{© 1998} by The National Academy of Sciences 0027-8424/98/9513448-5\$2.00/0 PNAS is available online at www.pnas.org.

Abbreviation: 2-MeImpG, guanosine-5'-phosphoro-2-methylimidazole.

 $\overline{*}$ To whom reprint requests should be addressed. e-mail: orgel@salk. edu.

FIG. 1. (*a*) Structures of D-2-MeImpG and L-2-MeImpG. (*b*) Schematic representation of the template-directed oligomerization of 2-MeImpG on D -L mixed oligodeoxynucleotide template $d(CC*)_5dC$, where dC* represents L-deoxycytidine. (*c*) Schematic representation of the template-directed oligomerization of 2-MeImpG and primer (dG) ₃G extension reaction with 2-MeImpG on D-L mixed oligodeoxynucleotide templates that have structure $d(C_4XC_4)$ where X can be dC*; d(C*)2, d(C*)3, d(C*CC*), or d(C*CC*CC*). (*d*) Schematic representation of the primer (dG)3G extension reaction with 2-Me-ImpG on D-L mixed oligodeoxynucleotide templates that have structure d(C₄XC₄), where X can be dC^{*}, (dC^{*})₂, (dC^{*})₃, d(C^{*}CC^{*}), or d(C*CC*CC*). In *b*–*d*, C is D-dC, C* is L-dC, and rG is D-2-MeImpG or L-2-MeImpG.

of the D- and L- enantiomers of 2-MeImpG yields identical patterns of products; only dimers and smaller amounts of trimers are formed (data are not shown). The results that we obtained by using a $d(C_{10})$ template and enantiomers of 2-MeImpG as substrates (Fig. 2 *a*, *b*, and *c*) are analogous to those reported for a poly(C) template (10). The presence of the $d(C_{10})$ template leads to efficient oligomerization of D-2-MeImpG (Fig. 2*a*). The major peaks on the HPLC profile correspond to all $3'-5'$ -linked oligo(G)_ns ranging in length from the dimer to 9-mer (5, 14–16). The addition of the 10th nucleotide to the 9-mer, as anticipated, leads to formation of a mixture of $2'$ -5'-linked and $3'$ -5'-linked products in low yield

(5). In striking contrast to these results, a $d(C_{10})$ template has no effect on the oligomerization of L-2-MeImpG (Fig. 2*b*).

The HPLC profile in Fig. 2*c* shows the product distribution obtained in the template-directed polymerization of an equimolar mixture of D- and L-2-MeImpG. The single peaks of the 3'-5'-linked products of the template-directed reaction of D-2-MeImpG are replaced by a shorter series of complex groups of peaks. A similar effect has been reported by Joyce *et al.* (10) for oligomerization of D-2-MeImpG on a poly(C) template and was designated as enantiomeric cross-inhibition. It was shown that, on a poly(C) template, L-2-MeImpG attaches to the $2'(3')$ terminus of D-oligo(G)_n by means of either a $2'$ -5' or a $3'$ -5' internucleotide bond and also to the 5' terminus of D -oligo $(G)_n$ by means of a pyrophosphate bond. The presence of these adducts, together with those formed with the D-isomer, explains the complexity of the elution profile.

Oligomerization of D-2-MeImpG on Templates Containing Dand L-dC. We studied the oligomerization of D-2-MeImpG on "mixed" D-L-oligo(dC) templates, which have the sequence $d(C_4XC_4)$, where X can be dC^* , $d(C^*)_2$, $d(C^*)_3$, $d(C^*CC^*)$, or d(C*CC*CC*) and dC* represents L-dC (Fig. 1*c*). The results obtained are shown in Fig. 3. The presence of the $d(C_8)$ template leads to efficient oligomerization of D-2-MeImpG (Fig. 3*a*). The major peaks on the HPLC profile correspond to all 3'-5'-linked $oligo(G)_n$ ranging in length from the dimer to the 7-mer. The insertion of one or more L-dC residues into the $d(C_8)$ template leads to a significant decrease in the oligomerization efficiency (Fig. 3 *b*–*f*). The largest peaks on the HPLC profiles (Fig. 3 *b*–*f*) in every case correspond to all $3'$ -5'-linked oligo(G)_n ranging in length from the dimer to the 5-mer. The presence of one or two L-dC residues in the middle of a $d(C_8)$ template does not stop extension of oligo(G)_ns completely (Fig. 3 *b* and *c*). All oligo(G)_ns from the dimer to the 8-mer or the 9-mer are detected in the HPLC profiles, although the yields of oligomers are somewhat reduced compared with those obtained with a $d(C_8)$ template. These products were shown to be exclusively $3'-5'$ -linked by cochromatography with corresponding oligo(G)*n*s synthesized in a reaction on a $d(C_8)$ template. It also was shown that purified oligomers from G_3 to G_7 are hydrolyzed completely by nuclease P1, an enzyme that is specific for $3'-5'$ linkages in RNA (17). The oligomerization of D-2-MeImpG yields very little product larger than G_5 when templates $d(C_4C^*C^*C_4^*)$, $d(C_4C^*CC^*C_4)$, and d(C4C*CC*CC*C4) are used (Fig. 3 *d*–*f*). Control experiments showed that a $d(C_4)$ (Fig. 3g,) or a $d(C_4AC_4)$ (Fig. 3*h*) template does not facilitate efficient oligomerization of D-2-MeImpG.

