Enzymatic Recognition of the Base Pair between Isocytidine and Isoguanosine†

Christopher Y. Switzer,‡ Simon E. Moroney,§ and Steven A. Benner*

Laboratory for Organic Chemistry, ETH Zurich, CH-8092 Switzerland

Received March 2, 1993; Revised Manuscript Received June 16, 1993®

ABSTRACT: The ability of various polymerases to catalyze the template-directed formation of a base pair between isoguanine (iso-G) and isocytosine (iso-C) in duplex oligonucleotides has been investigated. A new procedure was developed for preparing derivatives of deoxyisoguanosine suitable for incorporation into DNA using an automated DNA synthesizer. T7 RNA polymerase, AMV reverse transcriptase, and the Klenow fragment of DNA polymerase all incorporated iso-G opposite iso-C in a template. T4 DNA polymerase did not. Several polymerases also incorporated iso-G opposite T, presumably through pairing with a minor tautomeric form of iso-G complementary to T. In a template, iso-G directs the incorporation of both iso-C and T when Klenow fragment is the catalyst and only U when T7 RNA polymerase is the catalyst. Further, derivatives of iso-C were found to undergo significant amounts of deamination under alkaline conditions used for base deprotection after automated oligonucleotide synthesis. Both the deamination reaction of iso-C and the ambivalent tautomeric forms of iso-G make it unlikely that the (iso-C)·(iso-G) base pair was a part of information storage molecules also containing the A·T and G·C base pairs found in primitive forms of life that emerged on planet earth several billion years ago. Nevertheless, the extra letters in the genetic alphabet can serve useful roles in a contemporary laboratory setting.

The interaction between two complementary oligonucleotide strands remains one of the most remarkable examples of molecular recognition known to chemistry, especially because several features of nucleic acids appear *a priori* to suit them poorly for this process. Because oligonucleotides are flexible unidimensional polymers, duplex formation involves substantial loss of conformational entropy. Specificity is mediated by hydrogen bonds, problematic in an aqueous environment. Heterocyclic rings bearing exocyclic heteroatomic substituents form the hydrogen bonds. Yet such heterocycles often have accessible tautomeric forms that create hydrogen-bonding ambiguities (Elgureo *et al.*, 1976), a fact that has generated much confusion in the history of oligonucleotide research (Watson & Crick, 1953a).

Despite these features, nucleic acids perform superbly for the storage and transmission of encoded information. The fidelity of recognition between nucleic acids is extremely high, especially when assisted by enzymes. Most polymerases catalyze the replication of nucleic acids with fewer than one error per million. Only with this level of precision is life as we know it possible.

We and many others [see, for example, Joyce *et al.* (1987), Hashimoto and Switzer (1992), and Eschenmoser and Dobler (1992)] have for some time been interested in determining the extent to which the structure of nucleic acids can be redesigned without destroying recognition between complementary oligonucleotides. This work is also motivated by two more specific goals: to obtain antisense nucleic acid analogs that can bind to mRNA molecules and prevent their translation and to design self-replicating oligonucleotides. In these laboratories, work has explored the chemistry of oligonucleotide analogs with altered sugars (Schneider & Benner, 1990a) and altered linking groups joining the sugars (Schneider & Benner, 1990b; Huang *et al.*, 1991).

One of the more interesting variants of natural oligonucleotides are those that contain base pairs joined by "nonstandard" hydrogen-bonding patterns (Switzer *et al.*, 1989; Piccirilli *et al.*, 1990, 1991a,b; Moroney & Piccirilli, 1991). Six nonstandard base pairs can be readily written to conform to Watson–Crick geometry (Figure 1), each joined by hydrogen bonds that are (at least formally) mutually exclusive. Therefore, it is possible, in principle, to have 12 independently replicating "letters" in the nucleoside alphabet.

One interesting nonstandard nucleoside base pair is formed between isocytosine (iso-C) and isoguanine (iso-G);[1] the latter is known as a riboside as the natural product crotonoside (Cherbuliez & Bernhard, 1932; Pettit *et al.*, 1976). Isocytosine and isoguanine can form a Watson–Crick base pair with a standard geometry but with a hydrogen-bonding pattern unlike those found in the natural base pairs A·T and G·C. On these grounds, Rich suggested three decades ago (1962) that the (iso-G)·(iso-C) base pair might have been a component of primitive nucleic acids early in the development of life. Such an additional base pair might have added diversity to the structure of ribonucleic acids in the RNA world, making them more versatile as catalysts (Benner *et al.*, 1987).

More recently, Wissler and co-workers reported preliminary data indicating an extracellular growth factor containing RNA (Wissler *et al.*, 1989); this factor may contain iso-G as a constituent. If extracellular RNA molecules indeed act as

† We are grateful for support from the Swiss Nationalfond, the Forschungskommission of the ETH Zürich, and Sandoz AG. C.Y.S. was supported by a fellowship from the National Science Foundation (USA).

* To whom correspondence should be addressed in Zurich. Tel. (international) 41 1 258 2830; Fax (international) 41 1 262 24 37.

‡ Present address: Department of Chemistry, University of California, Riverside, CA. Tel. (714) 787-7266; Fax (714) 787-4713.

§ Present address: MorphoSys GmbH, Westendstrasse 125, 8000 Munich 2, Germany. Tel. (international) 49 89 85 78 39 75; Fax (international) 49 89 78 37 95.

® Abstract published in *Advance ACS Abstracts,* September 1, 1993.

[1] Abbreviations: iso-C, isocytosine; iso-G, isoguanine; iso-CMP, isocytidine monophosphate; iso-CDP, isocytidine diphosphate; iso-CTP, isocytidine triphosphate; iso-GMP, isoguanosine monophosphate; iso-GDP, isoguanosine diphosphate; iso-GTP, isoguanosine triphosphate; NPE, *p*-2-nitrophenethyl; NMP, 1-methyl-2-pyrrolidinone; RT, room temperature.

