ducted following procedures described in full detail elsewhere.[12] Eppendorf tubes containing the end-labeled DNA[12] (9 μL) in TE buffer (10 mM Tris, 1 mM EDTA, pH = 7.6) were treated with the agent in a solution of DMSO (1 μL at the specified concentration). Agent concentrations were measured by UV using known extinction coefficients. The reaction was mixed by vortexing and brief centrifugation and subsequently incubated at 4 or 37 °C for the specified time (generally 24 h). The DNA was separated from unbound agent by ethanol precipitation and resuspended in TE buffer (20 μL). The solution in an Eppendorf tube sealed with Teflon tape was warmed to 100 °C for 30 min to induce cleavage at the alkylation sites, allowed to cool to room temperature, and centrifuged. Formamide dye was added (10 μL) to the supernatant. Prior to electrophoresis, the samples were warmed at 100 °C for 5 min, placed in an ice bath, and centrifuged and the supernatant was loaded onto the gel. Sanger dideoxynucleotide sequencing reactions were run as standards adjacent to the agent treated DNA. Gel electrophoresis was carried out using an 8% sequencing gel (19:1 acrylamide:*N,N*-methylene bisacrylamide, 8 M urea). Formamide dye contained xylene cyanol FF (0.03%), bromophenol blue (0.03%), and aqueous Na_2EDTA (8.7%, 250 mM). Electrophoresis running buffer (TBE) contained Tris base (100 mM), boric acid (100 mM), and $Na_2EDTA{\cdot}H_2O$ (0.2 mM). Gels were prerun for 30 min with formamide dye prior to loading the samples. Autoradiography of dried gels was carried out at −70 °C using Kodak X-Omat AR film and a Picker Spectra intensifying screen. The relative intensity of DNA alkylation at the high affinity alkylation site within clone w794 [5′-d(AATTA)-3′] versus time was determined by measuring the optical density from the autoradiogram at that site using an LKB UltroScan XL scanning laser densitometer interfaced to LKB 2400 Gelscan XL integration software.

Acknowledgment. This work was supported by the National Institutes of Health (Grant No. CA41986 to D.L.B.) and the American Cancer Society (Grant No. PF-3311 to S.A.M.). We gratefully acknowledge Professor M. Bina for the gift of M13mp10 clones.

Nonenzymatic Template-Directed Synthesis on Hairpin Oligonucleotides. 2. Templates Containing Cytidine and Guanosine Residues

Taifeng Wu and Leslie E. Orgel*

Contribution from The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, California 92186-5800. *Received December 6, 1991*

Abstract: We have prepared hairpin oligonucleotides in which a 5′-terminal single-stranded segment contains cytidylate (C) and guanylate (G) residues. When these hairpin substrates are incubated with a mixture of cytidine 5′-phosphoro(2-methyl)imidazolide (2-MeImpC) and guanosine 5′-phosphoro(2-methyl)imidazolide (2-MeImpG), the 5′-terminal segment acts as a template to facilitate sequence-specific addition of G and C residues to the 3′-terminus of the hairpin. If an isolated G residue is present at the 3′-end of the template strand, it is copied regiospecifically in the presence of 2-MeImpC and 2-MeImpG to give a product containing an isolated C residue linked to its G neighbors by 3′−5′-internucleotide bonds. However, if only 2-MeImpC is present in the reaction mixture, very little reaction occurs. Thus, the presence of 2-MeImpG catalyzes the incorporation of C. If the template strand contains a short sequence of G residues, it is copied in the presence of a mixture of 2-MeImpC and 2-MeImpG. If only 2-MeImpC is present in the reaction mixture, efficient synthesis occurs to give a final product containing one fewer C residue than the number of G residues in the template.

Introduction

Template-directed reactions on oligonucleotides, oligodeoxynucleotides, and polynucleotides that contain cytidylic acid (C) and guanylic acid (G) residues have been studied extensively.[1-12] It has been established that random copolymer templates that contain a substantial excess of riboC residues over riboG residues, when incubated with a mixture of cytidine 5′-phosphoro(2-methyl)imidazolide (2-MeImpC) and guanosine 5′-phosphoro-(2-methyl)imidazolide (2-MeImpG), are able to direct an efficient synthesis of copolymers rich in G residues.[7] However, copolymers containing more or less equal numbers of G and C residues are not effective templates, presumably because of their stable intramolecular self-structure. Poly(G) fails to act as a template for a different reason. It forms a very stable tetrahelical self-complex,[13] which prevents binding by C residues. On the basis of our studies with random copolymer templates we speculated that the failure to copy G-rich random poly(CG) templates was probably due to the intra- and intermolecular secondary structures of the templates rather than to the inability of the underlying condensation reaction of activated monomers to proceed on templates containing single-stranded G-rich sequences.[7]

Studies of reactions on short templates of defined sequence are consistent with this interpretation, but do not confirm it unambiguously. Isolated G residues in a ribo[4] or deoxyribo[6,8,10] template direct the efficient incorporation of C residues, but short oligomers containing the sequences −CCGGCC− and −CCCGCGCC− do not function as templates. We attributed this failure to the formation of double-helical regions stabilized by four strong G:C base pairs since, in our experiments, we were obliged to use a template concentration of at least 0.01 M, a concentration high enough to stabilize a CGCG or CGGC double-helical segment at 0 °C.[8]

