VOLUME 114, NUMBER 21
OCTOBER 7, 1992

JOURNAL
OF THE
AMERICAN
CHEMICAL
SOCIETY

Nonenzymatic Template-Directed Synthesis on Hairpin Oligonucleotides. 3. Incorporation of Adenosine and Uridine Residues

Taifeng Wu and Leslie E. Orgel*

Contribution from the Salk Institute for Biological Studies, Post Office Box 85800, San Diego, California 92186-5800. *Received January 23, 1992*

Abstract: We have used [^{32}P]-labeled hairpin oligonucleotides to study template-directed synthesis on templates containing one or more A or T residues within a run of C residues. When nucleoside-5′-phosphoro(2-methyl)imidazolides are used as substrates, isolated A and T residues function efficiently in facilitating the incorporation of U and A, respectively. The reactions are regiospecific, producing mainly 3′-5′-phosphodiester bonds. Pairs of consecutive non-C residues are copied much less efficiently. Limited synthesis of CA and AC sequences on templates containing TG and GT sequences was observed along with some synthesis of the AA sequences on templates containing TT sequences. The other dimer sequences investigated, AA, AG, GA, TA, and AT, could not be copied. If A is absent from the reaction mixture, misincorporation of G residues is a significant reaction on templates containing an isolated T residue or two consecutive T residues. However, if both A and G are present, A is incorporated to a much greater extent that G. We believe that wobble-pairing between T and G is responsible for misincorporation when only G is present.

Introduction

We have recently used oligonucleotides with hairpin secondary structures to study the template-directed oligomerization of cytidine and guanosine 5′-phosphoro(2-methyl)imidazolides (2-MeImpC and 2-MeImpG).[1,2] We synthesize oligodeoxynucleotide sequences that form stem loop secondary structures (Figure 1a). The template sequences are the 5′-terminal single-strand portions of the hairpin. The 3′-terminal segments of the oligonucleotide, which are part of the double-helical stems, act as intramolecular "primers" for template-directed incorporation of monomers. Ribonucleotides are introduced as terminal residues at the 3′-ends of the oligomers to improve the efficiency of the reactions.[1] The hairpins are [^{32}P]-labeled at their 5′-termini, and the products are analyzed by autoradiography of gel electrophoretograms.

The use of the substrate oligonucleotides at nanomolar concentrations eliminates most of the problems caused by the association of two or more templates. The use of oligodeoxynucleotide hairpins also simplifies the analysis of the reaction products. The method has revealed many novel features of template-directed reactions involving C and G residues. In this paper we extend our studies to reactions in which T and A residues are incorporated into the template sequence.

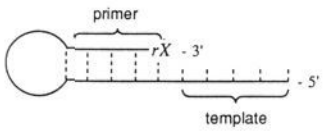
Results

Reactions on a Template Containing an Isolated T Residue—S_1. Figure 2a illustrates the products from the incubation of the hairpin oligomer S_1 (Figure 1b) with 2-MeImpA, 2-MeImpG or a mixture of 2-MeImpA and 2-MeImpG. After 1 day only a small amount of product is detectable if 2-MeImpA or 2-MeImpG is present separately. When both activated nucleotides are present together, adducts with up to four incorporated residues can be detected. After 7 days, only 28% of mono-A adducts are formed when A alone is present in the incubation mixture. With G alone, as much as 50% of tetra- and penta-G adducts are formed. When both nucleotides are present in the incubation mixture, the major products contain 4 or 5 added residues, while only 14% of the substrate survives unchanged. Figure 2b illustrates the time course of the reaction when both nucleotides are present. The substrate is converted fairly rapidly to mono-, di-, tri-, and tetraadducts.

(1) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 317–322.
(2) Wu, T.; Orgel, L. E. Submitted for publication.

0002-7863/92/1514-7963$03.00/0

a.



b.

Template	Sequence
S_1	CCCCTCCTAGTTTTTCTAGrG
S_2	CCCCACCTAGTTTTTCTAGrG
S_3	CCCATCCTAGTTTTTCTAGrG
S_4	CCCTACCTAGTTTTTCTAGrG
S_5	CCCAACCTAGTTTTTCTAGrG
S_6	CCCGACCTAGTTTTTCTAGrG
S_7	CCCAGCCTAGTTTTTCTAGrG
S_8	CCCTTCCTAGTTTTTCTAGrG
S_9	CCTTTCCTAGTTTTTCTAGrG
S_{10}	CCCTGCCTAGTTTTTCTAGrG
S_{11}	CCCGTCCTAGTTTTTCTAGrG
S_{12}	CCCCAACTAGTTTTTCTAGrU
S_{13}	CCCCGACTAGTTTTTCTAGrU
S_{14}	CCCCTACTAGTTTTTCTAGrU
S_{15}	CCCCCACTAGTTTTTCTAGrU
S_{16}	CCCCCTCTAGTTTTTCTAGrA

Figure 1. (a) Hairpin structure formed by the substrates used in this study. (b) Sequences of substrates referred to in the text.

The addition of a fifth residue is slow. The half-time for the disappearance of substrate is about 24 h.

Ribonuclease T1 is an enzyme that cleaves a 3′-5′-bond between a riboG residue and any ribonucleotide residue to yield products terminated by a 3′-phosphate and a 5′-hydroxyl group. The products resulting from the incubation of S_1 with a mixture of 2-MeImpA and 2-MeImpG for 2 weeks were isolated by preparative gel electrophoresis and subjected to degradation with ribonuclease T1. The sole product moved faster on the gel than the original substrate and must, therefore, correspond to the 3′-phosphate of the substrate (Figure 3a). This establishes that the bond from the terminal G of the substrate to the first incorporated base is a 3′-5′ internucleotide bond.