0006-2960/93/0432-10489$04.00/0

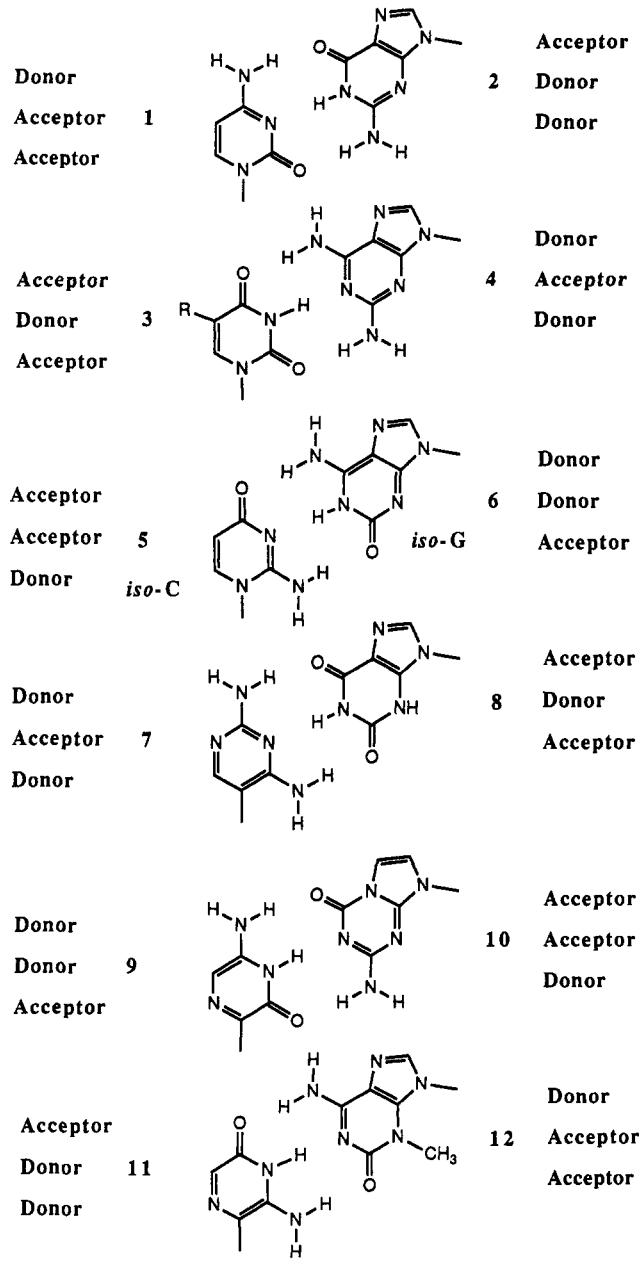

FIGURE 1: Structures of 12 bases capable of forming Watson–Crick base pairs joined by mutually exclusive hydrogen-bonding schemes incorporating three hydrogen bonds.

factors regulating growth and development in higher organisms, as suggested independently from the unusual biological activities observed with ribonucleases (Benner, 1988; Benner & Allemann, 1989), then non-standard nucleosides might be important in their function.

Still more recently, the (iso-G)·(iso-C) base pair was used to create an additional codon–anticodon set [(iso-C)AG)·(CU(iso-G)], which was used to incorporate the nonstandard amino acid iodotyrosine into a polypeptide synthesized by a ribosome through translation of a mRNA containing the additional codon (Bain *et al.*, 1992). In addition to showing the versatility of the ribosome, this work provided interesting insights into the mechanism by which translation terminates.

Future work in this area will rely heavily on methods for preparing oligonucleotides containing iso-C and iso-G, both chemical and enzymatic. We and others (Kazimierczuk *et al.*, 1991) have been exploring both aspects of the chemistry of this base pair. This paper reports efforts to explore the enzymological scope of the base pair between iso-C and iso-G, to better define conditions whereby it might be incorporated into duplex DNA and RNA by DNA and RNA polymerases.

EXPERIMENTAL PROCEDURES

Much of the synthesis of nucleosides and nucleotides containing both iso-C and iso-G followed earlier work in other laboratories. Experimental procedures are given when the procedure used here differed from that reported in earlier literature, or yielded a new compound.

(1) Ribose Derivatives of Isocytosine

Isocytidine. Riboisocytidine was prepared according to the method of Kimura *et al.* (1980).

Isocytidine Triphosphate. Tris(tributylammonium) pyrophosphate was prepared by treating pyrophosphoric acid (293 mg) in water (2.0 mL) with tributylamine (1.2 mL). The mixture was stirred vigorously at 0 °C for 10 min and then at RT for 1 h. The water was removed by lyophilization; the residue was further dried by coevaporation with pyridine (3 × 5 mL) and toluene (5 mL) and then by evacuation (0.25-mm pressure, 12 h). The tris(tributylammonium) pyrophosphate residue was dissolved in dimethylformamide (2.5 mL), and the solution was used for subsequent reactions.

Separately, isocytidine (80 mg, 0.33 mM) was suspended in trimethyl phosphate (0.822 mL, 7.0 mM) at 0 °C. The suspension was treated with $POCl_3$ (0.039 mL, 0.4 mM), and the mixture was stirred for 1.5 h. The solution of tris(tri-*n*-butylammonium) pyrophosphate in dimethylformamide from above was rapidly added, and the mixture was agitated vigorously for 1 min. A solution of Et_3NH^+ bicarbonate (1 M, pH 8.0, 2.0 mL) was then added. TLC (SiO_2, propanol:ammonia:water 11:7:2) and HPLC (100 mM Et_3NH^+ acetate pH 7.0 (A), acetonitrile (B), 100% A to 80% A/20% B over 30 min) showed a mixture of iso-C, iso-CMP, and iso-CTP. The mixture was resolved by chromatography on Sephadex A-25 (formate form, 30 mL), eluted with a gradient of Et_3NH^+ formate (0.2 to 1.5 M). The iso-CTP eluted last (overall yield 21%), was recovered by lyophilization, and was shown to be free of iso-C and iso-CMP by HPLC. ^{31}P NMR: δ −10.175, −11.0, −21.725, in water. FAB-MS: $(M-H)^-$ 482.2.

(2) 2′-Deoxyribose Derivatives of Isocytosine

2,5′-Anhydro-2′-deoxyuridine. 2,5′-Anhydro-2′-deoxyuridine was prepared according to the method of Watanabe *et al.* (1978).