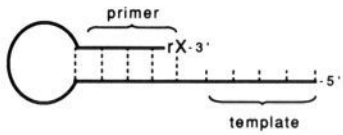
Recently we introduced a novel method for studying template-directed reactions, namely the extension of a hairpin oligonucleotide in which the 5′-end of the oligomer functions as a template for the extension of the 3′-end (Figure 1a).[14] A ribo-

(1) Inoue, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1981**, *103*, 7666-7667.
(2) Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1982**, *162*, 201-217.
(3) Inoue, T.; Orgel, L. E. *Science* **1983**, 219, 859-862.
(4) Inoue, T.; Joyce, G. F.; Grzeskowiak, K.; Orgel, L. E.; Brown, J. M.; Reese, C. B. *J. Mol. Biol.* **1984**, *178*, 669-676.
(5) Chen, C. B.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1985**, *181*, 271-279.
(6) Haertle, T.; Orgel, L. E. *J. Mol. Biol.* **1986**, *188*, 77-80.
(7) Joyce, G. F.; Orgel, L. E. *J. Mol. Biol.* **1986**, *188*, 433-441.
(8) Haertle, T.; Orgel, L. E. *J. Mol. Evol.* **1986**, *23*, 108-112.
(9) Grzeskowiak, K.; Orgel, L. E. *J. Mol. Evol.* **1986**, *23*, 287-289.
(10) Acevedo, O. L.; Orgel, L. E. *J. Mol. Biol.* **1987**, *197*, 187-193.
(11) Kanavarioti, A.; White, D. H. *Origins Life* **1987**, *17*, 333-349.
(12) Kanavarioti, A.; Bernasconi, C. F.; Doodokyan, L. D.; Alberas, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 7247-7257.

(13) *Basic Principles in Nucleic Acid Chemistry*; Ts'o, P. O. P., Ed.; Academic Press: New York and London, 1974; Vol. 1.

0002-7863/92/1514-5496$03.00/0

a.



b.

Template	Sequence
S_1	CCCCGCCTAGTTTTTCTAGrG
S_2	CCCCCGCTAGTTTTTCTAGrC
S_3	CCCGGCCTAGTTTTTCTAGrG
S_4	CCGGGCCTAGTTTTTCTAGrG
S_5	CCGGGGCCTAGTTTTTCTAGrG
S_6	CCGGGGGCCTAGTTTTTCTAGrG
S_7	CCCGGGGTGATTTTTTCACrC

Figure 1. (a) Hairpin structures of oligomers used in this work. The 3′-terminal residue is always a ribonucleotide. (b) Sequences of oligonucleotides S_1–S_7 printed in the 5′–3′ direction. The primer region is underlined and the template nucleotides are shown in bold type.

nucleotide is introduced at the 3′-end of the sequence to increase the efficiency of the reaction. We make use of denaturing gel electrophoresis to separate extension products from substrates. [^{32}P] labeling combined with autoradiography enables us to work with nanomolar concentrations of substrate and so to diminish greatly the consequences of intermolecular association between template molecules. The results obtained by applying this technique to oligomers containing C and G residues broadly confirm our earlier speculations.[7] In addition, they add a great deal of novel and detailed information, some of which is of general interest in the context of polynucleotide–mononucleotide interaction.

Results

Substrate with an Isolated G Residue in the Template Strand—Oligomer S_1. The results of an experiment involving the incubation of oligomer S_1 (Figure 1b) with 2-MeImpC, 2-MeImpG, or a mixture of 2-MeImpC and 2-MeImpG are illustrated in Figure 2a. After 1 day, only very small amounts of material corresponding to monoadducts are formed if 2-MeImpC or 2-MeImpG is present separately, but most of the starting material (72%) is converted to products when both activated nucleotides are present together. After 7 days small amounts of addition products containing 1–5 G residues are obtained (29%) when G only is present. The only products formed by incubation with 2-MeImpC are small amounts of monoadducts (20%). When both 2-MeImpC and 2-MeImpG are present, only a very small amount of the substrate survives (5%). The major products contain one C and 3–4 G residues. The time course of the reaction in the presence of C and G is illustrated in Figure 2b. The half-life for disappearance of starting material is 18 h.

Ribonuclease T1 is an enzyme that cleaves a 3′–5′ bond between a G residue and any nucleotide residue to yield products terminated by a 3′-phosphate and a 5′-hydroxyl group. Hydrolysis with ribonuclease T1 of the products formed by incubation of the template with 2-MeImpC and 2-MeImpG for 14 days gave a predominant product with higher electrophoretic mobility than S_1 (Figure 3a). It must be the 3′-phosphate derivative of S_1. This indicates that the first incorporated nucleotide, the C residue, is attached to the substrate mainly by a 3′–5′ bond. Weak bands corresponding to small amounts of material containing additional nucleotides are also observed. The corresponding products must contain a 2′–5′ linkage between the last G residue of S_1 and either a C residue or a G residue.