Ribonuclease H is an enzyme that cleaves phosphodiester bonds between ribonucleotides in hybrid double helices containing one oligoribonucleotide and one oligodeoxyribonucleotide chain. The products formed are terminated by a 5′-phosphate group and a 3′-hydroxyl group. Enzymatic degradation by ribonuclease H of the products obtained by incubating S_1 with a mixture of 2-MeImpA and 2-MeImpG gave a major product with mobility corresponding to a monoadduct of the substrate, but degradation was incomplete and a number of slower moving bands containing as much as 20% of total radioactivity remained even after prolonged incubation (Figure 3b).

The degradation with ribonuclease T1 of the products obtained by the incubation of S_1 with 2-MeImpG in the absence of 2-MeImpA gave two major products (Figure 3c). The most abundant product (~70%) had a greater mobility than the starting substrate and must be the 3′-phosphate of S_1. The less abundant (~30%) moved slightly slower than S_1 and is therefore the 3′-phosphate of the adduct of S_1 with a single G residue. This shows that G incorporation opposite a T residue in the template gives about 70% of 3′-5′-linked product and about 30% of 2′-5′-linked product.

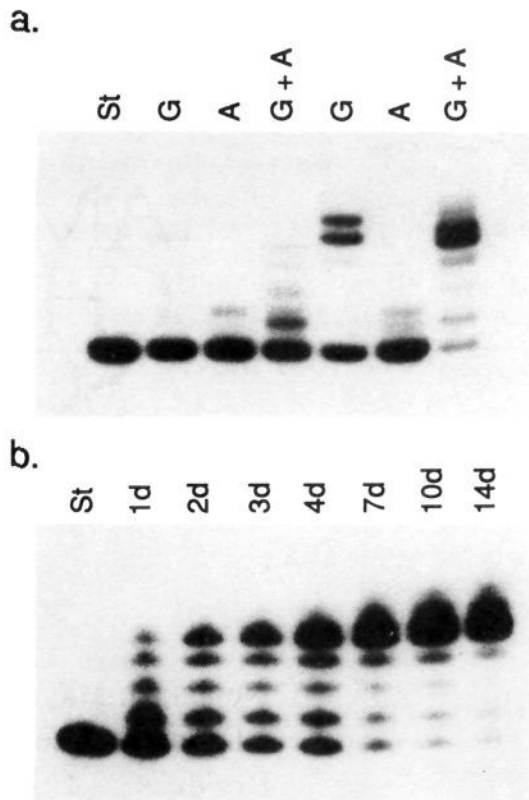


Figure 2. (a) Autoradiogram of the gel electrophoretogram of the products from the reactions of substrate S_1. Lane 1, the starting material, S_1; lanes 2–4 and lanes 5–7 are the products after reaction times of 1 and 7 days, respectively. The activated nucleotides present in the reaction mixture are indicated above the lanes. St indicates the starting substrate. (b) Time course of the reaction of substrate S_1 with a mixture of 2-MeImpG and 2-MeImpA.

We did not isolate the products from the reaction of S_1 with 2-MeImpA alone since the yield was very small. The electrophoretic patterns obtained from a crude reaction mixture before and after attempted ribonuclease T1 digestion were identical (data not shown). Hydrolysis of the reaction mixture with alkali resulted in the disappearance of the band corresponding to the monoA adduct and to the appearance of a band with greater mobility than the starting material. This shows that, in the monoadduct, the A residue is joined to S_1 by a 2′-5′-phosphodiester bond.

Reactions on a Template Containing an Isolated A Residue—S_2. Figure 4a illustrates the products from the reaction of S_2 (Figure 1b) with 2-MeImpG, 2-MeImpU, or a mixture of 2-MeImpG and 2-MeImpU. After 1 day very little product is formed with either reagent separately, but significant amounts of products with 1–4 incorporated nucleotides are formed if both are present. After 7 days little product is obtained with U alone (20%), but a greater yield of products with up to four incorporated residues is obtained with G alone (38%). When both activated nucleotides are present, most of the substrate (66%) is converted to adducts with three or more incorporated residues. The time course of the reaction with a mixture of both substrates is illustrated in Figure 4b. The half-life for disappearance of the substrate is about 4 days.

The products formed by the incubation of S_2 with a mixture of 2-MeImpG and 2-MeImpU for 2 weeks were isolated by preparative gel electrophoresis and degraded with ribonuclease T1 and ribonuclease A. Digestion with ribonuclease T1 gave a predominant product with a mobility greater than that of the starting substrate (Figure 5a). It must be the 3′-phosphate of S_2. Thus the bond from the primer to the first residue incorporated is predominantly 3′-5′-linked. Digestion with ribonuclease A was almost complete (Figure 5b). The product of degradation coelectrophoresed with the original substrate. It must be the 3′-phosphate of the pU adduct of the substrate. These results es-

