Nucleic acid duplexes incorporating a dissociable covalent base pair

Kui Gao and Leslie E. Orgel[†]

The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037

Contributed by Leslie Orgel, October 21, 1999

We have used molecular modeling techniques to design a dissociable covalently bonded base pair that can replace a Watson-Crick base pair in a nucleic acid with minimal distortion of the structure of the double helix. We introduced this base pair into a potential precursor of a nucleic acid double helix by chemical synthesis and have demonstrated efficient nonenzymatic template-directed ligation of the free hydroxyl groups of the base pair with appropriate short oligonucleotides. The nonenzymatic ligation reactions, which are characteristic of base paired nucleic acid structures, are abolished when the covalent base pair is reduced and becomes noncoplanar. This suggests that the covalent base pair linking the two strands in the duplex is compatible with a minimally distorted nucleic acid double-helical structure.

nucleic acid double helix | covalent base pair | nonenzymatic template-directed ligation

The importance of the nucleic acids in biology and biotechnology has been the motivation for many studies of oligonucleotide analogues. The sugar-phosphate backbones of RNA and DNA have been replaced by a variety of alternative backbones (1–4), a number of alternative or additional base pairs have been described (5, 6), and the possibility of using shape analogues of the bases has been explored (7, 8). Cross-linking of nucleic acid strands has also been discussed in many publications (9–14), but, as far as we are aware, reversible cross-linking leading to a "covalent base pair" without distortion of the structure of the double helix has not been described.

An ideal covalent base pair would mimic closely the geometry of a standard base pair. In particular, it should be approximately planar, and the glycosidic angles and the distance between the glycosidic carbon atoms should be as close as possible to those in a standard nucleic acid double helix. In the context of molecular replication and prebiotic chemistry, it is also important that the covalent base pair can be formed and dissociated reversibly. Devadas and Leonard (15) have synthesized glycosidic derivatives of fused heterocyclic systems that resemble pairs of nucleosides in a DNA double helix—for example, structure A. These molecules are coplanar and have geometry fairly close to that of a standard base pair, but they cannot form and dissociate reversibly. Another interesting base pair (B) has been described by Cowart and Benkovic (16) and by Webb and Matteucci (17). This adduct, although not coplanar, can adopt a coplanar structure in the context of a nucleic acid double helix, but it cannot dissociate reversibly.



Here we describe the design and synthesis of a covalent base pair that can replace a normal Watson-Crick base pair in nucleic acid with a minimum of structural perturbation. We have incorporated it into a DNA double helix by nonenzymatic aqueous solution chemistry and have shown that the helix can be formed and dissociated reversibly.

Materials and Methods

De Novo Design of Covalently Bonded Base Pair. We decided to study molecules in which two aromatic systems are joined by a C=N-N linker:



Systems of this kind are conjugated and, therefore, approximately coplanar and can be synthesized fairly easily from appropriate hydrazines and aldehydes. The hydrazone bond can be dissociated by treatment with excess hydrazine.

We investigated isolated five- and six-membered rings and fused bicyclic ring systems as possible aromatic components of the covalent base pair. For each possible structure, D, α_1 , and α_2 were calculated by using standard geometry parameters. Structures with D in the range 10.8 \pm 3.0 Å, and α_1 and α_2 in the range $50^{\circ} \pm 30^{\circ}$ were selected for further consideration. In each of the selected structures, the simple aromatic rings were systematically replaced by possible pyrimidine and related heterocyclic systems. These structures were then optimized with molecular mechanics, CFF91 force field (18), using Discover module in INSIGHTII software (Biosym Technologies, San Diego). In a second screen, the structures with D in the range 10.8 \pm 1.5 Å and α_1 and α_2 in the range $50^{\circ} \pm 15^{\circ}$ were selected for a final refinement by using modified neglect of differential overlap semiempirical quantum chemistry (19). The bestfit structure I was used to suggest a closely related structure II for experimental investigation. We chose II because it would be expected to have a geometry similar to that of I and can be synthesized more easily.