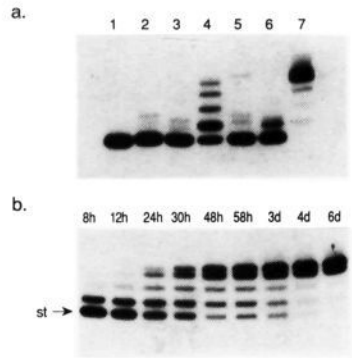


Figure 2. (a) Autoradiogram of the gel electrophoretogram of the products from the template reaction on substrate S_1. Lane 1, starting material S_1. Lanes 2–4 show products analyzed after 1 day: lane 2, S_1 with 2-MeImpG; lane 3, S_1 with 2-MeImpC; lane 4, S_1 with 2-MeImpC and 2-MeImpG. Lanes 5–7 correspond to the same reaction mixtures as in lanes 2–4 but analyzed after 7 days. (b) Time course of the reaction of S_1 with 2-MeImpC and 2-MeImpG.

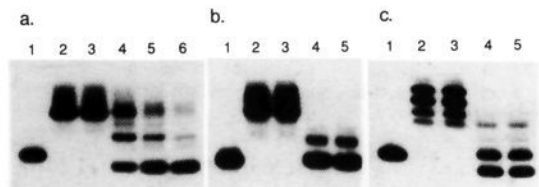


Figure 3. (a) Ribonuclease T1 analysis of the purified products from the reaction of S_1 with 2-MeImpC and 2-MeImpG (lane 2). Lane 1, S_1; lane 3, control in which the enzyme was omitted; lanes 4–6, the product from the reaction after treatment with ribonuclease T1 for 10, 20, and 40 min, respectively. (b) Ribonuclease A analysis of the products from the same reaction as (a) above. Lane 1, S_1; lane 2, purified products; lane 3, control in which the enzyme was omitted; lanes 4 and 5, after treatment with RNase A for 10 and 20 min, respectively. The samples used for ribonuclease T1 and ribonuclease A analysis were from different preparations. The latter contained a larger amount of pyrophosphate. (c) Ribonuclease T1 analysis of the product from the reaction of S_1 with 2-MeImpG alone. Lane 1, S_1; lane 2, the isolated products; lane 3, control without enzyme; lanes 4 and 5, after 10 and 20 min of digestion, respectively.

Ribonuclease A is an enzyme that cleaves 3′–5′ bonds between a pyrimidine residue and any nucleotide residue to yield products terminated by a 3′-phosphate and a 5′-hydroxyl group. Ribonuclease A hydrolysis of products from a 14-day reaction of S_1 with a mixture of 2-MeImpC and 2-MeImpG gave one predominant product with about the same electrophoretic mobility as the starting oligomer S_1 (Figure 3b). This product must be the 3′-phosphate of the monoC adduct of S_1. Thus the newly formed bond between C and a subsequent G is mainly 3′–5′ linked. A smaller amount of a product with lower mobility than the major product was also formed. This product was unaffected by alkaline phosphatase (data not shown), showing that the terminal 5′-phosphate was no longer free. It must, therefore, be an oligonucleotide capped by a pyrophosphate bond. Pyrophosphates accumulate slowly throughout the reaction and are significant in this case because we allowed the reaction to proceed for a longer than usual time (14 days) to maximize the yield of products.

The hydrolysis with ribonuclease T1 of the products obtained when G is the sole substrate gives two major products in roughly equal amounts (Figure 3c). One product has a higher mobility than S_1 and must be the 3′-phosphate of S_1. The other major product has a substantially lower mobility and corresponds to the 3′-phosphate of the monoG adduct of S_1. The small amounts of slower-moving material must also involve a 2′–5′ linkage to the primer. These results indicate that the first GpG bond to be formed is about 50% 2′–5′ linked and about 50% 3′–5′ linked. Most of the monoadduct obtained when the hairpin is incubated

(14) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 317–322.

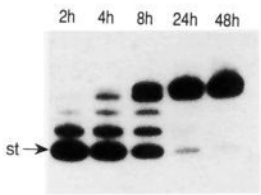


Figure 4. Time course for the reaction of S_2 with 2-MeImpG. St is the starting material.

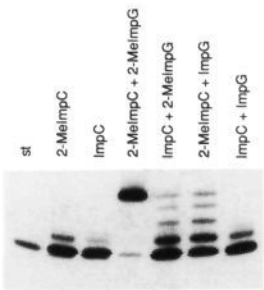


Figure 5. Reaction of S_1 with different combinations of nucleoside 5′-phosphoroimidazolides and nucleoside 5′-phosphoro(2-methyl)-imidazolides. The reactions were analyzed after 1 week. St is the starting material.

with 2-MeImpC alone is degraded by ribonuclease T1 (data not shown) to give a product with electrophoretic mobility greater than that of S_1, indicating that the C residue is predominantly 3′–5′ linked.

Substrate with an Isolated C Residue at Its 3′-Terminus—Oligomer S_2. The previous experiment establishes that when an isolated C residue is incorporated in the product strand it is possible to incorporate a G residue next to it efficiently and regiospecifically if the template strand contains an appropriately placed C residue. We studied the kinetics of a related reaction, the addition of G to the oligonucleotide S_2, which is terminated by a riboC residue. The time course of the addition is illustrated in Figure 4. The addition of G is efficient with a half-time of approximately 6 h.