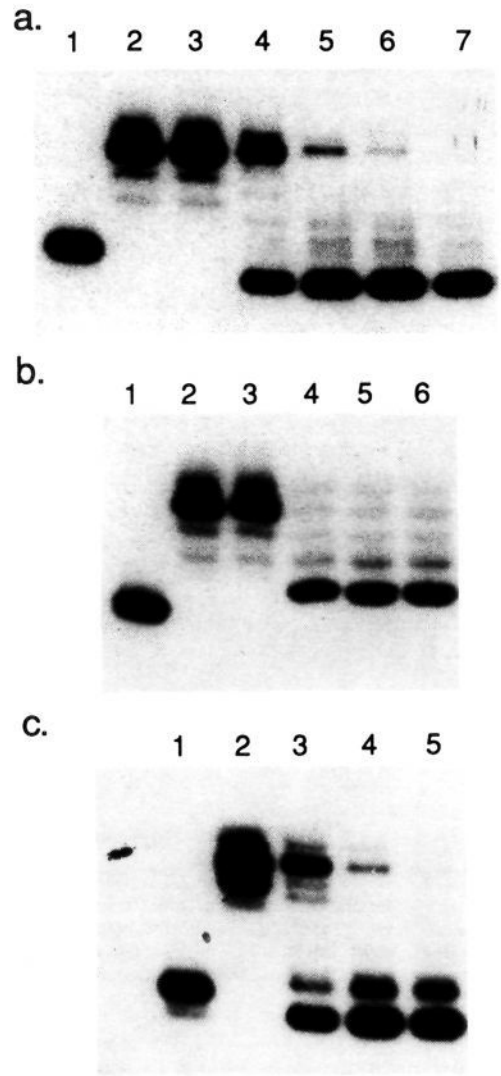


Figure 3. (a) Ribonuclease T1 analysis of the products isolated from the reaction of substrate S_1 with 2-MeImpG and 2-MeImpA after 2 weeks. Lane 1, substrate S_1; lane 2, the isolated products; lane 3, control incubated without enzyme; lanes 4–7, the product after treatment with ribonuclease T1 for 5, 20, 40, and 60 min. (b) Ribonuclease H analysis of the product from the same reaction as in (a) above. Lane 1, substrate S_1; lane 2, the isolated products; lane 3, control incubated without enzyme; lanes 4–6, the product after treatment with ribonuclease H for 10, 20, and 30 min. (c) Ribonuclease T1 analysis of the product isolated from the reaction of substrate S_1 with 2-MeImpG alone. Lane 1, substrate S_1; lane 2, the isolated products; lanes 3–5, the product after treatment with ribonuclease T1 for 5, 20, and 40 min.

tablished that the template directs the synthesis first of a 3′-5′-phosphodiester bond to U and then of a 3′-5′-phosphodiester bond to G.

The products formed by incubating S_2 with 2-MeImpG alone for 2 weeks were isolated by gel electrophoresis and then subjected to hydrolysis by ribonuclease T_1. We obtained two significant products (Figure 5c). The minor product (25%) had a mobility higher than that of S_2 and must be a 3′-phosphate of S_2. The major product (75%) had a mobility slightly lower than that of S_2 and must be the 3′-phosphate of the pG adduct of S_2. This implies that about 75% of the phosphodiester bonds formed in this reaction are 2′-5′-linked.

The product from the reaction of S_2 with 2-MeImpU alone was analyzed by a procedure very similar to that described above for the product obtained from S_1 and 2-MeImpA. We concluded that U is linked to S_1 predominantly by a 2′-5′-internucleotide bond.

Reactions on Templates Containing TA, AT, AA, AG or GA Sequences—Substrates S_{3-7}. Substrates S_3, S_4, S_5, S_6, and S_7

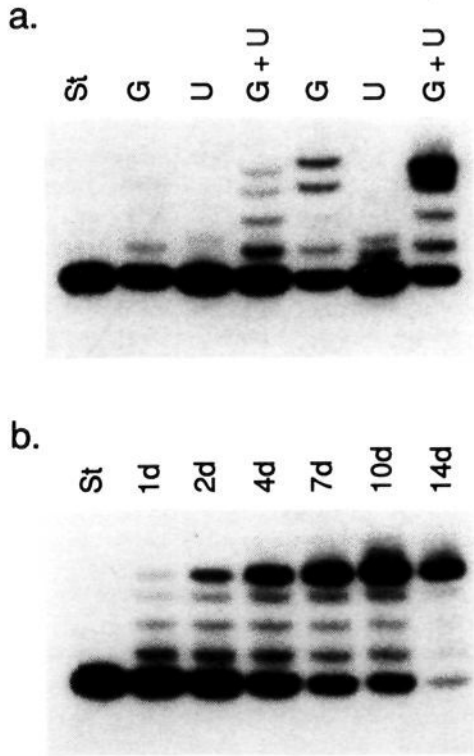


Figure 4. (a) Autoradiogram of the gel electrophoretogram of the products from the reactions of substrate S_2. Lane 1, starting material S_2; lanes 2–4 and 5–7, the products after 1 to 7 days reaction time, respectively. (b) Time course of the reaction of substrate S_2 with a mixture of 2-MeImpU and 2-MeImpG.

(Figure 1b), containing TA, AT, AA, AG, and GA sequences, respectively, in the template strand, did not react efficiently with any combination of activated nucleotides. Small amounts of monoadducts were formed after 1 week in many reactions, but in no case could we detect significant amounts of adducts containing two or more incorporated nucleotides.

Reaction on Templates with Two or Three Consecutive T Residues—Substrates S_8 and S_9. The incubation of substrate S_8 (Figure 1b) with 2-MeImpA alone gave very little product. However, incubation with either 2-MeImpG alone or a mixture of 2-MeImpG and 2-MeImpA for 7 days gave a mixture of products up to pentaadducts in moderate yield (Figure 6). The similarity between the electrophoretograms of the products obtained with G only and those obtained with G and A, lanes 4 and 6 or lanes 1 and 3 (Figure 6), left ambiguous the nature of the bases incorporated when both A and G were present. This ambiguity was resolved by subjecting the relevant products to ribonuclease T1 digestion.