2′-Deoxyisocytidine. Unprotected 2′-deoxyisocytidine was prepared according to a procedure adapted from the literature for the preparation of the ribose derivative (Kimura *et al.*, 1980). Methanol (50 mL) was saturated with dry NH_3 at 0 °C, and 2,5′-anhydro-2′-deoxyuridine (450.6 mg, 2.14 mmol) was added as a solid. The mixture was stirred at RT for 3 days. TLC (SiO_2, 20% $MeOH/CH_2Cl_2$) showed essentially complete conversion. Stirring was continued for an additional 5 days, the solvents were removed by evaporation, and deoxyisocytidine was recovered in essentially quantitative yield. FAB-MS: $(M + 1)^+$ 228.1.

N-Benzoyl-2′-deoxyisocytidine. A mixture of 2′-deoxyisocytidine (53 mg, 0.25 mmol), benzoic anhydride (59 mg, 0.25 mmol), and pyridine (2.0 mL) was stirred under N_2 overnight. The mixture was then diluted with water and extracted with CH_2Cl_2. The organic extracts were dried ($MgSO_4$) and evaporated, and the residue was chromatographed (SiO_2, 8% $MeOH/CH_2Cl_2$) to yield *N*-benzoyl-2′-deoxyisocytidine (32.7 mg, 40%; FAB-MS: $(M + H)^+$ 332.0805).

N-Benzoyl-2′-deoxy-5′-(dimethoxytrityl)isocytidine. *N*-Benzoyl-deoxyisocytidine (48 mg, 0.145 mmol) from the previous step was dried by coevaporation with pyridine (2 × 3 mL). To the residue was added dimethoxytrityl chloride (50 mg, 0.174 mmol) and (dimethylamino)pyridine (2 mg, 0.11 equiv) as solids. Pyridine (1.5 mL, distilled from CaH_2) and triethylamine (0.028 mL, 0.203 mmol, distilled from CaH_2) were added. The mixture was stirred at RT for 2 h, diluted with water (15 mL), and extracted with ether. The organic extracts were dried ($MgSO_4$) and evaporated, and the residue was chromatographed (SiO_2, 8% EtOAc/CH_2Cl_2 1:10) to yield *N*-benzoyl-2′-deoxy-5′-(dimethoxytrityl)isocytidine (63 mg, 68% FAB-MS: $(M + H)^+$ 634.3) as a yellow foam.

N²-Benzoyl-5′-(dimethoxytrityl)deoxyisocytidine Phosphoramidite. The phosphoramidite of protected 2′-deoxyisocytidine was prepared according to the procedure of Moore and Beaucage (1985). The tritylated base (31.7 mg, 0.05 mmol) was placed into a vial under N_2 (1 mL), together with 4,5-dichloroimidazole (24.5 mg, 0.179 mmol) and 1-methyl-2-pyrrolidinone (NMP, 0.3 mL). The solution was then transferred to another vial (1 mL) containing bis(diisopropylamino)methoxyphosphine (13.11 mg) under N_2. The first vial was washed with NMP (0.20 mL), and the washings were added to the second vial. The mixture was stored at RT under N_2 and used in the synthesis of the appropriate oligodeoxynucleotides.

2′-Deoxyisocytidine 3′-Monophosphate. 1,2,4-Triazole (207 mg, dried under high vacuum overnight) in dioxane (10 mL, distilled from sodium benzophenone) was treated under N_2 at RT with $POCl_3$ (0.092 mL) and then triethylamine (0.418 mL, distilled from CaH_2) over 15 min. The mixture was stirred under N_2 at RT for 40 min and then filtered under N_2 to yield a solution of tris(triazolo)phosphine oxide (0.1 M). *N*-Benzoyl-2′-deoxy-5′-(dimethoxytrityl)isocytidine (63.4 mg, 0.010 mmol) was dissolved in pyridine (0.40 mL) and treated with the solution of tris(triazolo)phosphine oxide (0.02 mmol). The mixture was stirred at RT for 3 h, diluted with water (0.4 mL), and stirred at RT overnight. NH_4OH (25%, 2.0 mL) was then added with pyridine (0.20 mL), and the mixture was heated at 60 °C for 12 h, kept at −20 °C overnight, and concentrated by evaporation under reduced pressure. The residue was treated with AcOH (80%, 1.0 mL) for 20 min and concentrated, first by evaporation under reduced pressure and then with absolute EtOH. The residue was dissolved in water (1.0 mL) and extracted with ether (3 × 1.0 mL). TLC (2-propanol:water:NH_4OH 11:2:7) showed a single predominant spot corresponding to 2′-deoxyisocytidine 3′-monophosphate (42% yield).

2′-Deoxyisocytidine Triphosphate. A suspension of 2′-deoxyisocytidine (75 mg, 0.33 mmol) in trimethylphosphate (0.822 mL) at 0 °C was treated with $POCl_3$ (0.039 mL) and stirred for 1.5 h. To this mixture was then rapidly added a solution of tris(tri-*n*-butylammonium) pyrophosphate (1.2 g) in dimethylformamide (2.5 mL), prepared as described above. The mixture was agitated vigorously for 1 min. A solution of Et_3NH^+ bicarbonate (1 M, pH 8.0, 2.0 mL) was then added, and the solvents were removed by evaporation. The residue was dissolved in water (10 mL), and the mixture was resolved by chromatography on Sepahadex A-25 (formate form, 30 mL), eluted with a gradient of Et_3NH^+ formate (0.2 to 1.5 M, pH 6.8). Further purification could be achieved by HPLC (TSK ODS-3 column, eluted with Et_3NH^+ acetate over 20 min in a gradient of 0 to 40% CH_3CN). FAB-MS: $(M + 3)^-$ 466.0. The ^1H NMR spectrum showed no contaminating dUTP. The triphosphate of deoxyisocytidine is unstable. After 6 weeks storage at −20 °C, an aqueous solution of d-iso-CTP contains only ca. 35% of the original material. Therefore, d-iso-CTP was used immediately after preparation, and ^1H NMR showed no detectable impurities in the samples used in the work reported here. FAB-MS: $(M + 3)^-$ 466.0.