Comparison of Imidazole and 2-Methylimidazole Derivatives as Nucleotide Donors. Substrate S_1 was incubated with 2-MeImpC and 2-MeImpG, ImpC and 2-MeImpG, 2-MeImpC and ImpG, or ImpC and ImpG under our standard conditions for 7 days. The nature of the products is illustrated in Figure 5. The substitution of ImpC for 2-MeImpC or of ImpG for 2-MeImpG substantially reduces the efficiency of the reaction. The substitution of ImpC for 2-MeImpC and of ImpG for 2-MeImpG prevents the formation of significant amounts of products.

Substrate with Two Consecutive G Residues in the Template Strand—Oligomer S_3. Figure 6a illustrates the products obtained when a substrate S_3 in which the template strand contains two consecutive G residues is incubated with 2-MeImpC, 2-MeImpG, or a mixture of both oligomers. Very little product is obtained by incubation of S_3 with 2-MeImpG, but virtually complete conversion to a single product is obtained when 2-MeImpC is used. The product is a monoC adduct. When S_3 is incubated with a mixture of 2-MeImpC and 2-MeImpG, it is converted efficiently to longer oligomers in which up to four or five nucleotides have been incorporated. The time course of the reaction with 2-MeImpC in the absence of 2-MeImpG is illustrated in Figure 6b. The half-time for the addition of the first C residue is about 24 h. The time course of the corresponding reaction when both 2-MeImpC and 2-MeImpG are present is illustrated in Figure 6c. The half-time for the disappearance of substrate is about 48 h. Enzymatic hydrolysis by ribonuclease T1 of the product obtained in the presence of 2-MeImpC along established that the C residue is linked to the primer by a 3′–5′-phosphodiester bond (data not shown).

Substrates with 3–5 Consecutive G Residues in the Template Strand—Oligomers S_4–S_6. Figure 7a–c illustrates the products formed when hairpin substrates containing three, four, or five

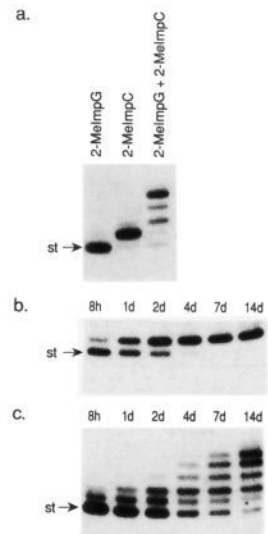


Figure 6. (a) Products from template reactions on substrate S_3 after 1 week. (b) Time course of the reaction of S_3 with 2-MeImpC. (c) Time course of the reaction of S_3 with 2-MeImpC and 2-MeImpG. St is the starting material.

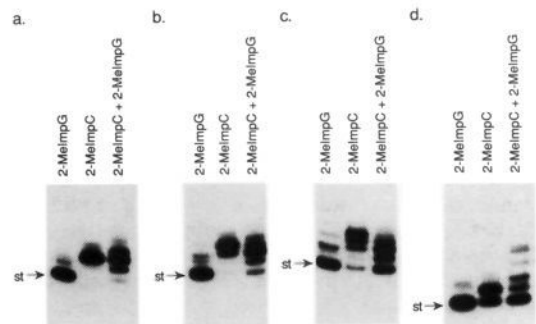


Figure 7. The products from the reaction of S_4–S_7 with 2-MeImpC and 2-MeImpG after 1 week: (a) S_4; (b) S_5; (c) S_6; (d) S_7. St is the starting material.

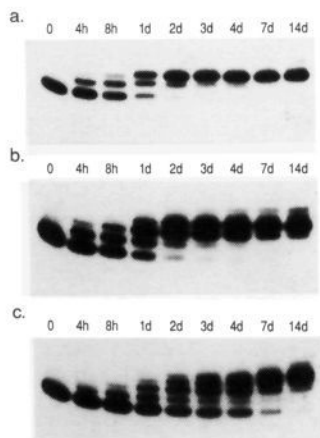


Figure 8. Time course of the reactions of substrates S_4–S_6 with 2-MeImpC only: (a) S_4; (b) S_5; (c) S_6.

consecutive G's in the template strand are incubated with 2-MeImpC, 2-MeImpG, or a mixture of 2-MeImpC and 2-MeImpG. Figure 8 illustrates the time courses of the reactions of S_4–S_6 with 2-MeImpC alone, and Figure 9 illustrates the time courses when both 2-MeImpC and 2-MeImpG are present. Comparison of the products formed on the three templates reveals many interesting trends.

Incubation of S_4 with 2-MeImpG alone yields small amounts of a monoadduct and little else. Incubation of S_5 under the same

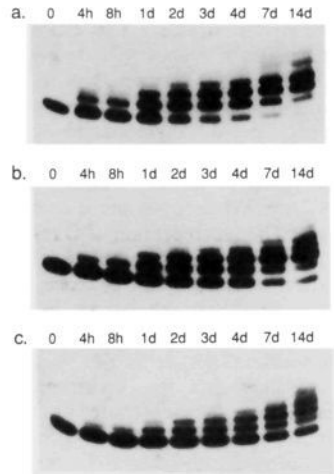


Figure 9. Time course of the reactions of substrates S_4–S_6 with 2-MeImpC and 2-MeImpG: (a) S_4; (b) S_5; (c) S_6.

conditions gives a similar yield of monoadduct and, in addition, detectable amounts of di-, tri-, and tetraadducts. Template S_6 gives rise to a significantly greater yield of 4 adducts that give an equally spaced ladder of 4 bands on the gel. A sequence of G residues apparently acts as a template for the oligomerization of 2-MeImpG.

