Nonenzymatic Oligomerization Reactions on Templates Containing Inosinic Acid or Diaminopurine Nucleotide Residues

by **Igor A. Kozlov** and **Leslie E. Orgel***[1])

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

Dedicated to Prof. Dr. *Frank Seela* on the occasion of his 60th birthday

The template-directed oligomerization of nucleoside-5′-phosphoro-2-methyl imidazolides on standard oligonucleotide templates has been studied extensively. Here, we describe experiments with templates in which inosinic acid (I) is substituted for guanylic acid, or 2,6-diaminopurine nucleotide (D) for adenylic acid. We find that the substitution of I for G in a template is strongly inhibitory and prevents any incorporation of C into internal positions in the oligomeric products of the reaction. The substitution of D for A, on the contrary, leads to increased incorporation of U into the products. We found no evidence for the template-directed facilitation of oligomerization of A or I through A – I base pairing. The significance of these results for prebiotic chemistry is discussed.

Introduction. – Because the nucleic acids occupy a central position in biochemistry, studies of template-directed synthesis have concentrated on the oligomerization of activated derivatives of the standard nucleotides, U, A, C, and G [1 – 7], although a few experiments with 2,6-diaminopurine nucleotide [8][9] and other nucleotide analogues [10] have been reported. Oligomerization reactions involving inosinic acid (I) or 2,6-diaminopurine nucleotide (D), although they have limited application to biochemistry, are of considerable interest for prebiotic chemistry, since potentially prebiotic syntheses that yield adenine and guanine typically also lead to the formation of hypoxanthine and 2,6-diaminopurine [11]. Studies of template-directed chemistry might help to explain why A and G were chosen as components of RNA (DNA) while I and D were excluded. It has been suggested that a primitive genetic system might have been based on A – I pairing [12]. Template-directed reactions of activated A-derivative on I-containing templates and *vice versa* might throw light on this hypothesis. Here, we investigate template-directed synthesis in a variety of relevant systems involving I and D, and compare our results to those obtained using the standard bases.

Results. – *Oligomerization of Activated Mononucleotides on C_4XC_4* (X = G, I, A, or D) *DNA Templates* (*Fig. 1,b*). The presence of a C_4XC_4 (X = G, I, A, or D) DNA template leads to oligomerization of guanosine-5′-(2-methylimidazol-1-yl phosphate) (2-MeImpG) (*Fig. 1,a*; X′ = G) alone to give G_3 and small amounts of G_4 (data not shown). An equimolar mixture of 2-MeImpG and 2-MeImpC in the presence of a C_4GC_4 template leads to generation of G_4CG_n products up to at least octamers (*Fig. 2,b*) but, in the presence of a C_4IC_4 template, gives no products longer than tri-

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

and tetramers (*Fig. 2,a*). An equimolar mixture of 2-MeImpG and 2-MeImpU in the presence of a C_4AC_4 template leads to formation of G_4UG_n products up to nonamers (*Fig. 2,c*). This reaction becomes about three times more efficient when C_4DC_4 is substituted for C_4AC_4 (*Fig. 2,d*).