(3) Ribose Derivatives of Isoguanosine

Isoguanosine Triphosphate. The procedure below was adapted from those of Mantsch *et al.* (1975) and Kazimierczuk and Shugar (1973). Hydrogen peroxide (30%, 3.5 mL) was added to a solution of Na_2CO_3 (2.6 g, 25 mmol) in water (25 mL). Maleic anhydride (2.45 g, 25 mmol) was then added, and the mixture was stirred at 0 °C for 30 min, at which point all of the maleic acid had dissolved. Concentrated H_2SO_4 (1.5 mL) in water (7.0 mL) was then added at 0 °C. The mixture was extracted with ether (8 × 25 mL), and the combined extracts were stored at 0 °C.

A solution of monopermaleic acid in water was obtained by evaporating 17.1 mL of the ether extracts in the presence of water (2.0 mL) in a stream of air. The pH of the solution was adjusted to 7.0 with NaOH (1 M), and a solution of the disodium salt of ATP (276 mg, 0.5 mmol, Fluka, in 1 mL of water, pH 7.0) was added. The reaction mixture was stirred for 24 h at RT, and the pH was adjusted to 4.5 with HCl (1 M). EtOH (absolute, 30 mL) was added, and the resulting precipitate was recovered by centrifugation and dissolved in water (2.0 mL), the pH adjusted to 4.5, and the product reprecipitated with ether. Chromatography (Dowex-1, bicarbonate form, eluted with a gradient of 0.01 to 1.0 M NH_4HCO_3 in water) yielded the *N*-oxide of ATP (208 mg, 73%).

The *N*-oxide of ATP (50 mg, 0.090 mmol) was dissolved in water (45 mL), and the solution was placed in a quartz tube. This was immersed in a second quartz tube containing aqueous acetic acid (10%). The distance between the walls of the tubes was 0.5 cm. The mixture was irradiated for 2 h (loss of absorbance at 235 nm, temperature ca. 45 °C), the pH adjusted to 10 (25% NH_4OH), and the mixture stirred at RT overnight. Water was then removed under reduced pressure at 30 °C, and the material was purified by HPLC (TSK ODS 120-T semipreparative column), eluted as described above (retention time 14.14 min). Fractions containing product were lyophilized three times with water to remove buffer, to yield isoguanosine triphosphate (8.18 mg, 12%), as determined by UV spectroscopy (λ_{max} = 291 nm).

Isoguanosine 5′-monophosphate was prepared according to essentially the same procedure as above, except that the *N*-oxide was submitted to photolysis directly after precipitation from ethanol/water. The final product was purified by HPLC as before (3% yield).

(4) Deoxyribose Derivatives of Isoguanosine

2′-Deoxyisoguanosine 3′-monophosphate was prepared according to the same procedure used for the preparation of isoguanosine 5′-triphosphate, except that the *N*-oxide was submitted to photolysis directly after precipitation from ethanol/water. The final product was purified by HPLC, and the amount of product was determined spectrophotometrically (λ_{max} = 292 nm, ϵ = 11 000, 0.0208 mmol, 18% for two steps).

2′-Deoxyisoguanosine 5′-triphosphate was prepared according to the same procedure used for the preparation of isoguanosine 5′-triphosphate, except that the *N*-oxide was submitted to photolysis directly after precipitation from ethanol/water. The final product was purified by HPLC.

The product (1.81 μmol), 2.6%) was determined by UV spectroscopy (λ_{max} = 292 nm, ϵ = 11 000). FAB-MS: 504.6. ^{31}P NMR indicated that the product was free of impurities containing phosphorus. However, a small amount (ca. 10%) of an unidentified unphosphorylated impurity remained. This could not be removed, either by HPLC or by extended irradiation. No similar impurity was observed in the preparation of either iso-GTP or 2′-deoxyiso-guanosine.

3′,5′-O-Bis(tert-butyldimethylsilyl)-2′-deoxyadenosine N-Oxide. To a solution of 3′,5′-*O*-bis(*tert*-butyldimethylsilyl)-2′-deoxyadenosine (Ogilvie, 1973) (6.0 g, 12.5 mmol) in a mixture of EtOH and H_2O (2.5:1) was added at RT monopermaleic acid (24 mmol). After 2 days, another 6 equiv of monopermaleic acid was added, and the reaction was continued for 2 weeks. Continuous monitoring (TLC on SiO_2, 10% MeOH in CH_2Cl_2) showed that the reaction did not proceed to completion, even though peracid was still detectable (KI test). After 2 weeks, CH_2Cl_2 was added, the organic phase was extracted with saturated $NaHCO_3$, and 3′,5′-*O*-bis(*tert*-butyldimethylsilyl)-2′-deoxyadenosine *N*-oxide (3.41 g, 55%) was isolated by chromatography (SiO_2, EtOAc:hexane 4:1), together with starting material (2.33 g, 39%).

3′,5′-O-Bis(tert-butyldimethylsilyl)-2′-deoxyisoguanosine. *N*-Oxide from the previous step (546 mg, 1.1 mmol) was dissolved in a mixture of EtOH (170 mL) and H_2O (170 mL). The solution was placed into a quartz reaction vessel with a mechanical stirrer and irradiated with a mercury lamp for 6 h. After irradiation, the pH of the mixture was adjusted to 10.0 (dilute NH_4OH), and the mixture was stirred at RT overnight. In the morning, the pH was adjusted to 7.0 (10% HCl), the EtOH was evaporated, the aqueous phase was extracted with CH_2Cl_2 and dried ($MgSO_4$), and 3′,5′-*O*-bis-(*tert*-butyldimethylsilyl)-2′-deoxyisoguanosine (330.9 mg, 60%) was isolated by chromatography (SiO_2, eluted first with EtOAc until all of the 3′,5′-*O*-bis(*tert*-butyldimethylsilyl)-2-deoxyadenosine eluted and then with 20% MeOH in CH_2Cl_2).

N^6-(N,N-Dibutylformamidyl)-3′,5′-O-bis(tert-butyldimethylsilyl)-2′-deoxyisoguanosine. The general procedure was adapted from that of McBride *et al.* (1986). 3′,5′-*O*-Bis(*tert*-butyldimethylsilyl)-2′-deoxyisoguanosine (292 mg, 0.59 mmol), *N*,*N*-dibutylformamide dimethylacetal (0.405 mL, 1.77 mmol Frohler & Mateucci, 1983), and pyridine (2.5 mL) were mixed and stirred at RT for 24 h. Complete conversion to a single product was indicated by TLC (5% MeOH in CH_2Cl_2). Solvents were removed by evaporation, and the residue was chromatographed (SiO_2, eluted with 5% MeOH in EtOAc) to yield the protected derivative of 2′-deoxyisoguanosine (270 mg, 73%). UV: λ_{max} = 350, 263 nm.