Incubation of S_4 with 2-MeImpC alone results in virtually complete conversion to a diadduct. When S_5 is incubated with 2-MeImpC the major product is the triadduct, but a significant amount of the diadduct is also detected. With substrate S_6, this trend continues and di-, tri-, and tetraadducts are all obtained in substantial yield.

When oligomers S_4–S_6 are incubated with a mixture of 2-MeImpC and 2-MeImpG, the G derivative inhibits the incorporation of C residues. This is seen from the increased survival of the substrate oligomers in the presence of 2-MeImpG and from an increase in the concentration of the adducts containing small numbers of C residues at the expense of those containing larger numbers (Figure 9). However, G residues are incorporated after a short run of C residues, as would be expected. The addition of G residues after longer runs of C residues, particularly on substrate S_6, is inefficient.

The products obtained by incubating S_5 with 2-MeImpC only were treated with ribonuclease T1 (Figure 10a) and ribonuclease A (Figure 10b). In each case the products were almost completely degraded, establishing that the reaction is highly regiospecific and produces 3′–5′-linked phosphodiester bonds.

Formation of an Initial C–C Linkage on Oligomer S_7. The results obtained using templates S_3–S_6 show indirectly that bonds between a 3′-terminal C residue and a subsequent C residue form efficiently if there is an appropriately placed G residue in the template strand. We were able to observe this addition directly by using substrate S_7 which carries a 3′-terminal C residue. The products formed are exactly those expected in light of the results described above (Figure 7d). The addition of 2-MeImpG is inefficient and gives only a monoadduct. 2-MeImpC also gives a monoadduct, but in good yield. Finally a mixture of 2-MeImpC and 2-MeImpG gives a series of products in which 2–5 residues are added to the substrate.

Discussion

Incorporation of C Residues. It has already been established that isolated G residues in a GC copolymer containing a substantial excess of C residues direct the regiospecific incorporation of isolated 3′–5′-linked C residues into a product oligomer containing an excess of G.[4,6,8,10] This result is confirmed by the results obtained with polymers S_1 and S_2. The time-course of the reaction of a mixture of 2-MeImpC and 2-MeInpG on S_1 shows that the half-time for addition of C on a primer terminated by riboG is 18 h. The results obtained with substrate S_2 show the half-life for the incorporation of G on a primer terminated by riboC is about 6 h. In a previous paper[14] we have shown that the half-time for the incorporation of G residues on homopolymeric C templates is only 3 h. Thus, reactions of templates containing isolated G residues, although regiospecific and efficient, are slower than the reaction of 2-MeImpG on oligo(C) templates.

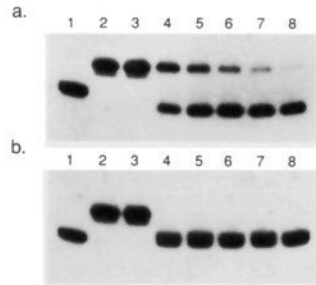


Figure 10. (a) Ribonuclease T1 analysis of the products from the reaction of S_5 with 2-MeImpC. Lane 1, S_5; lane 2, the isolated product from the template reaction; lane 3, control in which the enzyme was omitted; lanes 4–8, after treatment with ribonuclease T1 for 10, 20, 30, 40, and 60 min. (b) Ribonuclease A analysis of the same product as in (a) above. The lanes correspond to those given for (a).

Earlier studies of template-directed synthesis on random copolymers of riboC and riboG suggested strongly that, in the absence of self-structure, short oligo(G) sequences would direct the synthesis of products containing 2 and possibly more consecutive C residues.[7] This tentative conclusion is confirmed by our new results. Substrates S_3–S_6 contain runs of 2–5 consecutive G residues in the template strand. Those with 2–4 G residues, when incubated with 2-MeImpC and 2-MeImpG, give products in which the C residues are incorporated efficiently and G's are added after an appropriate run of C's. The reaction on S_6 did not give significant yields of products containing 5 C residues followed by G residues. We believe this is due to the long time required for completing a strand of 5 C residues rather than to a specific block to the incorporation of G after 5 C additions.

The results obtained when substrates S_3–S_6 are incubated with 2-MeImpC in the absence of 2-MeImpG are novel, and in some ways surprising. It is well-known that monomeric uridine derivatives do not form helical structures with polyadenylic acid.[13] It has not been possible to obtain evidence concerning the potential stacking of monomeric C derivatives on polyguanylic acid (poly(G)), since poly(G) forms a very stable tetrahelical self-structure.[13] Because of this lack of experimental evidence it has usually been supposed that poly(G):monoC helices are intrinsically unstable. Our results suggest otherwise. We find that a sequence of up to 5 consecutive G residues is an effective template for the formation of oligo(C)'s. This strongly suggests, but perhaps does not prove, that poly(G):monoC helices would be stable if self-association of poly(G) could be prevented.