Figure 7 illustrates the products obtained by ribonuclease T1 digestion of the products obtained by incubating S_8 with G alone (lanes 3–5) or A and G (lanes 7–9). It is clear from lane 5 that the majority of the products resist hydrolysis to the 3′-phosphate of S_8. It follows that the bonds formed from the primer to a misincorporated G residue must be mainly 2′-5′-linked. The results in lane 9 show that the first residue attached to the primer when both A and G are present is mainly hydrolyzed by the enzyme to the 3′-phosphate of S_8 and is therefore 3′-5′-linked. It follows from these two results that, when both bases are present, correct incorporation of A is the main reaction. We cannot tell whether the presence of some slower moving material in lane 9 reflects misincorporation of G or the incorporation of A via a 2′-5′-bond.

Incubation of S_9 (Figure 1b), a substrate in which the template strand includes a sequence of three consecutive T residues, with 2-MeImpA or 2-MeImpG alone or a mixture of both resulted in the formation of only small amounts of monoadducts.

Reactions on Templates Containing GT and TG Sequences—Substrates S_{10} and S_{11}. The efficiency of reactions on substrates

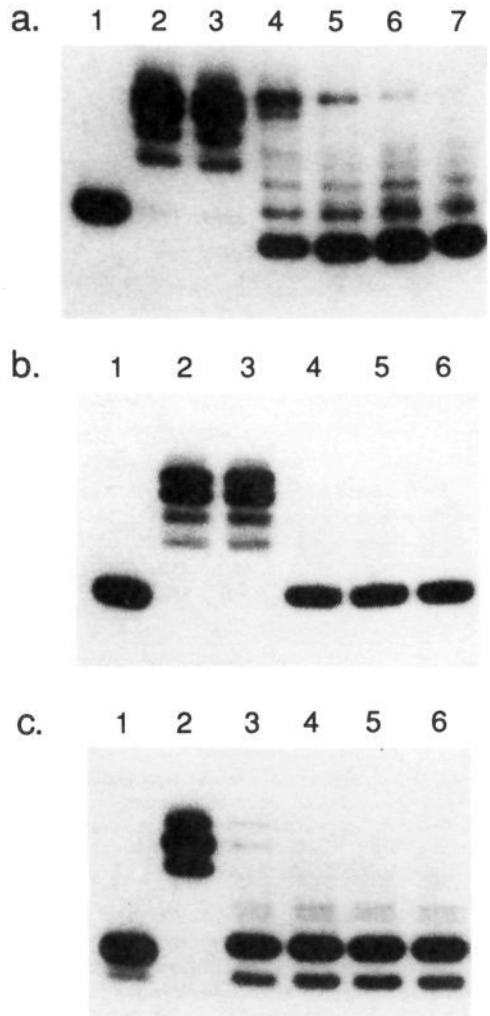


Figure 5. (a) Ribonuclease T1 analysis of the product isolated from the reaction of substrate S_2 with a mixture 2-MeImpG and 2-MeImpU. Lane 1, substrate S_2; lane 2, the isolated products; lane 3, control incubated without enzyme; lanes 4–7, the product after treatment with ribonuclease T1 for 5, 25, 40, and 60 min. (b) Pancreatic ribonuclease analysis of the same product as in (a). Lane 1, substrate S_2; lane 2, the isolated products; lane 3, control incubated without enzyme; lanes 4–6, the products after treatment with enzyme for 10, 20, and 40 min. (c) Ribonuclease T1 analysis of the purified product from the reaction of substrate S_2 with 2-MeImpG only. Lane 1, substrate S_2; lane 2, the isolated products; lanes 3–6, product after treatment with ribonuclease T1 for 5, 10, 30, and 60 min.

S_{10} and S_{11} (Figure 1b) in which the template strands contain GT and TG sequences, respectively, is much lower than the efficiency on templates containing isolated T or G residues. After incubation for 1 week, only mono- and diadducts could be detected (data not shown), but after incubation for 2 weeks with a mixture of 2-MeImpC, 2-MeImpG, and 2-MeImpA, substantial yields of mono- and diadducts were formed as well as smaller amounts of higher adducts (Figure 8a and b).

Template-Directed Addition of U, C, or A Residues to a Primer Terminated by rU—Substrates S_{12}, S_{13}, and S_{14}. The experiments described above show that AA, AT, TA, and TT sequences on the template strand block template-directed synthesis. To investigate this effect further we prepared substrates in which the primer is terminated by a riboU residue and studied the template-directed incorporation of a U, A, or C residue next to the primer. The incubation of substrate S_{12} (Figure 1b) with 2-MeImpG, 2-MeImpU or a mixture of the two failed to yield significant amounts of products (Figure 9a). Substrates S_{13} and S_{14} (Figure 1b), on the other hand, gave modest yields of products (59% and 68%, respectively) when incubated with an appropriate

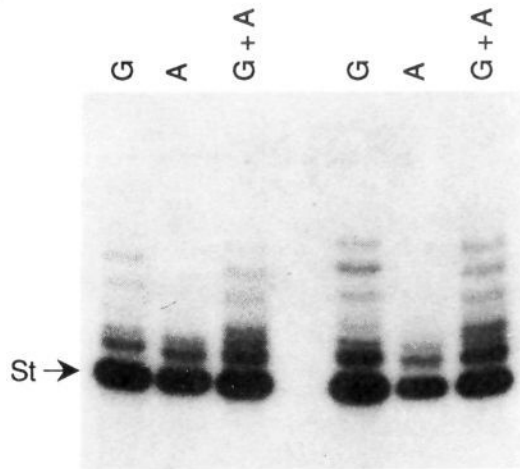


Figure 6. Autoradiogram of the gel electrophoretogram of the products from template reactions on substrate S_8 analyzed after 1-week (lanes 1–3) and 2-week reaction (lanes 4–6).