Removal of Formamidine from Deoxyisoguanosine. Formamidyldeoxyisoguanosine (25 mg) was suspended in aqueous ammonia (25%, 1 mL) and ethanol (1 mL) and stirred at RT for 24 h. Solvents were then evaporated, and the residue was purified by HPLC (TSK OD5-3 column, eluted with aqueous Et_3NH^+ acetate in a gradient of 0 to 40% CH_3CN over 20 minutes). The half-life for deprotection at 25 °C was 8 min. FAB-MS: $(M + H)^+$ 268.1.

N^6-(N,N-Dibutylformamidyl)-O^2-(2-p-nitrophenethyl)-3′,5′-O-bis(tert-butyldimethylsilyl)-2′-deoxyisoguanosine. The general procedure was adapted from that of Himmelsbach *et al.* (1984). A mixutre of formamidylated bis(TBDMS)d-iso-G (150 mg, 0.24 mmol), *p*-nitrophenethyl iodide [133 mg, 0.48 mmol, prepared by treating 2-(*p*-nitrophenyl)ethanol with triphenylphosphite methiodide (Landauer & Rydon, 1953)], Ag_2CO_3 (132 mg), and toluene (2.5 mL) was heated (80 °C) for 75 min. The mixture was then allowed to stand at RT for 1 h. Duplicate runs of the reaction were combined and filtered, and the protected product (243.8 mg, 65%) was isolated by chromatography (SiO_2, hexane/EtOAc 65:35). FAB-MS: $(M + H)^+$ 784.6.

N^6-(N,N-Dibutylformamidyl)-O^2-(2-p-nitrophenethyl)-2′-deoxyisoguanosine. The product of the previous step (219 mg, 0.28 mmol) in THF (0.19 mL) was treated with tetrabutylammonium fluoride on SiO_2 (43 mg, 0.84 mmol, Fluka). The reaction mixture was quenched via addition of $NaHCO_3$ (saturated solution) and extracted with CH_2Cl_2, and the extracts were dried ($MgSO_4$) to yield *N*,*N*-dibutyl-*N*-formamidyl-*O*^{2}-(2-*p*-nitrophenethyl)-2′-deoxyisoguanosine (101.5 mg, 65%), isolated by chromatography (SiO_2, eluted with 10% MeOH in CH_2Cl_2). The low yield was caused by partial removal of the NPE group under these conditions.

N^6-(N,N-Dibutylformamidyl)-O^2-(2-p-nitrophenethyl)-2′-deoxy-5′-(dimethoxytrityl)isoguanosine. A mixutre of the protected derivative of 2′-deoxyisoguanosine (93.3 mg, 0.168 mmol) from above, dimethoxytrityl chloride (68.3 mg, 0.202 mmol), and pyridine (2.0 mL) was stirred at RT for 2 h. MeOH (3 drops) was added, the solvents were removed by evaporation, the residue was taken up in CH_2Cl_2, and the organic layer was washed with aqueous Na_2CO_3 (5%). The organic layer was concentrated, and *N*^{6}-(*N*,*N*-dibutylformamidyl)-*O*^{2}-(2-*p*-nitrophenethyl)-2′-deoxy-5′-(dimethoxytrityl)isoguanosine (105.3 mg, 73%) was isolated by chromatography (SiO_2, eluted with CH_2Cl_2:MeOH:pyridine 100:1:2).

Preparation of Deoxyisoguanosine Phosphoramidite. Protected deoxyisoguanosine (105.3 mg, 0.122 mmol), bis(*N*,*N*-diisopropylamino)-(2-cyanoethoxy)phosphine (Aldrich, 0.046 mL, 0.146 mmol, 1.2 equiv), and diisopropylammonium tetrazolide (McBride *et al.*, 1986; 10.4 mg, 0.06 mmol) were dissolved in CH_3CN (2.0 mL) and stirred for 1 h. The reaction was monitored by TLC (SiO_2, eluted with EtOAc:CH_2Cl_2:triethylamine 45:45:10). An additional portion (0.02 mL) of bis(*N*,*N*-diisopropylamino)-(2-cyanoethoxy)phosphine was then added, and stirring was continued for an additional hour. Water (2 drops) was added, the mixture was stirred for 15 min and then diluted with CH_2Cl_2 (30 mL), and the organic layer was washed with aqueous Na_2CO_3 (2%) and dried (Na_2SO_4). The phosphoramidite (120.3 mg, 93%) was isolated by chromatography (SiO_2, EtOAc:CH_2Cl_2:triethylamine 45:45:10 as eluant). ^{31}P NMR: 149.74, 149.69.

(5) Oligonucleotides

DNA synthesis was performed on an Applied Biosystems 380B automated oligonucleotide synthesizer. All oligonucleotides were purified by HPLC (Macherey Nagel SS 250 Nucleosil 300-7 C-4; Et_3NH^+ acetate (A) and CH_3CN (B), 80% (A)/20% (B) to 70% (A)/30% (B) over 20 min, 4 mL/min flow rate), and a sample of each purified template was labeled with P-32 (polynucleotide kinase with [γ-^{32}P]ATP) and examined by gel electrophoresis to establish its purity.

Incorporation of nonstandard bases into synthetic oligonucleotides was verified by digestion of samples according to the procedure of Eritja *et al.* (1986), chromatography of the resulting products, and identification by comparison with authentic standards. Sequences of oligoribonucleotides were determined by the method of Randerath *et al.* (1980).