The disappearance of substrates in the presence of 2-MeImpC and 2-MeImpG is slower than in the presence of 2-MeImpC only. Thus the presence of 2-MeImpG inhibits the incorporation of C. This may indicate the formation of hydrogen-bonded base-pairs between 2-MeImpC and 2-MeImpG in 0.1 M aqueous solution at 0 °C. Displacement of C from its binding site adjacent to the primer by stacked and possibly hydrogen-bonded G residues and the formation of tetrahelical complexes involving template oligo(G) segments and 2-MeImpG monomers are other possibilities.

Regiospecificity. We have used ribonuclease T1 and ribonuclease A to examine the nature of the bonds formed in the oligomerization reaction on a number of templates. We have confirmed our previous conclusion that the incorporation of an isolated C residue on a template containing an isolated G residue is highly regiospecific, leading to the formation of 3′–5′-linked CG and GC neighbors (Figure 3). In addition, we have now shown that all of the bonds formed on templates containing a series of consecutive G residues are predominantly 3′–5′ bonds. It is particularly striking that the regiospecificity of the addition of C to a primer that is terminated by C has the same high regiospecificity as the addition of G to G on a poly(C) template.

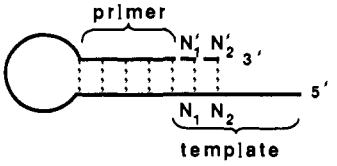


Figure 11. Neighboring base effect: the rate of incorporation of the nucleotide N_1' into products is strongly dependent on the presence of the 5'-phosphoro(2-methyl)imidazolide of the downstream base N_2'.

The misincorporation of G opposite G in the template is not regiospecific, presumably because the phosphate group of the incoming residue does not adopt an appropriate position relative to the 3'-OH group of the primer.

Effect of Backbone Structure. In our previous studies we have found that oligopyrimidines and the corresponding oligodeoxypyrimidines are almost equivalent as templates. This was not unexpected, since DNA–RNA hybrids and double-stranded RNA both adopt the A DNA structure. In the stem–loop structure, the double-stranded deoxynucleotide segment presumably adopts the standard B DNA structure. The strong similarity in the template properties of hairpin substrates and oligoribonucleotides or oligodeoxynucleotides implies that the B structure does not propagate into the ribonucleotide-containing helix.

Effect of a Nearest Neighbor Hydrogen-Bonded Monomer. The addition of a C residue to substrate S_1 is a slow and inefficient reaction in the absence of 2-MeImpG. In the presence of 2-MeImpG, the substrate is transformed efficiently to products. Since the first step in this reaction sequence is the incorporation of C, it follows that the presence of 2-MeImpG greatly accelerates the rate of incorporation of C into products.

The results obtained with oligomers S_2, S_4, S_5, and S_6 point in the same direction. When there are n consecutive G residues in the template, incubation of the substrate with 2-MeImpC alone results in the incorporation of a maximum of $(n - 1)$ C residues in the product. Thus a C residue is incorporated efficiently only if it is possible to stack another C residue downstream on the template.

The importance of the occupancy of the neighboring downstream position on the template (Figure 11) on the rate of template-directed reactions seems to be a general one. We have obtained results analogous to those described above on a variety of hairpin substrates containing A and T residues (unpublished results). The slow rate of addition of the last G residue on a substrate containing an oligo(C) sequence[14] can be considered another example of this nearest neighbor effect.

There are two obvious explanations of the effect of neighboring bases. The first proposes that the occupation of the site adjacent to the "primer" strand of the substrate is low unless the next downstream site is occupied. Stacking interaction between two or more monomers attached to the template could supply the free energy needed to guarantee the occupancy of the site next to the primer. In the second explanation, it is proposed that the presence of a neighboring base on the template "catalyzes" the addition reaction at the "primer" terminus. The increased efficiency is attributed to a more favorable reaction rate at the "primer" terminus when the site is occupied, rather than to a greater occupancy of the reaction site.

The experiments described here do not address the issue directly. However, we doubt that the first model, alone, can account for the quantitative differences that we observe, particularly those for substrate S_3 which include only two consecutive G residues on the template strand. Stacking between the terminal G residue of the primer strand and a C residue should be stronger than stacking between two consecutive C monomers. Thus if the occupancy of an isolated C site next to G is low, the ratio of double occupancy to single occupancy on substrate S_3 should also be low. Thus our results suggest that there is a significant effect of a downstream hydrogen-bonded monomer over and above any effect on the occupation of the site adjacent to the primer. The results obtained when one or both of the two 2-methylimidazolides was replaced by an unsubstituted imidazolide support this proposal.