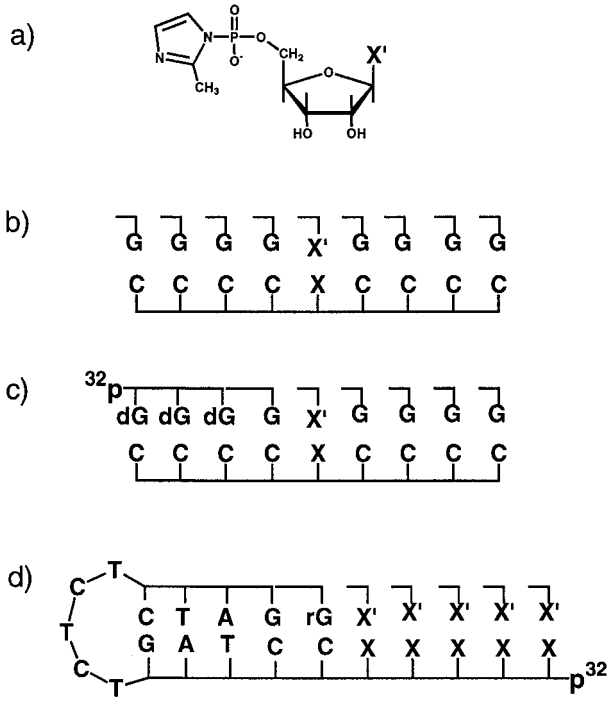

Fig. 1. a) *Structure of activated nucleoside 5′-phosphates* 2-MeImpX′ (X′ = G, C, A, I, or U). b) *Schematic representation of oligomerization of activated monomers 2-MeImpG and 2-MeImpX′* (X′ = C or U) *on a C_4XC_4 template* (X = G, I, A, or D). c) *Schematic representation of the primer $^{32}p(dG)_3G$ extension reaction with 2-MeImpG and 2-MeImpX′* (X′ = C or U) *on a C_4XC_4 template* (X = G, I, A, or D). d) *Schematic representation of the oligomerization of 2-MeImpX′* (X′ = C, U, A, or I) *on a deoxyribonucleotide hairpin template 5′-XXXXXCCTAGTCTCTCTAGrG-3′* (X = D, I, or A).

Extension of a $^{32}p(dG)_3G$ Primer on C_4XC_4 (X = G, I, A, or D) *DNA Templates* (*Fig. 2,c*). The product distributions in the reactions of ^{32}P-labeled $p(dG)_3G$ with 2-MeImpG, or an equimolar mixture of 2-MeImpG with 2-MeImpC or 2-MeImpU on C_4XC_4 (X = G, I, A, or D) DNA templates are shown in *Fig. 3*. Significant extension of the primer $p(dG)_3G$ with 2-MeImpG alone does not take place on any of the DNA templates (*Fig. 3*, *Lanes 1, 3, 5,* and *7*). The extension of the primer $p(dG)_3G$ with an equimolar mixture of 2-MeImpG and 2-MeImpC in the presence of a C_4GC_4 DNA template leads to conversion of more than 80% of the primer to $p(dG)_3GCG_n$ ($n = 1$–4) products (*Fig. 3*, *Lane 2*). However, the same reaction on a C_4IC_4 DNA template leads to conversion of only *ca.* 30% of the primer to a $p(dG)_3GC$ product, and does not yield any longer products (*Fig. 3*, *Lane 4*). The extension of the primer $p(dG)_3G$ with an equimolar mixture of 2-MeImpG and 2-MeImpU on a C_4AC_4 DNA template leads to conversion of less than 5% of the primer mainly to a $p(dG)_3GUG_3$ product (*Fig. 3*,

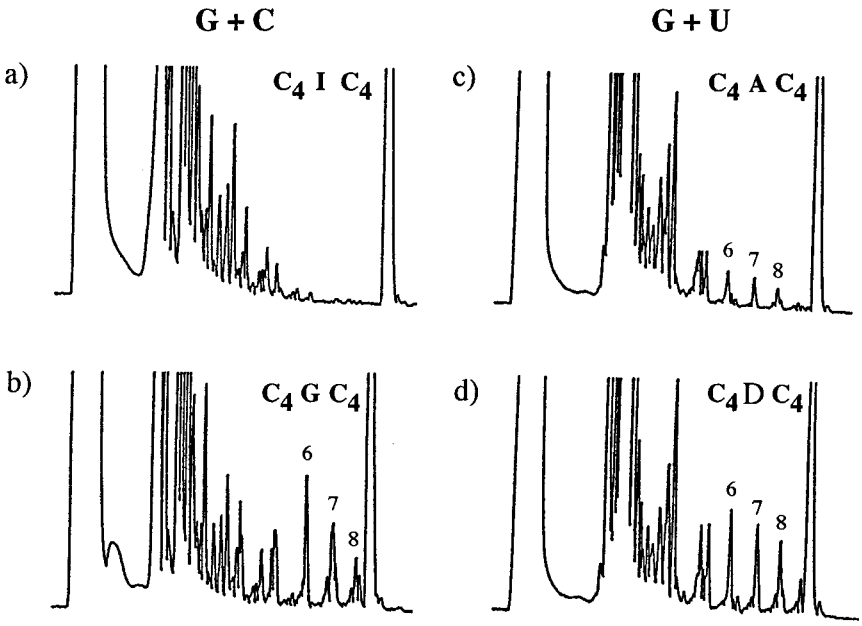

Fig. 2. *Elution profiles from an* RPC5 *column of the products from the oligomerization of an equimolar mixture of 2-MeImpG and 2-MeImpC* (G + C), *or an equimolar mixture of 2-MeImpG and 2-MeImpU* (G + U) a) *on a* C_4IC_4 *template*, b) *on a* C_4GC_4 *template*, c) *on a* C_4AC_4 *template*, d) *and on a* C_4DC_4 *template*. The reaction time was 14 days. The numbers above the peaks indicate the length of the all 3′-5′-linked oligoribonucleotide products.