Extension of p(dG)3G Primer with D-2-MeImpG or L-2- MeImpG on D-L Mixed Oligo(dC) Templates. The products distribution in reaction of ^{32}P -labeled p(dG)₃G with D-2-MeImpG on mixed D-L oligo(dC) templates (Fig. 1*d*) are shown in Fig. 4. The results obtained confirm the results for the oligomerization of D-2-MeImpG on these templates. A $d(C_8)$ template directs extension of the primer up to $p(dG)_{3}G_{5}$. The presence of one or two L-dC residues does not stop the

FIG. 2. Oligomerization of D-2-MeImpG and L-2-MeImpG on d(C10) template. (*a*) D-2-MeImpG. (*b*) L-2-MeImpG. (*c*) A mixture of D-2-MeImpG and L-2-MeImpG. The numbers of the peaks represent the length of the oligo(G)_n products; T represents the position of the template.

FIG. 3. $(a-f)$ Oligomerization of D-2-MeImpG on D-L mixed oligo(dC) templates. (*a*) d(C₈). (*b*) d(C₄C*C₄). (*c*) d(C₄C*C^{*}C₄). (*d*) $d(C_4C^*C^*C^*C_4)$. (*e*) $d(C_4C^*CC^*C_4)$. (*f*) $d(C_4C^*CC^*CC^*C_4)$. (*g* and *h*) Control templates. (*g*) $d(C_4)$. (*h*) $d(C_4AC_4)$. The numbers of the peaks represent the length of the oligo(G)*n* products; T represents position of the template. The positions of the template on diagrams *c*–*f* are not shown.

extension of the primer whereas the presence of three L-dC residues prevents the attachment of more then one nucleotide. The presence of alternating sequences $d(C^*CC^*)$ or $d(C^*CC^*CC^*)$ in the middle of template allows the first nucleotide to be attached with high efficiency, and then only small amounts of second and third nucleotides.

The product distribution in reaction of a $p(dG)_{3}G$ primer with L-2-MeImpG on different mixed D-L templates is shown in Fig. 5. The $d(C_8)$ template directs attachment of one L nucleotide to the $5'$ terminus of the primer through a pyrophosphate bond to give $G^*pp(dG)_3\tilde{G}$, where G^* is L-G. A single L-nucleotide also adds to the $2'(3')$ terminus of the primer $p(dG)_{3}G$, but less efficiently, to produce a small amount of $p(dG)_{3}GG^*$. Templates containing one or more L-dC residue behave differently. Attachment of one Lnucleotide to the $2'(3')$ terminus of the primer is the main reaction, and the formation of $G^*pp(dG)_3G$ is somewhat less efficient. These differences are attributable to differences in the ways that primer and substrate can be accommodated on the various templates (see *Discussion*).

The extension of a $p(dG)_{3}G$ primer with a mixture of D- and L-2-MeImpG on the mixed oligo(dC) templates gives a complicated mixture of products. A $d(C_8)$ template directs the

formation of a mixture of $p(dG)_{3}G_{n}$ and $Gpp(dG)_{3}G_{n}$ products containing both D- and L-G. Templates containing one or more L-dC residues allow attachment of one mononucleotide to the $2'(3')$ terminus of the primer and also direct attachment of one mononucleotide to the $5'$ terminus of the primer through a pyrophosphate bond (data not shown).

Oligomerization of the Enantiomers of 2-MeImpG on an Alternating D-L Template. The presence of an alternating D-L $d(CC*)₅dC$ template in which dC^* represents L-dC has no effect on the oligomerization of D-2-MeImpG, L-2-MeImpG, or a racemic mixture of D- and L-2-MeImpG (data not shown).