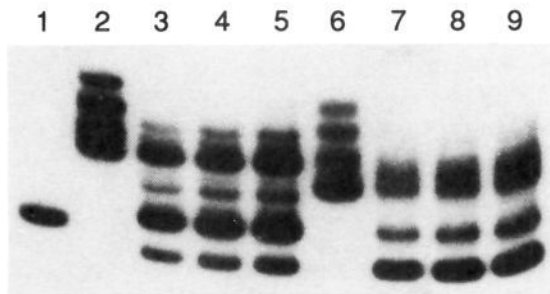


Figure 7. Ribonuclease T1 analysis of the products from the reaction of substrate S_8 with 2-MeImpG and 2-MeImpA and from the reaction with 2-MeImpG alone. Lane 1, substrate S_8; lane 2, the product isolated from the reaction of S_8 with 2-MeImpG alone after 2 weeks; lanes 3–5, products formed after treatment of the isolated material illustrated in lane 2 with ribonuclease T1 for 20, 40, and 60 min; lane 6, the product isolated from the reaction of substrate S_8 with 2-MeImpA and 2-MeImpG; lanes 7–9, products formed after treatment of the isolated material in lane 6 with ribonuclease T1 for 25, 40, and 60 min.

mixture of two activated monomers but gave very little product when incubated with either activated monomer separately (Figure 9b,c).

Template-Directed Addition of G Residues to Primers Terminated by rU or rA—Substrates S_{15} and S_{16}. In an earlier paper we showed that the template-directed incorporation of a G residue after a riboC residue at the terminus of the primer is rapid. We have now extended this work to study the incorporation of a G residue next to a terminal rU or rA residue. The time course of the addition of 2-MeImpG to S_{15} (Figure 1b), a substrate containing a terminal rU residue, is illustrated in Figure 10a. The half-time for the disappearance of the substrate is about 8 h. The time course of the addition of 2-MeImpG to S_{16} (Figure 1b), a substrate terminated by a riboA residue, is illustrated in Figure 10b. The half-time for the disappearance of substrate is about 5 h.

Discussion

Comparison with Earlier Work. Detailed studies of the copolymerization of a mixture of 2-MeImpA and 2-MeImpG on random copolymers of riboC and riboU clearly demonstrate that product oligomers containing A and G are formed in good yield, provided that the template contained an excess of C residues over U residues.[3-5] The efficiency of incorporation of A is only about

(3) Inoue, T.; Orgel, L. E. *Science* **1983**, *219*, 859–862.
(4) Joyce, G. F.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1984**, *176*, 279–306.
(5) Joyce, G. F. In *Cold Spring Harbor Symposia on Quantitative Biology*; Cold Spring Harbor Laboratory: Cold Spring Harbor, New York, 1987; Vol. LII, pp 41–51.

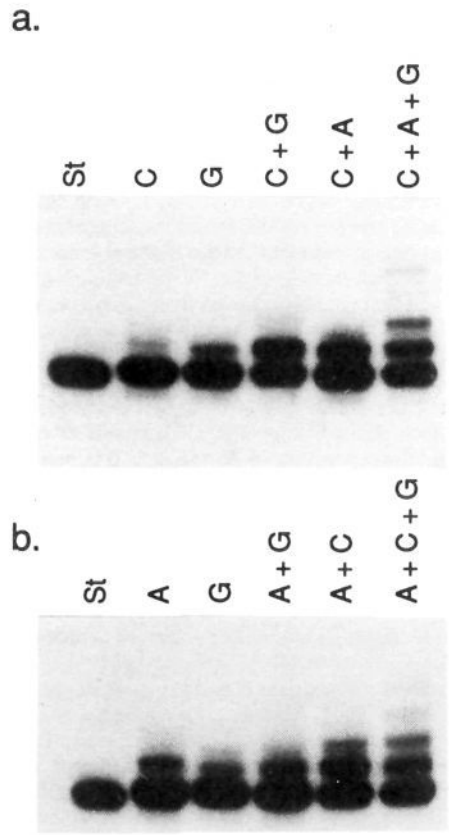


Figure 8. Autoradiogram of the gel electrophoretogram of the products from the template reaction on (a) substrate S_{10} and (b) substrate S_{11}, analyzed after the reaction time of 2 weeks.

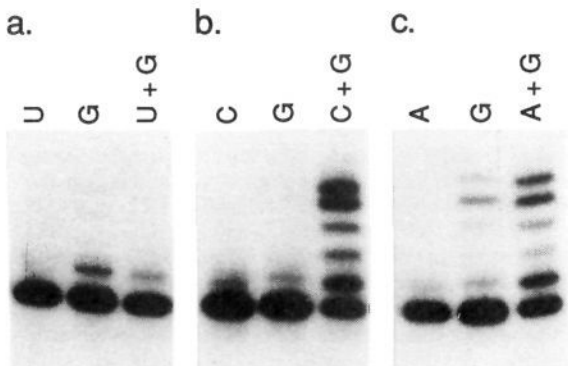


Figure 9. Autoradiogram of the gel electrophoretogram of the products from the template reactions on (a) substrate S_{12}, (b) substrate S_{13}, and (c) substrate S_{14}, analyzed after a reaction time of 2 weeks.

half the efficiency of incorporation of G. Surprisingly, we found that the number of long runs of G formed in the product when only 2-MeImpG was present in the reaction mixture was substantially greater than the number of long runs of C residues in the template. We attributed this either to the association of G with U by wobble pairing or to the "looping out" of U residues in the template.