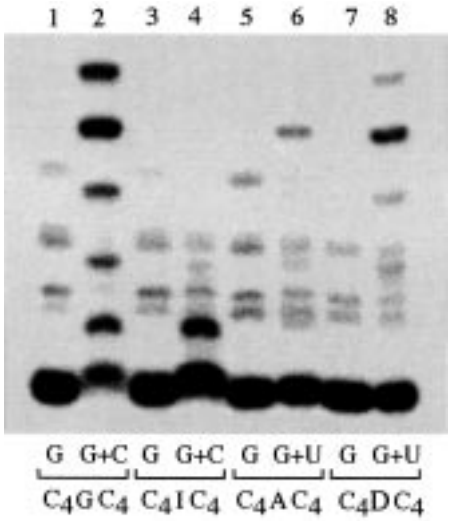

Fig. 3. *Extension of a* $^{32}p(dG)_3G$ *primer on* C_4XC_4 *DNA templates* (X = G, I, A, or D) *after 5 days*. G represents 2-MeImpG, G + C represents an equimolar mixture of 2-MeImpG and 2-MeImpC, G + U represents an equimolar mixture of 2-MeImpG and 2-MeImpU. The fastest-moving band in the diagram corresponds to the $^{32}p(dG)_3G$ primer.

Lane 6). The same reaction on a C_4DC_4 DNA template leads to conversion of more than 10% of the primer to $p(dG)_3GUG_3$ and $p(dG)_3GUG_4$ products in a ratio of *ca.* 10 : 1 (*Fig. 3*, *Lane 8*).

Oligomerization Reactions of 2-MeImpU and 2-MeImpC on Templates Containing Several D or I Residues. We studied oligomerization reactions of 2-MeImpU and 2-MeImpC on oligodeoxynucleotide D_{10} and I_{10} templates, respectively. We also studied 'primer extension' reactions in oligodeoxynucleotide hairpin templates, 5′-XXXXXCCTAGTCTCTCTAGrG-3′ (*Fig. 1*,*d*; X = D or I, X′ = U or C) that are 3′-terminated with a single ribonucleotide [6][7]. No template-directed facilitation of oligomerization was detected.

Oligomerization Reactions of 2-MeImpA and 2-MeImpI on Templates Containing Several I or A Residues. We studied in some detail the oligomerization reactions of 2-MeImpA and 2-MeImpI on oligodeoxynucleotide I_{10} and A_{10} templates, respectively. We also used oligodeoxynucleotide hairpin templates (*Fig. 1*,*d*, X = I or A, X′ = A or I) and oligodeoxynucleotide hairpin templates containing A – I base pairs (*Fig. 4*) to explore template-directed primer extension. In no case were we able to detect an effect of the template on oligomerization or primer extension.

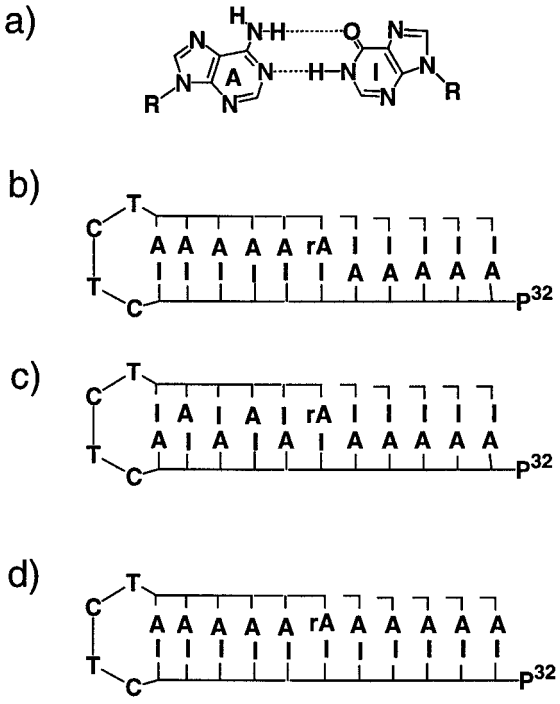

Fig. 4. a) *Schematic representation of an A – I base-pair.* Dotted lines represent H-bonds. b) – d) Schematic representation of attempted oligomerizations of 2-MeImpA or 2-MeImpI on deoxyribonucleotide hairpin templates containing A – I base pairs.