DISCUSSION

The template-directed reactions of D-2-MeImpG on a variety of oligo(C) and oligo(dC) templates, and of racemic 2-Me-ImpG on poly(C), have been studied extensively $(5, 7, 10, 10)$ 14–16, 18, 19). Our results on the oligomerization of D-, L- and racemic 2-MeImpG on $d(C_{10})$ are consistent with previous findings. D-2-Me \overline{ImpG} yields $3'$ -5'-ligated oligomers up to the 9-mer rapidly and the 10-mer more slowly and less regiospecifically. L-2-MeImpG does not oligomerize significantly on $d(C_{10})$. Finally, racemic 2-MeImpG oligomerizes less effi-

FIG. 4. Extension of p(dG)₃G primer with D-2-MeImpG on D-L mixed oligo(dC) templates. Lanes: 1, no template; 2, $d(C_8)$; 3, d(C₄C*C₄); 4, d(C₄C*C*C₄); 5, d(C₄C*C*C*C₄); 6, d(C₄C*CC*C₄); and 7, $d(C_4C^*CC^*CC^*C_4)$. ${}^*p(dG)_3G$ is ${}^{32}p(dG)_3G$ (5'-32P-labeled primer).

ciently than D-2-MeImpG on $d(C_{10})$ and yields a complex mixture of products because of enantiomeric cross-inhibition (10).

Although L-2-MeImpG inhibits the synthesis of long D-G oligomers on $poly(C)$, it seemed possible that an oligo(dC) containing alternating D- and L- residues might act as an efficient template when incubated with racemic 2-MeImpG. Our experiments show that alternating the D-,L-template does not catalyze polymerization of D- or L-2-MeImpG or the racemate.

The oligomerization of D-2-MeImpG and the elongation of a $p(dG)_{3}G$ primer on templates of the form $d(C_{4}XC_{4})$, where C represents a D-dC residue and X represents a sequence

FIG. 5. Extension of p(dG)3G primer with L-2-MeImpG on D-L mixed oligo(dC) templates. Lanes: 1, no template; 2, $d(C_8)$; 3, d(C₄C*C₄); 4, d(C₄C*C^{*}C*C₄); 5, d(C₄C*C^{*}C*C₄); 6, d(C₄C*CC^{*}CC^{*}C₄); and 7, $d(C_4C^*CC^*CC^*C_4)$. G^* is L-G. $Sp(dG)_3G$ is ${}^{32}p(dG)_3G$ (5⁷- ${}^{32}P$ -labeled primer).

initiated by an L-dC residue, are unexpected. In every case, $3'$ -5'-linked-oligomers up to $G₅$ are obtained in substantial yield. When X is $L-dC$ or $(L-dC)_2$, synthesis continues to the end of the template strand, but when X is $(L-dC)₃$, synthesis stops at G_5 . In control experiments, we found that $d(C_4)$ and $d(C_4AC_4)$ do not support template-directed synthesis in the same way. They yield only G_3 and small amounts of G_4 . Many other published experiments indicate that a mismatch between template and substrate prevent efficient primer extension, except when wobble pairing between T and G is possible (18, $20 - 22$).

The results presented above indicate that D-2-MeImpG interacts specifically with an L-dC residue in the template in such a way as to permit elongation of an all D-primer. It seems almost certain that base pairing between L-dC and D-2- MeImpG occurs in the normal way with the formation of three hydrogen bonds. Then, because the syn-conformation of pyrimidine nucleotides is inaccessible, one possible way of bringing the 5'-phosphate of D-2-MeImpG into contact with the 3'-OH of the primer is by holding the D-2-MeImpG in the syn-configuration. It is also possible that D-G could form a Watson–Crick pair with L-dC in the manner suggested by Urata and coworkers (23).

Attempts to extend a $p(dG)_{3}G$ primer with L-2-MeImpG on the templates used in our experiments were unsuccessful. When the template contained one or more L-dC residues, small amounts of the primer extended at the 3'-end by one residue were the major products, together with lesser amounts of the 5'-capped pyrophosphate. On a $d(C_8)$ template, pyrophosphate formation was the main reaction, presumably because the primer and the L-2-MeImpG residue could be accommodated on a continuous sequence of D-residues of the template (Fig. 6). A similar pyrophosphate-capping reaction on poly(C) was reported by Joyce *et al.* and was discussed by them in detail (10). The attachment of an L-nucleotide to the 5' terminus of the D-primer through a pyrophosphate bond would not contribute to the enantiomeric cross-inhibition of primer extension at the 3' terminus.