The Effect of the 2-Methyl Groups of the 2-Methylphosphoroimidazolide Substrates. The efficiency with which 5'-phosphoroimidazolides of guanosine undergo template-directed oligomerization of poly(C) is very sensitive to the nature of the substituents on the imidazole.[1] In particular, 2-MeImpG undergoes efficient regiospecific polymerization to give 3'–5'-linked products while ImpG reacts poorly and gives comparable amounts of 2'–5'- and 3'–5'-linked products. Recently we have found that the effect of the 2-methyl group is more complicated than we originally believed. The incorporation of C residues from ImpC or 2-MeImpC on a GCGCGCG template is enhanced by the presence of 2-MeImpG but not by the presence of ImpG.[15] Thus, methyl substitution on a base other than the one that is to be incorporated in the product can have a profound effect on the efficiency of the reaction.

The results obtained when S_1 reacts with various combinations of 2-methylimidazolides and unsubstituted imidazolides of C and G provide further evidence for a "neighbor" effect (Figure 5). The combination of 2-MeImpC with 2-MeImpG, as we have seen, leads to the efficient synthesis of product oligomers. As anticipated, the combination of ImpC and ImpG gives very little product. However, the combination of ImpC with 2-MeImpG and the combination of 2-MeImpC with ImpG gave substantial yields of products. Surprisingly, the effect of methyl substitution in the G substrate was almost as pronounced as substitution in the C substrate. The exchange of imidazole and 2-methylimidazole subsequent to the partial hydrolysis of the phosphoroimidazolides cannot explain this result. The half-time for the hydrolysis of 2-MeImpG under almost the same conditions that we used is about 50 days,[16] while the equivalence of the mixture of 2-MeImpG + ImpC or of ImpG and 2-MeImpC is evident after only 2 days (data not shown). Our results, therefore, establish that a 2-methyl substitutent on the residue one position downstream from the residue that adds to the primer (Figure 11) plays an important role in the template-directed reactions of 2-methylimidazole derivatives.

The nature of the catalysis by the downstream residue is unclear. It seems to us unlikely that it is a purely steric effect in which the downstream neighbor reorients the reacting C residue—such a large effect of a methyl group in a relatively unconstrained system would be unprecedented. Perhaps the imidazole moiety of the neighboring residue acts as an acid–base catalyst. Interaction with the 2-methylimidazole group or with the pentacovalent phosphate group in the transition state is possible and might be sensitive to the stereochemical effect of methyl substitution.

Misincorporation of G Opposite G. The use of hairpin oligonucleotides of defined sequence enables us to monitor relatively inefficient misincorporation reactions. The misincorporation of C opposite C in the template does not occur to a significant extent. However, after 7 days, it is possible to detect the products formed by misincorporation of G opposite G residues in the template strand in appreciable yield.

A comparison of the extent of misincorporation of G opposite sequences of G of different lengths leads to an unexpected conclusion. The longer the string of consecutive G residues in the template, the more efficient is the misincorporation reaction. Our results, therefore, suggest that a sequence of consecutive G residues, provided it cannot undergo self-association, can act as a template for oligo(G) synthesis. This is an example of template-directed synthesis based on identity pairing.

General Implications. The experiments reported in this paper show that template strands contain strings of consecutive G residues that when appropriately primed can direct the incorporation of several consecutive 3'–5'-linked C residues. There is already much evidence that C-rich copolymers direct the synthesis of complementary G-rich oligomers. Taken together these findings imply that if the tetrahelical self-structure of G-rich oligomers could be prevented, the development of pairs of complementary

(15) Ng, K.-m. E.; Orgel, L. E. *J. Mol. Evol.* **1989**, *29*, 101–107.
(16) Kanavarioti, A.; Chang, S.; Alberas, D. *J. Mol. Evol.* **1990**, *31*, 462–469.

sequences, each of which directs the synthesis of the other, would be possible. A demonstration of complementary syntheses of this kind would have important implications for discussions of the origin of biological information storage.

We do not see any way of destabilizing the oligo(G) self-structure without greatly weakening the interaction of oligo(G) with C-containing monomers. However, the substitution of 7-deaza-G[17] (or 7-deaza-8-aza-G)[18] for G in a polynucleotide prevents the formation of multi-strand self-structures. We speculate that poly(7-deazaG) may form a stable double-helix with appropriate monomeric C derivatives and that it may be possible to develop pairs of complementary templates containing riboC and ribo-7-deaza-G residues.

We emphasize that the above suggestion is not directly relevant to the origins of life. We do not believe that 2-MeImpC, 2-MeImpG, and 7-deaza-G were present in large amounts on the primitive earth. However, the development of pairs of efficient, complementary templates would represent on important step toward the realization of a nonenzymatic replicating system that stores information.

(17) Seela, F.; Tran-Thi, Q.-H.; Franzen, D. *Biochemistry* **1982**, *21*, 4338-4343.
(18) Seela, F.; Driller, J. *Nucleic Acids. Res.* **1989**, *17*, 901-910.

Experimental Section

Materials and methods used in this work, with the exception of those described explicitly below, were the same as those described in part 1 of the series.[14]

Ribonuclease A was purchased from Boehringer Mannheim. Cytidine 5'-phosphoro(2-methyl)imidazolide (2-MeImpC), guanosine 5'-phosphoro(2-methyl)imidazolide (2-MeImpG), cytidine 5'-phosphoroimidazolide (ImpC), and guanosine 5'-phosphoroimidazolide (ImpG) were prepared by a published procedure.[19]

Ribonuclease A digestion of products was carried out as follows. An aliquot of the gel-purified product (3500 cpm) was dissolved in 14 μL of RNA sequencing buffer (7 M urea, 20 mM sodium citrate, 1 mM EDTA; pH 5). Ribonuclease A (5 units, 1 μL) was added and the solution was incubated at 37 °C. Aliquots (3 μL) were withdrawn after 10, 20, 40, 60, and 90 min and kept over a dry ice/ethanol mixture until they were analyzed by electrophoresis on polyacrylamide gels.