Discussion. – The results obtained in this study imply that replacement of the A – U base pair by the D – U base pair improves the efficiency of nonenzymatic template-directed oligomerization reactions. In some cases, D – U base pairs support reactions that are comparable in efficiency to those involving G – C base pairs (compare *Fig. 2,b* and *d*). The difference between A – U and D – U pairs is probably attributable to the presence of three H-bonds in the D – U pair compared with only two H-bonds in the A – U pair (*Fig. 5*). In this respect, the D – U pair resembles a G – C pair, both pairs having three H-bonds.

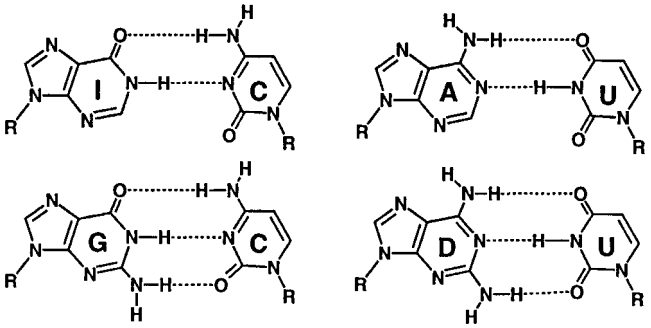

Fig. 5. *Schematic representation of base-pairings I – C, G – C, A – U, and D – U.* Dotted lines represent H-bonds.

The substitution of a G – C base pair by an I – C pair leads to a very large decrease in oligomerization efficiency. The I – C pair is much less efficient than the A – U pair in facilitating oligomerization (*cf. Fig. 2,a*, and *c*; *Fig. 3, Lanes 4* and *6*), although both are held together by two H-bonds. Our results show that the I – C base pair must adopt a conformation that inhibits template-directed synthesis strongly, especially at the stage when a primer is terminated by a C residue opposite I and needs to be extended by the addition of a G residue. We do not understand the structural basis of this inhibition. Our attempts to oligomerize 2-MeImpU and 2-MeImpC on D_{10} and I_{10} templates were unsuccessful, as were analogous attempts at primer-extension with oligodeoxyribonucleotide hairpin templates (*Fig. 1,d*, X = D or I, X′ = U or C).

Our results confirm that the G – C base pair is exceptional in providing the necessary conformation for efficient nonenzymatic RNA synthesis using 2-methyl imidazolides of nucleoside 5′-phosphates as substrates. Substitution of I for G in this reaction leads to poor incorporation of C and negligible extension of the resulting primer terminated by C. It is possible that a different activated derivative of C might polymerize efficiently, but our results suggest that the absence of I from replicating nucleic acids may have a basis in the conformation of double helices containing I – C base pairs.

A similar explanation to that given above cannot account for the exclusion of D from nucleic acids. Our results suggest that the replacement of A by D would lead to more efficient synthesis (*cf. Fig. 2,c* and *d*; *Fig 3*, *Lanes 6* and *8*), so the choice of A rather than D is likely to reflect factors other than efficiency of replication. Availability in the prebiotic environment is one possibility. Alternatively, optimization rather than maximization of the stability of double-helical RNA may have led to the selection of A rather than D.

Purine nucleosides can be synthesized more easily than pyrimidine nucleosides under prebiotic conditions [11]. Consequently, it has been suggested that purine–purine, A–I, pairing may have made possible the development of the first nucleic acid genetic system [12]. Our extensive efforts to demonstrate facilitation of adenosine nucleotide oligomerization on templates containing I, and *vice versa*, like several less complete earlier studies in our laboratory, have failed. We cannot exclude the possibility that experiments using a different form of activation would succeed, but our results do not provide any evidence supporting A–I pairing as a mechanism of complementary replication. They argue, although not conclusively, against the hypothesis of a genetic system based on A–I pairing.

Experimental Part

Unless otherwise noted, all chemicals were reagent grade, were purchased from commercial sources and used without further purification. Nucleoside 5′-(2-methylimidazol-1-yl phosphates) (2-MeImpX′, X′ = G, C, A, U, I) were synthesized by a published method in at least 95% yield [13]. The oligodeoxyribonucleotides were synthesized and purified as previously described [14]. 2-Amino-2-deoxyadenosine-β-cyanoethyl phosphoramidite and deoxyinosine-β-cyanoethyl phosphoramidite (*Glen Research*) were used under standard conditions to introduce D or I residues into oligodeoxyribonucleotides.