Nearest-neighbor analysis was performed by a modification of the procedure of Sgaramella and Khorana (1972). The product from incubation with [α-^{32}P]dCTP or [α-^{32}P]dGTP was isolated by gel electrophoresis and digested with micro-

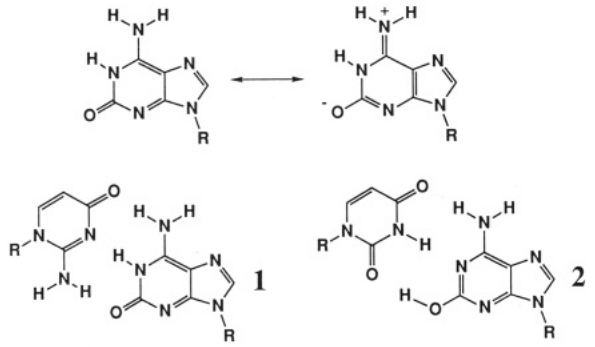

FIGURE 2: (Top) Two resonance structures of the N^1–H tautomeric form of isoguanosine. The form involving charge separation (right) is minor, yet because it brings formal aromaticity to the imidazole ring, it is more important than the corresponding tautomeric form in guanosine. (Bottom) Tautomeric forms of isoguanosine. Left is the N^1–H tautomeric form paired with isocytidine. Right is the O^2–H tautomeric form paired with uracil. The impact of the tautomeric forms on fidelity of incorporation of the (iso-C)·(iso-G) base pair is discussed in the text.

coccal nuclease and spleen phosphodiesterase. The digestion mixture was then diluted with carrier nucleoside 3′-monophosphates, the mixture separated by HPLC, and the radioactivity determined by scintillation counting.

Polymerase reactions were performed according to the method of Cobianchi and Wilson (1987). Reaction with AMV reverse transcriptase was performed according to the method of Larson *et al.* (1985). Boehringer Mannheim and Pharmacia both provided Klenow fragment of DNA polymerase (1 U/mL). Polynucleotide kinase was from Pharmacia. T7 RNA polymerase, AMV reverse transcriptase, and T4 DNA polymerase were from Pharmacia.

RESULTS

The syntheses proceeded mostly as expected on the basis of general precedents in nucleoside chemistry. Three exceptions are worth noting. First, exploratory efforts to acylate the exocyclic amino group of isoguanosine with isobutyric anhydride failed under conditions that isobutyrylate guanosine. The decreased nucleophilicity of the exocyclic amino group in iso-G can be understood in terms of a zwitterionic resonance structure that restores formal aromaticity to the imidazole nucleus of the purine ring system (Figure 2). This resonance structure is expected to be more important in isoguanine than in guanine, where the imidazole ring has formal aromaticity in the non-zwitterionic resonance form. Presumably, this resonance structure also accounts for the unusual UV absorbance of iso-G (λ_{max} = 291 nm). In any case, the decreased nucleophilicity of the exocyclic amino group prompted the development of a new protecting procedure based on the formamidine protecting group (McBride *et al.*, 1986).

Second, templates into which d-iso-C had been built were found to contain ca. 15% 2′-deoxyuridine, arising apparently from deamination of iso-C at some point during the synthesis or deprotection of the oligonucleotide. This product was observed in control experiments with deoxyisocytidine, formamidyl-protected deoxyisocytidine, and *N*-benzoylated deoxyisocytidine; in each case, treatment with aqueous ammonia (25%, 12 h, 60 °C) yielded 12–15% 2′-deoxyuridine as well as deprotected iso-C. An analogous (but slower) deamination of cytosine is well known.

Finally, the triphosphate of 2′-deoxyisocytidine was found to be rather unstable in comparison with the triphosphate of the ribose derivative isocytidine. After being stored for 6 weeks at –20 °C, an aqueous solution of d-iso-CTP contains only ca. 35% of the original material; the decomposition products were not identified. The corresponding ribo-iso-CTP is stable under these conditions.

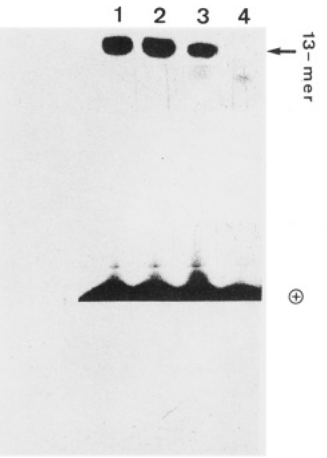

FIGURE 3: Incorporation of d-iso-GTP into an oligonucleotide opposite iso-C in a template, catalyzed by the Klenow fragment of DNA polymerase. Concentrations are 100 μM for each nucleotide component. Lanes 1 and 2 used the template–primer combination 5′-GATTTTGA-3′ and 3′-CTAAAACTGGCGA-5′. Lane 1: incubated with dGTP, dCTP, and [α-^{32}P]TTP. Lane 2: incubated with dGTP, dCTP, d-iso-GTP, and [α-^{32}P]TTP. Lanes 3 and 4 used the template–primer combination 5′-GATTTTGA-3′ and 3′-CTAAAACTGG(iso-C)GA-5′. Lane 3: incubated with dCTP, d-iso-GTP, and [α-^{32}P]TTP. Lane 4: incubated with dCTP and [α-^{32}P]-TTP.

To explore the biochemical scope of the base pair between iso-C and iso-G, a series of template–primer and promoter–template combinations (the first for DNA polymerases, the second for RNA polymerases) were prepared. In most cases, the templates were designed so that a unique base bearing an α-^{32}P label was the final base incorporated. Following incubation with polymerase, the products were resolved by gel electrophoresis and visualized by autoradiography. A radioactive band at the position corresponding to the full-length product identified successful reading through the nonstandard base in the template. The oligonucleotide products were then isolated and analyzed to determine how read-through was achieved. As controls, incubations were run with templates containing only standard nucleotides or with templates containing the nonstandard nucleotide in the absence of the triphosphate of the complementary nonstandard nucleoside.

Different polymerases behaved differently when challenged to incorporate d-iso-GTP into a growing oligonucleotide chain opposite iso-C in a template. For example, the Klenow fragment of DNA polymerase yields a 13-mer containing d-iso-G when filling in a template-primer with iso-C in the single-stranded region (Figure 3). Incorporation of d-iso-GTP opposite iso-C was 52% as efficient (based on yield of product) as incorporation of dGTP opposite C (Figure 3). Nearest-neighbor analysis (Sgaramella and Khorana, 1972) using [α-^{32}P]CTP yielded three labeled nucleotides in approximately equimolar amounts: 3′-d-iso-GMP, 3′-dCMP, and 3′-dAMP. Each of these bases occurs once 5′ relative to a C in the product, establishing incorporation of iso-G at the expected position.