Reaction mixtures contained each relevant nucleoside phosphoroimidazolide at a concentration of 0.1 M.

Acknowledgment. This work was supported by Grant NAWG-1660 from the National Aeronautics and Space Administration. We thank Aubrey Hill for technical assistance and Sylvia Bailey for manuscript preparation.

(19) Joyce, G. F.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1984**, *176*, 279-306.

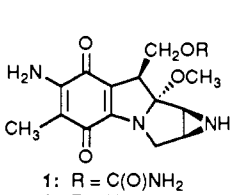
Recognition of Mitomycin C-DNA Monoadducts by UVRABC Nuclease

Harold Kohn,*,† Ven-Shun Li,† and Moon-shong Tang*,‡

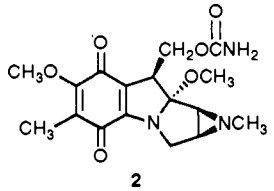
Contribution from the Department of Chemistry, University of Houston, Houston, Texas 77204-5641, and The University of Texas System M. D. Anderson Cancer Center, Science Park—Research Division, Smithville, Texas 78957. *Received December 20, 1991*

Abstract: The *Escherichia coli* UVRABC nuclease assay has been shown to be an excellent method to monitor mitomycin C-DNA monoalkylation transformations. Analysis of the mitomycin C-induced and 10-decarbamoylmitomycin C-induced DNA incision sites from five different DNA restriction fragments after UVRABC treatment revealed that all the drug-induced UVRABC incisions can be attributed to drug bonding at guanine residues and that bonding proceeded in a highly sequence selective manner. The densitometric data indicated that both the 5' and 3' nearest neighbor bases surrounding the guanine site affect its susceptibility to drug modification. Mitomycin lesions occurred predominantly at 5'CG sequences, and 5'CGG sequences were the preferred trinucleotide units for mitomycin C monoalkylation transformations. Potential contributing factors (i.e., C(10) carbamate moiety, C(8) hydroquinone hydroxyl group) responsible for the sequence selectivity of this process have been examined and a rationale offered for the DNA bonding specificity. The advantages of the UVRABC assay versus the previously employed λ-exo protocol for the detection of mitomycin C monoalkylation sites are discussed.

Mitomycin C (**1**) is a clinically significant antineoplastic agent.[1] Mechanisms have been advanced that suggest that drug function proceeds by initial reductive activation of **1** followed by attachment of mitomycin C to DNA to generate first mono-alkylated and then bis-alkylated products.[2] The ability of **1** to bond to complementary strands of DNA has been associated with its pronounced cytotoxicity and antitumor activity.[2a,f] Accordingly, reports have appeared concerning the DNA base sequence specificity for the formation of *cross-linked* mitomycin C-DNA adducts.[3-5] Evidence has been provided that these lesions are generated at complementary 5'CG sequences and explanations have been offered to account for this phenomenon. Recently, we demonstrated that the initial mitomycin C *monoalkylation* event to DNA oc-



1: R = C(O)NH$_2$
3: R = H



2

† University of Houston.
‡ The University of Texas System M. D. Anderson Cancer Center.

(1) Carter, S. K.; Crooke, S. T. *Mitomycin C. Current Status and New Developments*; Academic Press: New York, 1979.
(2) (a) Iyer, V. N.; Szybalski, W. *Science* **1964**, *145*, 55. (b) Remers, W. A. *The Chemistry of Antitumor Antibiotics*; Wiley: New York, 1979; Vol. 1, pp 221-276. (c) Franck, R. W.; Tomasz, M. In *The Chemistry of Antitumor Agents*; Wilman, D. E. V., Ed.; Blackie and Sons, Ltd.: Glasgow, Scotland, 1990; pp 379-394. (d) Fisher, J. F.; Aristoff, P. A. *Prog. Drug Res.* **1988**, *32*, 411. (e) Moore, H. W.; Czerniak, R. *Med. Res. Rev.* **1981**, *1*, 249 and references therein. (f) Zwelling, L. A.; Anderson, T.; Kohn, K. W. *Cancer Res.* **1979**, *39*, 365.
(3) Teng, S. P.; Woodson, S. A.; Crothers, D. M. *Biochemistry* **1989**, *28*, 3901.
(4) (a) Weidner, M. F.; Millard, J. T.; Hopkins, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 9270. (b) Millard, J. T.; Weidner, M. F.; Raucher, S.; Hopkins, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 3637. (c) Weidner, M. F.; Signurdsson, S. T.; Hopkins, P. B. *Biochemistry* **1990**, *29*, 9225.
(5) Borowy-Borowski, H.; Lipman, R.; Tomasz, M. *Biochemistry* **1990**, *29*, 2999.

0002-7863/92/1514-5501$03.00/0