Although II has several degrees of freedom because of its flexibility, the conformation shown here is always found to be the most stable by systematic search and optimization at PM3 semiempirical quantum chemistry level (20). The geometry of the proposed covalent base pair, calculated by using PM3

BIOCHEMISTR

Abbreviation: EDC, 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride. [†]To whom reprint requests should be addressed. E-mail: orgel@salk.edu.

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



semiempirical quantum chemistry, closely resembles that of a standard base pair (Fig. 1). The estimated values of D, and α_1 and α_2 are 10.5 Å and 48.5°/57.1°, very close to those for the DNA double helix (10.8 Å and 52°). Dynamic simulation and optimization suggests that a double helix containing this covalent base pair should be stable.

Overview of the Synthesis of a Nucleic Acid Incorporating a Covalent Base Pair. The structure of a single-stranded nucleic acid modified to include a covalently bonded base pair is illustrated in Fig. 2. The base pair in adduct Template—F is formed by condensing a nucleoside analogue F containing an aldehyde group (V) with a N^4 -aminocytidine residue in the target single-stranded nucleic acid. Outlines of the synthesis of the dialdehyde and details of the conversion of a cytidine residue in the target strand to a N^4 -aminocytidine residue are given below.

Synthesis of Compound V. The synthesis of compound V is illustrated in Fig. 2. 4-Methyl-1-(2,3,5-tri-*O*-benzoyl- β -D-ribo-furanosyl)-2-pyrimidinone (III) is synthesized efficiently by the reaction of silylated 4-methyl-2-pyrimidinone with 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl- β -D-ribofuranose. The methyl group of the pyrimidinone is then converted to an α -diformylmethyl group by the Vilsmeier reaction (21). The product is deprotected to give the nucleoside analogue (compound V, designated as F) as a light yellow powder in 45% overall yield. The product was characterized by NMR, mass spectroscopy, and UV spectroscopy.

Incorporating a Covalent Base Pair into a Single-Strand DNA. All standard oligodeoxynucleotides were synthesized by using an Applied Biosystems DNA Synthesizer, were deprotected in concentrated ammonia at 55°C, and were concentrated by using an OP 120 OligoPrep instrument (Savant). They were then purified on a 20% denaturing polyacrylamide gel (PAGE in 1× TBE buffer) and were eluted with TE buffer (0.01 M Tris·HCl/0.001 M EDTA, pH 8). A NENSORB 20 column (DuPont) was used to desalt the oligomers. 5'-³²P-labeled oli-

gomers were prepared by using T4 polynucleotide kinase (New England Biolabs) and $[\gamma^{-32}P]ATP$ (Amersham Pharmacia) (22). Enzymatic synthesis of 5'-³²P-labeled oligonucleotide 3',5'-diphosphates was carried out as described (23). Oligomers used in this work and various derivatives of them are listed in Table 1. Oligomers L₃ and L₅ are complementary to the 3' and 5' segments, respectively, of template Tem.

In Fig. 3, we illustrate our general approach to the incorporation of a covalent base pair into a presynthesized oligodeoxynucleotide. The incorporation protocols were developed by using a short oligodeoxynucleotide T₃CT₃ as substrate. The conversion of a cytidine residue to N^4 -aminocytidine was achieved by using a slight modification of a published procedure (24). A reaction solution containing the oligonucleotide (1 nM to 1 mM), 4 M hydrazine, 0.3 M sodium bisulfite, and 0.1 M phosphate buffer at pH 5.0 (the pH was adjusted with concentrated HCl) in a 50- μ l volume, was incubated at 50°C for 3 hours. The reaction was stopped by dilution with 500 μ l TET buffer (0.1 M Tris·HCl/0.001 M EDTA/0.01 M triethylamine, pH 8) and was desalted on a NENSORB column. The N⁴-aminocytidine residue is unstable and decomposes slowly, even at -20° C. The freshly formed oligomer containing a N⁴-aminocytidine residue must be condensed immediately with the aldehyde.