With the Klenow fragment, d-iso-C was not complemented by dGTP, dTTP, or dCTP. However, in the presence of dATP, the full-length product was seen to the extent of 16% of that observed when d-iso-GTP was used to complement d-iso-C in

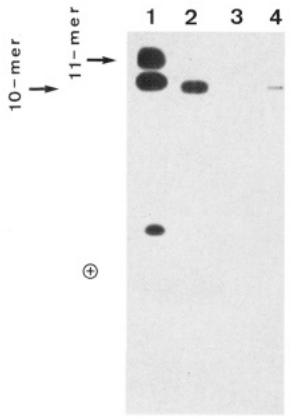

FIGURE 4: Incorporation of iso-GTP into an oligonucleotide opposite d-iso-C in a template, catalyzed by T7 RNA polymerase. Concentration are 2.5 mM for each nucleotide component, except as otherwise noted. Lane 1 used the promoter–template sequence 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGCCCGA-5′, incubated with GTP, CTP, and [α-^{32}P]UTP. Lanes 2–4 used the promoter–template sequence 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGC(iso-C)CGA-5′. Lane 2: incubated with GTP, CTP, iso-GTP, and [α-^{32}P]UTP. Lane 3: incubated with GTP, CTP, iso-GTP (12.5 mM), and [α-^{32}P]UTP. Lane 4: incubated with GTP, CTP, and [α-^{32}P]UTP.

the template. The full-length product observed with dATP in the absence of d-iso-GTP almost certainly does not arise because A complements d-iso-C. Rather, the template was found to contain ca. 15% 2′-deoxyuridine arising by deamination of d-iso-C noted above, and the template containing 2′-deoxyuridine is almost certainly the source of the full-length product obtained with ATP in the absence of d-iso-GTP.

T7 RNA polymerase also incorporates iso-GTP into a template opposite iso-C (Figure 4). As has been frequently observed in products of transcriptions catalyzed by T7 RNA polymerase (Milligan *et al.*, 1987), short failure sequences were often observed, and the full-length product was contaminated with oligonucleotides one base longer in the control reaction when a template with only natural bases was used. Products with an extra base were not observed with the template containing iso-C, however; this behavior is not understood. These experiments suggested that either iso-GTP itself or an impurity in the preparation inhibited the reaction, as incorporation was not detected at high concentrations of iso-GTP.

To establish that iso-GTP was incorporated into the RNA products, the full-length reaction product was partially degraded according to the procedure of Randerath *et al.* (1980). The products were resolved by gel electrophoresis to yield a ladder of fragments of defined length. These were separated, 5′-labeled with ^{32}P, and totally hydrolyzed with nuclease P1 to yield a set of labeled nucleoside 5′-monophosphates, each arising from the 5′-end of a ladder fragment of defined length. These were identified by ion-exchange chromatography using the appropriate synthetic standards, yielding a sequence of the product RNA. This analysis showed that iso-G was indeed incorporated at the expected position opposition iso-C (Figure 5). As the fragments were detected by radioactive label, the method was rather sensitive. More than a few percent misincorporation could be excluded.

Incorporation of the triphosphate of the pyrimidine component of the nonstandard base pair (iso-C) opposite the purine component (iso-G) was then explored. Surprisingly, the Klenow fragment of DNA polymerase incorporates *both* d-iso-CTP and TTP opposite iso-G in the template (Figure 6).

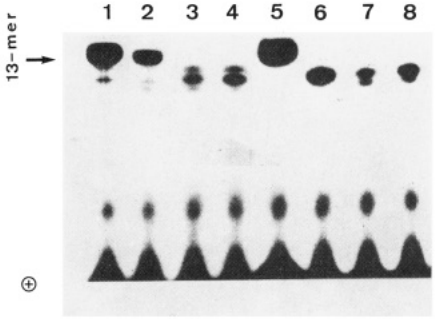

FIGURE 5: Fidelity of incorporation of d-iso-GTP opposite iso-C in a template in a polymerization reaction catalyzed by the Klenow fragment of DNA polymerase. Concentrations are 100 μM for each nucleotide component. Lane 1–4 used the template–primer combination 5′-GATTTTGA-3′ and 3′-CTAAAACTGG(iso-C)GA-5′. Lane 1: incubated with dCTP, d-iso-GTP, and [α-^{32}P]TTP. Lane 2: incubated with dCTP, dATP, and [α-^{32}P]TTP. Lane 3: incubated with dCTP, dGTP, and [α-^{32}P]TTP. Lane 4: incubated with dCTP and [α^{32}-P]TTP. Lanes 5 and 6 used the template–primer combination 5′-GATTTTGA-3′ and 3′-CTAAAACTGGTGA-5′. Lane 5: incubated with dCTP, d-iso-GTP, and [α^{32}P]TTP. Lane 6: incubated with dCTP and [α-^{32}P]TTP. Lanes 7 and 8 used the template–primer combination 5′-GATTTTGA-3′ and 3′-CTAAAACTGGCGA-5′. Lane 7: incubated with dCTP, d-iso-GTP, and [α-^{32}P]TTP. Lane 8: incubated with dCTP and [α-^{32}P]TTP.

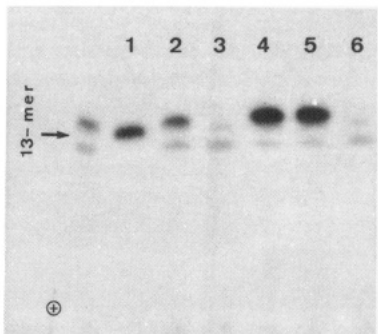

FIGURE 6: Incorporation of both d-iso-CTP and TTP into an oligonucleotide opposite iso-G in a template, catalyzed by the Klenow fragment of DNA polymerase I. Concentrations are 100 μM for each nucleotide component. The minor lower band in lanes 2–6 is presumably the failure sequence arising from misincorporation of dA opposite d-iso-G. The slightly different mobility of full-length product in lane 1 arises from its different composition. Lane 1 used the primer–template combination 5′-GATTTTGA-3′ and 5′-CTAAAACTGGCGA-3′, incubated with dGTP, dCTP, and [α-^{32}P]-TTP. Lanes 2–6 used the template–primer combination 5′-GATTTTGA-3′ and 3′-CTAAAACTCC(iso-G)CT-5′. Lane 2: incubated with dGTP, d-iso-CTP, and [α-^{32}P]ATP. Lane 3: incubated with dGTP and [α-^{32}P]ATP. Lane 4: incubated with dGTP, dTTP, and [α-^{32}P]ATP. Lane 5: incubated with dGTP, d-iso-CTP, TTP, and [α-^{32}P]ATP. Lane 6: incubated with dGTP, dCTP, and [α-^{32}P]ATP.