Reaction conditions for the oligomerization of 2-MeImpG (or its mixture with an equal amount of 2-MeImpC or 2-MeImpU) on DNA C_4XC_4 (X = G, I, A or D) templates were chosen to permit comparison with earlier published work [5][15]. Reactions were run at 0° for 14 days in 0.2M 2,6-lutidine-HCl buffer (pH 7.9 at 25°) containing 1.2M NaCl, 0.2M $MgCl_2$, and 0.5 mM of a template. In one set of reactions, the soln. also contained 0.1M 2-MeImpG. In another set, the soln. contained 0.1M 2-MeImpC and 2-MeImpG with the C_4GC_4 or C_4IC_4 template, or 0.1M 2-ImpU and 0.1M 2-MeImpG with the C_4AC_4 or C_4DC_4 templates. The same conditions were used for reactions on oligodeoxynucleotide I_{10}, D_{10}, and A_{10} templates; the concentration of activated substrate was always 0.1M. The reaction mixtures were analyzed by HPLC on an *RPC5* column as previously described [15].

Reaction conditions for $p(dG)_3G$ primer extension reactions on DNA C_4XC_4 (X = G, I, A, or D) templates were again chosen to permit comparison with earlier published work [5–7][15]. The reactions were 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, and 20 nM of the primer. In one set of reactions, the solns. also contained 50 mM 2-MeImpG. In a second set of reactions, the soln. contained not only 50 mM 2-MeImpG but also 50 mM of 2-MeImpC with C_4GC_4 and C_4IC_4 templates or 50 mM of 2-MeImpU with C_4AC_4 and C_4DC_4 templates. The same conditions were used in experiments with DNA hairpin templates (*Fig. 1,d*; X = A, D, I; *Fig. 4*). In these experiments, the concentration of template was 20 μM and the concentration of the corresponding activated monomer, 50 mM. The reaction mixtures were analyzed by electrophoresis in 20% PAG containing 8M urea as previously described [14].

We thank *Aubrey R. Hill, Jr.* for technical assistance and *Bernice Walker* for manuscript preparation. This work was supported by grants from *NASA* NAG5-4118 and NAG5-4546 NSCORT/Exobiology.

REFERENCES

[1] A. Kanavarioti, *J. Org. Chem.* **1998**, *63*, 6830.
[2] A. R. Hill Jr., L. E. Orgel, T. Wu, *Origins Life Evol. Biosphere* **1993**, *23*, 285.
[3] M. Kurz, K. Göbel, C. Hartel, M. W. Göbel, *Angew. Chem., Int. Ed.* **1997**, *36*, 842.
[4] M. Kurz, K. Göbel, C. Hartel, M. W. Göbel, *Helv. Chim. Acta* **1998**, *81*, 1156.
[5] J. G. Schmidt, P. E. Nielsen, L. E. Orgel, *Nucl. Acid Res.* **1997**, *25*, 4797.
[6] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 7964.
[7] T. Wu, L. E. Orgel, *J. Am. Chem. Soc.* **1992**, *114*, 5496.
[8] T. R. Webb, L. E. Orgel, *Nucl. Acids Res.* **1982**, *10*, 4413.
[9] K. Grzeskowiak, T. R. Webb, L. E. Orgel, *J. Mol. Evol.* **1984**, *21*, 81.
[10] H. Rembold, R. K. Robins, F. Seela, L. E. Orgel, *J. Mol. Evol.* **1994**, *38*, 211.

[11] S. L. Miller, L. E. Orgel, 'The Origins of Life on the Earth', Prentice-Hall, Inc., Englewood Cliffs, New Jersey 1974.
[12] F. H. C. Crick, *J. Mol. Biol*. **1968**, *38*, 367.
[13] G. F. Joyce, T. Inoue, L. E. Orgel, *J. Mol. Biol*. **1984**, *176*, 279.
[14] I. A. Kozlov, S. Pitsch, L. E. Orgel, *Proc. Natl. Acad. Sci. U.S.A*. **1998**, *95*, 13448.
[15] I. A. Kozlov, B. De Bouvere, A. Van Aerschot, P. Herdewijn, L. E. Orgel, *J. Am. Chem. Soc*. **1999**, *21*, 5856.

Received August 14, 1999