Oligomer $T_3C'T_3$ (0.5 nM to 5 mM), prepared as described above, was incubated with a 0.01 M solution of the nucleoside analogue F in 0.05 M phosphate buffer (pH 5.0) at room temperature for 36 hours (reaction volume 20 μ l). The reaction was stopped by mixing with gel loading buffer (a mixture of 900 μ l of deionized formamide, 25 μ l of 2% xylene cyanol, 25 μ l of 2% bromophenol blue, and 100 μ l of 10× TBE buffer) and 10 M urea in the ratio of 2:1:2. The hydrazone derivative was purified by gel electrophoresis. The identity of the adduct, $T_3(C=F)T_3$, was confirmed by matrix-assisted laser desorption ionization (MALDI) mass spectrometry: $C_{81}H_{104}N_{18}O_{52}P_6$ (calculated 2,346.45, found 2,346). The unmodified oligomer T_3CT_3 did not react with analogue F under identical conditions.

The protocols developed by using T_3CT_3 as described above were then used to synthesize a longer adduct, Tem=F, and the ³²P-labeled adduct, p*Tem=F. The hydrazone bond in adduct p*Tem=F could be reduced to a C—N single bond by treatment with sodium borohydride. The adduct p*Tem=F (10,000 cpm in 10 μ l of reaction volume) was treated with 0.1 M NaBH₄ and 0.001 M EDTA in 0.025 M phosphate buffer (pH 6.8) at room temperature for 45 minutes. The reduction reaction was terminated by 20-fold dilution with TET buffer at 0°C. The product p*Tem—F was desalted immediately on a NENSORB column and was used in the ligation reactions without further purification.



Fig. 1. (A) The structure of the proposed covalently bonded base pair. (B) The structure of a standard base pair. (C) Superposition of the covalently bonded base pair (darker) with a Watson-Crick base pair (lighter).



Fig. 2. Scheme for the synthesis of an oligonucleotide incorporating a covalent base pair (the hydrazone bond is indicated by the arrow). i, SnCl₄/HMDS/TCS/CH₃CN. ii, NaHCO₃. iii, POCl₃/DMF. iv, H₂O. v, MeONa/MeOH. vi, NH₂NH₂/NaHSO₃/phosphate buffer (pH 5). vii, F/phosphate buffer (pH 5). HMDS, hexamethyldisilazane; TCS, trimethylchlorosilane; Template, oligonucleotide template; Template', oligonucleotide incorporating a N⁴-aminocytidine residue; Template==F, template covalently linked to the nucleoside analogue F by means of a hydrazone bond. Either T₃CT₃ or Tem (Table 1) can be used as the Template.

Ligation Reactions. The labeled oligomer p^*L_{3p} (1 × 10³ to 8 × 10⁵ cpm in 10 μ l of reaction volume) was treated with 0.11 M 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDC) in 0.1 M imidazole·HCl solution (pH 6.0) at room temperature for 1 hour to give the phosphorimidazolide of p^*L_{3p} , Im(p^*L_{3p}). An aliquot of the solution containing Im(p^*L_{3p}) (4 μ l), without further purification, was then added to the ligation reaction mixture (total reaction volume 40 μ l),

Table 1. The oligodeoxynucleotides and their derivatives used in these experiments

Oligodeoxynucleotides and their derivatives

derivatives	Symbol
5'TTTCTTT3'	T₃CT₃
5′TTT-C-TTT3′ NH2	$T_3C'T_3$
5′TTT-C-TTT3′ F	T ₃ (C=F)T ₃
5'TGTGTGTTGTG-C-GTTGGTGGTTTGGT3'	Tem
5'TGTGTGTTGTG-C-GTTGGTGGTTTGGT3' F	Tem=F
5'TGTGTGTTGTG-C-GTTGGTGGTTTGGT3' F	Tem=F
5'CACAACACA3'	L ₅
5'GCACAACACACA3'	GL ₅
5'CAAACCACCAAC3'	L ₃
5'pCAAACCACCAACpGCACAACACACA3'	pL₃pGL₅
5'TGTGTGTTGTG-C-GTTGGTGGTTTGGT3'	Tem=L₅pF
5'TGTGTGTGTGTG-C-GTTGGTGGTTTGGT3'	Tem=FpL₃p
3′FpCAACCACCAAACp5′	
5'TGTGTGTTGTG-C-GTTGGTGGTTTGGT3'	Town I what w
3΄ΑCACACAACACpFpCAACCACCAAACp5′	Tem=L5pFpL3p
5'pCAAACCACCAACpF3'	FpL₃p
5'FpCACAACACA3'	L₅pF