Our new results are consistent with these earlier findings since the yield of products from the incubation of substrates S_1 and S_8 with 2-MeImpG in the absence of 2-MeImpA are substantial. In both cases, the total yield of products is almost the same whether 2-MeImpA is present in or absent from the reaction mixture. A detailed examination of the products makes possible a distinction between wobble pairing and looping out as the explanation of the anomalous incorporation of G residues when S_1 or S_8 is incubated with G alone. On the one hand, looping out of two U residues in template S_8 would lead to the rapid incorporation of at most

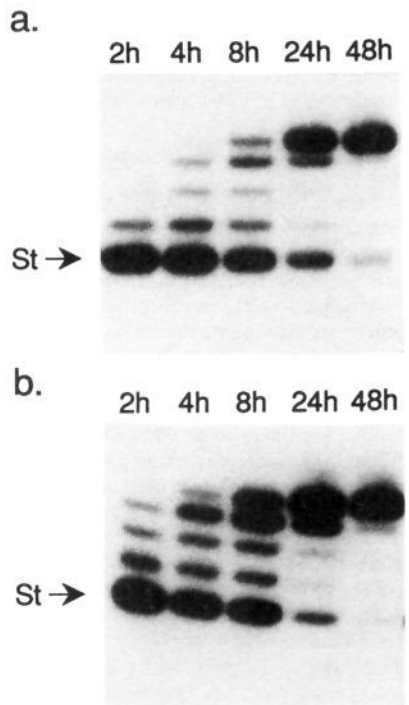


Figure 10. Autoradiogram of the gel electrophoretogram showing (a) time course of the reaction of substrate S_{15} with 2-MeImpG and (b) time course of the reaction of substrate S_{16} with 2-MeImpG.

2 G residues, and the slower incorporation of one further G residue. Wobble pairing, on the other hand, would lead to the rapid incorporation of up to 4 G residues and to the slower incorporation of a fifth. Figure 2 clearly indicates that, when S_1 is the substrate, up to 4 G residues are incorporated. Similar results are obtained with substrate S_8. This shows that it is wobble pairing rather than looping out that permits sequences containing T residues to act as effective templates when incubated with 2-MeImpG alone. Presumably, the same is true for the reactions on random copolymers of rU and rC that we reported earlier.

The copolymerization of 2-MeImpU and 2-MeImpG on random copolymer templates containing riboC and riboA is less efficient than reactions on equivalent templates containing C and G or C and U residues.[3,5,6] Furthermore, the incorporation of U is much less efficient than the incorporation of G. We concluded from a detailed analysis of the spectrum of products that "runs of consecutive adenine residues would not be suitable for use in a chemical self-replicating system". This conclusion is confirmed by the results reported here, particularly by the difficulty of adding a U residue to S_{12}.

Efficiency of Incorporation. The results obtained in this and two previous papers[1,2] show clearly that the efficiency with which a template strand facilitates the synthesis of its complement is strongly sequence-dependent. The incorporation of a G residue at the end of an oligo(G) primer on a poly(C) template, the reaction discovered first, is very efficient and regiospecific. The incorporation of an isolated C, A, or T residue into a product oligomer on a template of the type –CCNCC–, where N is G, A, or T, involves two steps. The first is the incorporation of C, T, or A at the end of a primer terminated by G; the second is the incorporation of a G residue at the end of a primer terminated by C, T, or A. Our results confirm earlier studies.[3-6] They show that each of these incorporation reactions occurs efficiently, although the rates of addition vary from case to case. Thus copying of sequences consisting mainly of C residues but including isolated

(6) Joyce, G. F.; Orgel, L. E. *J. Mol. Biol.* **1988**, *202*, 677–681.
(7) Kuchino, Y.; Nishimura, S. In *Methods in Enzymology. RNA Processing. Part A. General Methods*; Dahlberg, J. E., Abelson, J. N., Eds.; Academic Press, Inc.: New York, 1989; Vol. 180.

G, A, or T residues occurs efficiently under our conditions.

Copying of a template containing two consecutive residues different from C, involves three successive reactions. Copying of the TT sequence in –CCTTCC, for example, involves adding A to a terminal G, then A to a terminal A, and finally G to a terminal A. We find that the sequence GG is copied with high efficiency;[2] the sequences TT, GT, and TG, with intermediate efficiency; and the sequences AA, AT, TA, AG, and GA, with very low efficiency.

It is impractical to study systematically all template sequences containing three successive residues chosen from G, A, and T. We have shown that sequences of several consecutive G residues are copied efficiently. A reasonable inference from the results described above is that no other sequences of three consecutive non-C residues within an environment of C residues (–$CCN_1N_2N_3CC$–) is copied efficiently, and only sequences containing two G residues and one T residue are likely to be copied at all.

Regiospecificity. The incorporation of a U residue opposite an isolated A residue in the template is clearly highly regiospecific since the products of elongation of the substrate S_2 in the presence of a mixture of 2-MeImpU and 2-MeImpG are almost completely degraded by ribonuclease T1 and ribonuclease A to give the 3′-phosphate of the substrate and the 3′-phosphate of the U adduct of the substrate, respectively.