It is striking, and possibly relevant to the origin of the RNA world, that primer elongation with D-2MeImpG continues, even if at a somewhat slower rate, past one or two L-dC residues in the template. If the RNA world was the first organized biological world, as often has been proposed, it must have arisen in an environment containing racemic nucleotides. Our findings do not overcome the problems presented by enantiomeric cross inhibition because none of our templates are copied efficiently when the monomeric substrate is racemic. However, our results do suggest that once a ''predominantly D-metabolism'' was in place, a small proportion of

FIG. 6. Schematic representation of the extension of a $p(dG)_{3}G$ primer by reaction with L-2-MeImpG (G^*) . (*a*) On a $d(C_8)$ template. Pyrophosphate formation is the main reaction, presumably because the primer and the L-2-MeImpG residue can be accommodated on a continuous sequence of D-residues of the template. (*b*) On D-L mixed oligo(dC) templates that have sequences $\overline{d(C_4XC_4)}$, where X is a sequence beginning with L-dC; once the primer occupies its most stable position on the template, there is no place to accommodate an additional L -2-MeImpG residue at the 3' terminus of the template.

L-monomers in the template or the substrate would not lead to the termination of replication. Nonenzymatic replication may be more resistant to poisoning by the incorrect enantiomer than previously seemed likely.

We thank Aubrey R. Hill, Jr. for technical assistance and Bernice Walker for manuscript preparation. This work was supported by National Aeronautics and Space Administration Grant NAG5-4118 and National Aeronautics and Space Administration Specialized Center of Research and Training/Exobiology Grant NAG5-4546.

- 1. Mason, S. F. (1984) *Nature (London)* **311,** 19–23.
- 2. Miller, S. L. & Orgel, L. E. (1974) *The Origins of Life on the Earth* (Prentice–Hall, Englewood Cliffs, NJ).
- 3. Joyce, G. F. (1989) *Nature (London)* **338,** 217–224.
- 4. Joyce, G. F. & Orgel, L. E. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 1–25.
- 5. Chen, C.-H. B., Inoue, T. & Orgel, L. E. (1985) *J. Mol. Biol.* **181,** 271–279.
- 6. Pitsch, S. (1997) *Helv. Chim. Acta* **80,** 2286–2312.
- 7. Joyce, G. F., Inoue, T. & Orgel, L. E. (1984) *J. Mol. Biol.* **176,** 279–306.
- 8. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
- 9. Chu, B. C. F. & Orgel, L. E. (1984) *Biochim. Biophys. Acta* **782,** 103–105.
- 10. Joyce, G. F., Visser, G. M., van Boeckel, C. A. A., van Boom, J. H., Orgel, L. E. & van Westrenen, J. (1984) *Nature (London)* **310,** 602–604.
- 11. Schmidt, J. G., Nielsen, P. E. & Orgel, L. E. (1997) *J. Am. Chem. Soc.* **119,** 1494–1495.
- 12. Bo¨hler, C., Nielsen, P. E. & Orgel, L. E. (1995) *Nature (London)* **376,** 578–581.
- 13. Schmidt, J. G., Nielsen, P. E. & Orgel, L. E. (1997) *Nucleic Acids Res.* **25,** 4797–4802.
- 14. Bridson, P. K. & Orgel, L. E. (1980) *J. Mol. Evol.* **144,** 567–577.
- 15. Inoue, T. & Orgel, L. E. (1981)*J. Am. Chem. Soc.* **103,** 7666–7667.
- 16. Inoue, T. & Orgel, L. E. (1982) *J. Mol. Biol.* **162,** 201–217.
- 17. Sawai, H., Totsuka, S., Yamamoto, K. & Ozaki, H. (1998) *Nucleic Acids Res.* **26,** 2995–3000.
- 18. Inoue, T. & Orgel, L. E. (1983) *Science* **219,** 859–862.
- 19. Inoue, T., Joyce, G. F., Grzeskowiak, K., Orgel, L. E., Brown, J. M. & Reese, C. B. (1984) *J. Mol. Biol.* **178,** 669–676.
- 20. Hill, A. R., Jr., Orgel, L. E. & Wu, T. (1993) *Origins Life Evol. Biosphere* **23,** 285–290.
- 21. Wu, T. & Orgel, L. E. (1992) *J. Am. Chem. Soc.* **114,** 5496–5501.
- 22. Wu, T. & Orgel, L. E. (1992) *J. Am. Chem. Soc.* **114,** 7963–7969.
- 23. Urata, H., Ueda, Y., Suhara, H., Nishioka, E. & Akagi, M. (1993) *J. Am. Chem. Soc.* **115,** 9852–9853.