which contained the analogue Tem=F(Imp*L₅:Tem=F \approx 1:1.5), 0.1 M KCl, 0.05 M MgCl₂, and 0.1 M Tris·HCl (pH 7.4) (see Fig. 3, a). The solution was first heated to 70°C, was allowed to cool to 25°C in 45 minutes, and then was left at room temperature for 4 to 7 days. The reaction was terminated by adding 0.5 M EDTA (10 μ l). The reaction products were desalted and analyzed by gel electrophoresis. Experiments in which EDC or MgCl₂ was omitted were used as controls. A standard ligation experiment with unmodified oligonucleotides GL_5 and $Im(p^*L_3p)$ aligned on the unmodified template Tem was carried out under identical conditions to permit comparison between a standard ligation reaction and ligation to a covalent base pair. A ligation reaction between Tem=F and $p*L_5$ (Fig. 3, b) was carried out in the same way. In addition, we purified the product Tem=FpL3p* of the first ligation reaction by gel electrophoresis and used it in a second ligation reaction with p*L₅ (Fig. 3, c).

The ligation products, Tem= L_5p^*F and Tem= FpL_3p^* (10,000 cpm), were dissociated by treating them with 0.1 M hydrazine at room temperature for 10 hours (5- μ l reaction volume). A parallel reaction omitting hydrazine was used as a control.

The orientations of the products of the ligation reactions were determined by dephosphorylation with alkaline phosphatase (Boehringer Mannheim) (22). Ligation products, Tem= L_5p^*F and Tem= FpL_3p^* , were first dissociated with excess hydrazine, and the product oligomers, L_5p^*F and FpL_3p^* , which contain the nucleotide analogue F, were purified by gel electrophoresis. L_5p^*F and FpL_3p^* (1,000 cpm) were then treated with alkaline phosphatase (0.1 unit) in 50 mM Tris·HCl, and 0.1 mM EDTA in 10 μ l (pH 8.5). The reaction products were analyzed by PAGE.

Results and Discussion

We used computer modeling techniques to design *de novo* a dissociable covalent base pair II that can be formed by reaction between two nucleoside analogues, the dialdehyde F and N⁴-aminocytidine. We synthesized F in 45% overall yield from 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl- β -D-ribofuranose and 4-methyl-2-pyrimidinone. A N⁴-aminocytidine residue in an oligonucleotide was obtained from a cytidine residue by treating the

BIOCHEMISTRY



Fig. 3. General approach to the synthesis of nucleic acid double helices incorporating a covalently bonded base pair by nonenzymatic template-directed ligation. The oligonucleotide segments are those listed in Table 1.

oligomer with hydrazine in the presence of bisulfite as a catalyst. The nucleoside F reacts readily with the N^4 -aminocytidine residues in a single-stranded nucleic acid to form a hydrazone derivative (covalent base pair). Treatment of 5'-³²P-labeled oligomer T₃CT₃ with hydrazine under conditions that convert cytidine to N^4 -aminocytidine does not change its mobility (Fig. 4, lane 2). However, subsequent treatment of the N^4 -aminocytidine-containing oligomer with the nucleoside analogue F results in the formation of the adduct p*T₃(C=F)T₃, which has a reduced electrophoretic mobility (Fig. 4, lane 3). The yield is ≈85%.