The results obtained in the enzymatic degradation of the products by incubating substrate S_1 with a mixture of 2-MeImpA and 2-MeImpG are less clearcut. Degradation with ribonuclease T1 establishes that the A residue that is first incorporated is 3′-5′-linked to the primer. Ribonuclease H cleaves the products after the A residue, but degradation is incomplete. As much as 20% of the products survived. Our results are hard to interpret—incomplete degradation of the products by ribonuclease H could be attributed either to the inability of the enzyme to degrade completely hairpins with very short DNA–RNA duplexes or to the presence of a substantial proportion of 2′-5′-linkages. On the basis of the ribonuclease H results we can only be sure that at least 80% of the A–G bonds are 3′-5′-linked. It is clear that ribonuclease U and ribonuclease PhyM do not hydrolyze stem-loop products formed in our reactions, even if they are 3′-5′-linked. This follows from the failure of these enzymes to degrade the 80% of the products that are cut by ribonuclease H. Thus these two ribonucleases do not give useful information about linkage isomerism in our system.

The results established in this and the previous papers of the series, taken together, show that all of the efficient template-directed reactions that we have studied give predominantly 3′-5′-linked products with the possible exception of the addition of G to a primer terminated by riboA. In that case we can only say that at least 80% of the bonds are 3′-5′-linked.

The regiospecificity of the misincorporation of G opposite a T or an A residue in the template was also studied. Incorporation of G opposite T in the template, presumably through wobble pairing, gives 70% of 3′-5′- and 30% of 2′-5′-linked bonds. Incorporation of G opposite an A residue gives a majority of 2′-5′-linked products and only about 25% of the 3′-5′-linked products. In a previous paper we showed that when G is incorporated opposite to G in the template, 2′-5′-linked and 3′-5′-linked products are obtained in approximately equal amounts.[2]

The phosphoro(2-methyl)imidazolide of G was chosen as an activated G derivative because it oligomerized efficiently to give almost exclusively 3′-5′-linked products. We believe that the regiospecificity of the reactions of the other nucleoside phosphoro(2-methyl)imidazolides, when they are attached by Watson–Crick pairing to the template, is a consequence of the geometrical equivalence of the Watson–Crick base pairs. If the C:G base pair places the 5′-phosphate of the incoming base in a position to react with the 3′-hydroxyl group at the end of the primer, so also will G:C, A:T, and T:A base pairs. The geometry of non-standard base pairs is different from that of the Watson–Crick base pairs. Consequently, the positioning of the 5′-phosphate group of an incoming base relative to the 2′- and 3′-hydroxyl groups of the primer when mispairing occurs is significantly different from that achieved with Watson–Crick base pairing. This may lead to the formation of a large proportion of 2′-5′-internucleotide bonds.

Misincorporation. The use of hairpin substrates to study template-directed reactions is not always suited to the determination of the fidelity of synthesis, since electrophoresis does not distinguish readily between the anticipated template-directed products and other oligomers of the same length but with different sequences. Thus our results are mainly concerned with the misincorporation of an incorrect base in the absence of competition from the correct base.

It is clear that G residues are easily misincorporated on a variety of templates. The misincorporation of G on substrate S_1 is very significant. After short incubation times (1 day) misincorporation of G opposite T is already detectable. After 7 days 50% of the substrate has reacted. It is striking that the product terminated by a misincorporated G opposite a T is rapidly extended to longer products. Incorporation of A opposite T is much faster than misincorporation of G, but elongation of the first-formed product by incorporation of further G residues is faster after a misincorporated G residue than after a correctly incorporated A residue (Figure 2b). The incorporation of G residues opposite A in the template is less efficient than incorporation opposite T (Figures 2a and 4a). Incorporation of G opposite A or T, however, is significantly faster than misincorporation of G opposite an isolated G residue in the template.[2]

Consecutive misincorporation of G residues opposite most pairs of consecutive non-C residues does not occur to any significant extent. However, there is one important exception. Misincorporation of two G residues opposite a TT sequence in the template in the presence of 2-MeImpG alone occurs to almost the same extent as incorporation in the presence of 2-MeImpA and 2-MeImpG together (Figure 6).

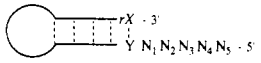
The ability of G alone to "copy" the template strands in S_1 and S_8 raises the question of fidelity. When A and G are both present in equimolar amounts, which is incorporated opposite T or TT? In the case of substrate S_1, it is clear that correct incorporation of A is the dominant reaction, since the final products obtained when 2-MeImpG and 2-MeImpA are present have very different mobilities from the mobilities of the products obtained with 2-MeImpG alone (Figure 2). Unfortunately, the mobilities of the products obtained from substrate S_8 with 2-MeImpG alone or 2-MeImpG and 2-MeImpA are similar (Figure 6). A careful examination of the original gels suggest that in this case also the predominant reaction when 2-MeImpG and 2-MeImpA are present involves the incorporation of two A residues opposite the TT sequence rather than the misincorporation of G residues. This result was confirmed by examining the digestion of the products with ribonuclease T1 (Figure 7).

In summary, misincorporation of G residues in template-directed reactions opposite non-C residues is often detectable. It is particularly significant opposite a T residue or a TT sequence, presumably because of the stability of the T–G wobble pair. Our experiments do not enable us to determine quantitatively the fidelity of the reactions, for example, the extent of misincorporation of G opposite T in the presence of A. Qualitatively, misincorporation seems a minor side reaction.

Effect of Downstream Bases. In an earlier paper[2] we reported that the incorporation of C opposite G residues in the template strand depends on the presence of the phosphoro(2-methyl)-imidazolide of the base that is complementary to the next downstream base in the template. This is also true for the incorporation of U and A residues. The incorporation of A opposite T on substrate S_1 is inefficient unless 2-MeImpG is present in the reaction mixture. Similarly, the incorporation of U opposite A on substrate S_2 is dependent on the presence of G. Similar effects were seen with templates S_{13} and S_{14}. In these reactions, addition of A or C to a primer terminated by U depended on the availability of 2-MeImpG to pair with a downstream C base on the template strand (Figure 9b,c).