Nonenzymatic template-directed ligation reactions can be used to join two oligonucleotides aligned on a complementary template. If the structure of the covalent base pair resembles the structure of a standard base pair closely enough, the F residue that is covalently linked to the template should be able to replace the complex of the template with one of the complementary oligomers in this reaction. If the presumed covalent base pair differs greatly in structure from a standard base pair, it would not be expected to ligate efficiently. The oligomer Tem is a 26-mer with a single C residue at position 12. Nonenzymatic templatedirected ligation between the nucleoside analogue F covalently linked to the unique C residue of the template and complementary oligodeoxynucleotide sequences was used to test whether the covalent base pair disturbs the double helical structure of DNA. Fig. 3 depicts the general design of the experiments.

We preactivated the p^*L_3p by treating it with EDC and imidazole HCl at pH 6. The yield of the 5'-³²P-labeled phosphorimidazolide of p^*L_{3p} , Im (p^*L_{3p}) varied from 70% to 90% (data not shown). The formation of the ligation product Tem= $Fp*L_3p$ of Tem=F with $Im(p*L_3p)$ after 7 days at room temperature is shown in Fig. 5, lane 1 (\approx 44% yield). In a control reaction in which the activation step was omitted, no ligation product is observed (Fig. 5, lane 2). When Mg^{2+} is omitted from the reaction mixture, the yield of ligation product is much reduced (Fig. 5, lane 3). We carried out a parallel standard ligation reaction using ≈ 0.7 equivalent of p*L₃p and a 10-fold excess of GL₅ to guarantee that the 5'-end of the template Tem was fully occupied. The yield of the standard ligation product was only 7% (Fig. 5, lane 4). Ligation in the covalent template is clearly much more efficient ($\approx 6 \times$) than standard ligation between two oligonucleotides. Presumably, this is partly attributable to incomplete hybridization in the standard ligation reaction, but the efficiency of the reaction with the covalent base pair is nonethe less surprising. The corresponding ligation reaction involving the 5'-phosphate-terminated oligomer (p^*L_5) and Tem=F gave a similar yield (48%) of the ligation product Tem= L_5p^*F .

Next, we synthesized a nucleic acid duplex containing an internal covalent base pair. We prepared Tem= FpL_3p^* as described above and purified it by gel electrophoresis. This product was then incubated with activated p^*L_5 under the standard ligation condition. Gel electrophoresis (Fig. 5, lane 5) revealed the formation of the higher molecular weight product, Tem= $L_5p^*FpL_3p^*$. A control ligation reaction of p^*L_3p and p^*L_5 in the absence of template gave no detectable ligation product (Fig. 5, lane 8). Tem= $L_5p^*FpL_3p^*$ can also be formed by ligation of Tem= L_5p^*F with p^*L_3p , but with lower efficiency, presumably because an isolated 5'-hydroxyl group is a less good nucleophile than a cis-glycol (data not shown).

To establish that efficient ligation depends on the geometry of the covalent base pair and not only on its proximity to the activated phosphate group of its neighbor, we next attempted a ligation reaction to a covalently linked oligomer with altered structure. Reduction with NaBH₄ converts Tem—F to Tem—F in which the C—N double bond (hydrazone bond) has been reduced to a C—N single bond. After reduction, the base pair is no longer coplanar. In Fig. 5, lane 9, we show the product p*Tem—L₅p*F from a ligation reaction between 5'-³²P-labeled Tem—F and activated 5'-³²P-labeled L₅ and, in Fig. 5, lane 10, the product from the corresponding reaction with the reduced



Fig. 4. Incorporation of a covalent base pair into a DNA strand. Oligomer $p*T_3CT_3$ (lane 1) was treated with hydrazine (lane 2) and then was condensed with nucleoside analogue F (lane 3).



Fig. 5. Autoradiograms of the products of ligation reaction using the 26-mer template Tem. A compares ligation to the covalently bonded base pair to standard nonenzymatic template-directed ligation. Lanes: 1, ligation of Tem—F with p^*L_3p ; 2, control omitting EDC; 3, control omitting Mg²⁺; 4, standard ligation of GL₅ with p^*L_3p . *B* illustrates the preparation of the duplex containing an internal covalently bonded base pair (Tem—L₅ $p^*FpL_3p^*$, lane 5). Lanes: 6, ligation of Tem—F with p^*L_5 ; 7, ligation of Tem—F with p^*L_3p and the ligation product Tem—FpL₃ p^* was used to ligate to p^*L_5 shown in lane 5; 8, control experiment, in which Im(p^*L_3p) and Im p^*L_5 were incubated together in absence of template. *C* illustrates the effect of the reduction of the hydrazone bond on the yield of ligation products. Lanes: 9, ligation of L₅ with p^*Tem —F; 10, ligation of L₅ with p^*Tem —F, the reduction product of p^*Tem —F.

form p*Tem—F. Clearly, reduction of the C—N double bond results in a complex that ligates with much decreased efficiency.

The hydrazinolysis of Tem= L_5p^*F is illustrated in Fig. 6. Tem= L_5p^*F can be dissociated by treatment with excess hydrazine. As anticipated, the product Fp^*L_5 of hydrazinolysis (Fig. 6, lane 2) has an electrophoretic mobility less than that of the labeled starting material p^*L_5 (Fig. 6, lane 3).

Ligation of Tem F with p^*L_5 should produce an internally labeled phosphodiester group. In Fig. 7, lane 2, we show that the dissociated product L_5p^*F is not dephosphorylated by alkaline phosphatase, an enzyme that cleaves phosphomonoesters but not phosphodiesters, confirming that L_5p^*F does indeed contain an internally labeled phosphate group. The corresponding ligation product, FpL_3p^* , includes a labeled 5'-terminal phosphomonoester group, which is, as anticipated, removed by alkaline phosphatase (Fig. 7, lane 4). These experiments confirm that the complementary pairing of the short oligomers on the template is essential for efficient ligation reaction.

The experiments described above establish that efficient ligation occurs between the covalently attached F residue and an adjacent activated oligonucleotide. We have shown that an oligodeoxynucleotide incorporating a single covalent base pair resembles an oligodeoxynucleotide-primer complex. Nonenzymatic ligation to an appropriate complementary oligonucleotide is more efficient than the corresponding ligation reaction between pairs of oligonucleotides. In an important series of controls, we find that reduction of the C—N double bond in the covalent base pair, which abolishes its planarity, greatly reduces the efficiency of the ligation reactions. We believe that these results establish that covalent cross-linking occurs to give a



BIOCHEMISTRY



Fig. 7. Dephosphorylation with alkaline phosphatase (AP) of products obtained by dissociating duplexes incorporating the covalent base pair with hydrazine. Lanes: 1, L₅p*F; 2, attempted dephosphorylation of L₅p*F with alkaline phosphatase; 3, FpL₃p*; 4, dephosphorylation of FpL₃p* with alkaline phosphatase.

double helix, the structure of which is little different from the standard nucleic acid duplex structure, although, without structural data, we cannot quantitate the extent of distortion.

Minimally distorted, cross-linked nucleic acid duplexes should prove useful in exploring the mechanisms of enzymes that act on double-helical nucleic acids because they would help to establish the extent to which strand separation is an essential feature of the reaction mechanisms. The inclusion of one or more covalent base pairs should also permit that stabilization of normally inaccessible secondary and tertiary structure. The experimental

- 1. Eschenmoser, A. & Leowenthal, E. (1992) Chem. Soc. Rev. 21, 1-16.
- Joyce, G. F. & Orgel, L. E. (1999) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 49–77.
- Bolli, M., Litten, J. C., Schutz, R. & Leumann, C. (1996) J. Chem. Biol. 3, 197–206.
- Koshkin, A. A., Nielsen, P., Meldgaard, M., Rajwanshi, V. K., Singh, S. K. & Wengel, J. (1998) J. Am. Chem. Soc. 120, 13252–13253.
- Lutz, M. J., Held, H. A., Hottiger, M., Hubscher, U. & Benner, S. A. (1996) Nucleic Acids Res. 24, 1308–1313.
- Strobel, S. A., Cech, T. R., Usman, N. & Beigelman, L. (1994) *Biochemistry* 33, 13824–13835.
- 7. Morales, J. C. & Kool, E. T. (1998) Nat. Struct. Biol. 5, 950-954.
- 8. Guckian, K. M., Krugh, T. T. & Kool, E. T. (1998) Nat. Struct. Biol. 5, 954-959.
- Knorre, D. G. & Vlassov, V. V. (1985) Prog. Nucleic Acid Res. Mol. Biol. 32, 291–320.
- 10. Webb, T. R. & Matteucci, M. D. (1986) Nucleic Acids Res. 14, 7661-7664.
- Dong, Q., Barsky, D., Colvin, M. E., Melius, C. F., Ludeman, S. M., Moravek, J. F., Colvin, O. M., Bigner, D. D., Modrich, P. & Friedman, H. S. (1995) *Proc. Natl. Acad. Sci. USA* 92, 12170–12174.