General Implications. The results reported in this paper and in the two earlier papers of the series define, for the first time,

Table I. Approximate Half-Times for Various Template Reactions Studied in This and Previous Papers[1,2]



rX	N_1'	$N_1N_2N_3N_4N_5$	$t_{1/2}$ (h)
G	G	C C C C C	3
A	G	C C C C C	5
C	G	C C C C C	6
U	G	C C C C C	8
G	C	G C C C C	18
G	A	T C C C C	24
G	U	A C C C C	96
G	C	G G C C C	24[a]
			48[b]
C	C	G G C C C	24

[a] In the presence of 2-MeImpC only. [b] In the presence of 2-MeImpC and 2-MeImpG.

the family of C-rich oligonucleotides that can be copied efficiently. We have now shown that runs of G residues may be present, but that A and T residues must be isolated from each other and from G residues. The sequences TT, GT, and TG are partial barriers to copying, and the sequences AT, TA, AA, GA, and AG are almost total barriers.

The half-times of some of the reactions that we have studied are collected in Table I. They illustrate the relative rates of the most successful copying reactions. The incorporation of G after any of the four (isolated) bases is rapid. The incorporation of the three other bases after G is substantially slower. The slow rate of incorporation of U after G is already a substantial obstacle to observing efficient template reactions involving incorporation of U in our system. After 4 days the hydrolysis of the activated nucleotides is significant, so the incorporation of several isolated U residues in a product could only be accomplished by replenishing the supply of 2-MeImpU from time to time.

There are other obstacles to copying, in addition to those revealed by the present studies. Intramolecular self-structure is one such obstacle. It is likely to be particularly important for copolymers that include large proportions of both C and G. Another obstacle is the aggregation of G-rich copolymers via the formation of tetrahelical segments.

A minimal self-replicating system consists of two complementary sequences each of which is capable of facilitating the synthesis of the other. Is such a system possible within the framework of our model system? Clearly a copolymer of C and G cannot qualify. If the two complements contain roughly equal amounts of C and G they are disqualified because they give rise to extensive intramolecular self-structure. If one polymer is rich in C, the other must be rich in G and so is disqualified because it forms a stable intermolecular tetrahelical self-structure.

Can the introduction of isolated A and T residues save the situation? This seems unlikely, but one cannot be sure without more experimental information on the efficiency of template-directed synthesis on oligoG sequences that contain an isolated A or T residue, for example, -GGAGG- and related sequences. We are accumulating the information needed to answer this question.

Experimental Section

Materials. The sources of all reagents, and all experimental procedures, were the same as those reported in earlier papers of the series.[1,2]

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

Registry No. S_1, 143214-82-8; S_2, 143214-80-6; S_3, 143214-86-2; S_4, 143214-87-3; S_5, 143214-83-9: S_6, 143214-84-0; S_7, 143214-85-1; S_8, 143214-90-8; S_9, 143214-91-9; S_{10}, 143214-88-4; S_{11}, 143214-89-5; S_{12}, 143214-76-0: S_{13}, 143214-77-1; S_{14}, 143214-78-2; S_{15}, 143214-75-9; S_{16}, 143214-81-7; 2-MeImpC, 85179-51-7; 2-MeImpG, 80242-42-8; 2-MeImpA. 31008-65-8; 2-MeImpU, 85179-50-6.

Peptide-Titanium Complex as Catalyst for Asymmetric Addition of Hydrogen Cyanide to Aldehyde

Hideaki Nitta, Donghai Yu, Masanobu Kudo, Atsunori Mori, and Shohei Inoue*

Contribution from the Department of Synthetic Chemistry, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan. *Received January 27, 1992*

Abstract: The complex of titanium ethoxide and an acyclic dipeptide ester whose terminal amino group is modified to a salicylal-type Schiff base catalyzes the asymmetric addition of hydrogen cyanide to aldehydes with high enantioselectivity. In the reaction of benzaldehyde and hydrogen cyanide, (*R*)-mandelonitrile is obtained with an enantiomeric excess of 90% when *N*-((2-hydroxy-1-naphthyl)methylene)-(*S*)-valyl-(*S*)-tryptophan methyl ester is employed. In place of the dipeptide, the amide derivatives of an amino acid modified by substituted salicylaldehyde, such as *N*-(3,5-dibromosalicylidene)-(*S*)-valine piperidide, exhibit an entirely opposite stereoselectivity to yield *S*-cyanohydrins with optical purities up to 97% ee. This novel peptide-titanium complex, therefore, enables us to afford optically active cyanohydrins of both absolute configurations by using natural *S*-amino acids as chiral auxiliaries.

Introduction

The development of a method to synthesize optically active compounds has always received much attention in various areas of organic and biological chemistry.[1] In recent years the design of chiral metal complexes as catalysts for asymmetric organic reactions has been widely studied.[2] In particular, highly stereoselective procedures have been explored in various organic reactions using the ligands of C_2-symmetry, as represented by the

(1) For a review, see: *Asymmetric Synthesis*; Morrison, J. D., Ed.; Academic Press: New York, 1985; Vol. 5.

(2) Noyori, R.; Kitamura, M. In *Modern Synthetic Methods*; Scheffold, R., Ed.; Springer: Berlin, 1989; Vol. 5, pp 115-198.

0002-7863/92/1514-7969$03.00/0