results presented have established that it is possible to form covalent base pairs in a fairly simple way. This would justify the considerable effort that would be needed to develop protected nucleoside analogues for the routine solid-phase synthesis of cross-linkable oligomers. The present covalent base pair would not be useful in antisense applications because both the antisense oligomer and its target would need to be modified. However, the possibility of developing base-analogues that cross link reversibly to a normal base to generate a covalent base pair needs to be explored.

The covalent base pair that we have synthesized is not of direct relevance to prebiotic chemistry because there are no plausible prebiotic syntheses of the component hydrazine-containing and aldehyde-containing nucleotides. However, they do show in a more general way that a genetic system base on covalent rather than hydrogen bonding is possible (25). If life originated at high temperatures, this would be an attractive option.

We thank Aubrey R. Hill, Jr. for technical assistance, Bernice Walker for manuscript preparation, Senyon Choe for access to a workstation (Silicon Graphic, Inc.), and Gerald F. Joyce and Xiaochang Dai for assistance with mass spectroscopy. We also thank Dr. Alex Rich, Guoqiang Shi, Bin Zhou, Jian Li, Guangcheng Jiang, Shaokai Jiang, Barbara C. F. Chu, and Zhiyuan Zhang for helpful discussions. This work is supported by the National Aeronautics and Space Administration Grant NAGW-1660 and National Aeronautics and Space Administration Specialized Center of Research and Training/Exobiology Grant NAGW-2881.

- Huang, H., Zhu, L., Reid, B. R., Drobny, G. P. & Hopkins, P. B. (1995) Science 270, 1842–1845.
- 13. Coleman, R. S. & Pires, R. M. (1997) Nucleic Acids Res. 25, 4771-4777.
- Fagan, P. A., Spielmann, P., Sigurdsson, S. T., Rink, S. M., Hopkins, P. B. & Wemmer, D. E. (1996) Nucleic Acids Res. 24, 1566–1573.
- 15. Devadas, B. & Leonard, N. J. (1990) J. Am. Chem. Soc. 112, 3125-3135.
- 16. Cowart, M. & Benkovic, S. J. (1991) Biochemistry 30, 788-796.
- 17. Webb, T. R. & Matteucci, M. D. (1986) J. Am. Chem. Soc. 108, 2764-2765.
- Maple, J., Dinur, U. & Hagler, A. T. (1988) Proc. Nat. Acad. Sci. USA 85, 5350–5354.
- 19. Dewar, M. J. S. & Thiel, W. (1977) J. Am. Chem. Soc. 99, 4899-4907.
- 20. Stewart, J. P. (1989) J. Comp. Chem. 10, 209-220.
- Marson, C. M. & Giles, P. R. (1994) Synthesis Using Vilsmeier Reagents (CRC, Boca Raton, FL).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY).
- 23. Chu, B. C. F. & Orgel, L. E. (1994) Methods Mol. Biol. 26, 145-165.
- Negishi, K., Nitta, N., Yamashita, Y., Takahashi, M., Harada, C., Oohara, K., Nishizawa, M., Wataya, Y. & Hayatsu, H. (1983) *Nucleic Acid Res.* 12, 29–30.
- 25. Weber, A. L. (1989) Origins Life Evol. Biosphere 19, 179-186.