Interestingly, TTP is the more efficient substrate, at least judging from the relative amounts of full-length product observed. A nearest-neighbor analysis (again using [α-^{32}P]-CTP as source of label) of products obtained in the absence of TTP gave a ratio of 3′-d-iso-CMP, 3′-dAMP, and 3′-dGMP in a ratio of 1:1:1, again confirming site-specific incorporation.

In contrast to the Klenow enzyme, T7 RNA polymerase proved less willing to accept iso-CTP. Scarcely any full-length product was observed with a template containing iso-G presented with iso-CTP in the incubation mixture. When given the opportunity, however, T7 RNA polymerase efficiently incorporates UMP opposite iso-G (Figure 7).

AMV reverse transcriptase was also observed to incorporate iso-G opposite iso-C in a template (Figure 8). With this

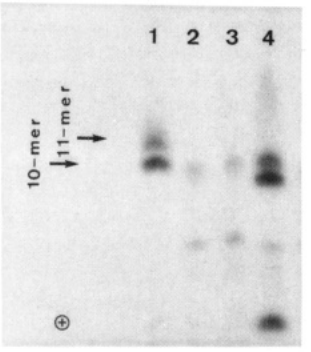

FIGURE 7: Incorporation of UTP but not iso-CTP into an oligonucleotide opposite iso-G in a template, catalyzed by T7 RNA polymerase. Concentrations are 2.5 mM for each nucleotide component. Failure sequences represented by the lower bands are not uncommon with T7. Lane 1 used the template–promoter combination 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGCCCGA-5′, incubated with GTP, CTP, and [α-^{32}P]UTP. Lanes 2–4 used the template–promoter combination 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGC(iso-G)CGT-5′. Lane 2: incubated with GTP, CTP, iso-CTP, and [α-^{32}P]ATP. Lane 3: incubated with GTP, CTP, and [α-^{32}P]ATP. Lane 4: incubated with GTP, CTP, UTP, and [α-^{32}P]ATP.

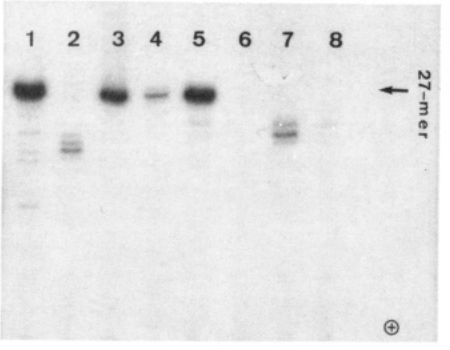

FIGURE 8: Experiments with AMV reverse transcriptase (lanes 1–4) and T4 DNA polymerase (lanes 5–8). Concentrations are 50 μM for each nucleotide component. AMV reverse transcriptase: lanes 1 and 2 used the primer–template combination 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGCCCGA-5′. Lane 1: incubated with dGTP, dCTP, and [α-^{32}P]TTP. Lane 2: incubated with dCTP and [α^{32}-P]TTP. AMV reverse transcriptase: lanes 3 and 4 used the primer–template combination 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGC(iso-C)CGA-5′. Lane 3: incubated with dGTP, dCTP, d-iso-GTP and [α-^{32}P]TTP. Lane 4: incubated with dGTP, dCTP, and [α-^{32}P]TTP. T4 DNA polymerase: lanes 5 and 6 used the primer–template combination 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGCCCGA-5′. Lane 5: incubated with dGTP, dCTP, and [α-^{32}P]TTP. Lane 6: incubated with dCTP and [α-^{32}P]TTP. T4 DNA polymerase: lanes 7 and 8 used the primer–template combination 5′-TAATACGACTCACTATAG-3′ and 3′-ATTATGCTGAGTGATATCGCGGC(iso-C)CGA-5′. Lane 7: incubated with dGTP, dCTP, d-iso-GTP), and [α-^{32}P]TTP. Lane 8: incubated with dGTP, dCTP, and [α-^{32}P]TTP.

polymerase, small amounts of full-length product were observed with the template containing iso-C even in the absence of d-iso-GTP when dGTP and dCTP were present in the incubation mixture. In contrast, T4 DNA polymerase did not incorporate any detectable amounts of d-iso-GTP when challenged with a template containing iso-C (Figure 8).

DISCUSSION

The results reported here show that various polymerases handle the (iso-C)·(iso-G) base pair with their own characteristic idiosyncrasies. T4 DNA polymerase refused to accept the nonstandard base pair under the conditions examined. The remaining polymerases tolerated the nonstandard base pair to different degrees, with iso-G preferred more as the triphosphate than iso-C. The Klenow fragment appears to accept the new base pair with the greatest facility.

Interestingly, some polymerases also pair iso-G with T (or U). Indeed, when presented with a template containing iso-G, T7 RNA polymerase appears to prefer UTP over iso-CTP. With the Klenow fragment, incorporation of d-iso-CTP occurs together with incorporation of TTP, and both d-iso-GTP and dATP can complement T in a template. The mismatching of iso-G and T is perhaps not surprising. Shugar and co-workers pointed out some time ago (Sepiol *et al.*, 1976) that iso-G exists in two tautomeric forms (Figure 2). The N^1–H tautomer (having a keto group, **1**) contributes ca. 90% of the total in aqueous media, while the phenolic O^2–H tautomer (**2**) contributes ca. 10% (Figure 2). While the N1–H tautomer is complementary to iso-C in a standard Watson–Crick geometry, the phenolic tautomer is complementary to T (or U), with three hydrogen bonds being formed instead of the two found in the A·T base pair.

We cannot say at this point whether T7 RNA polymerase prefers to pair U opposite iso-G (rather than iso-C) because it prefers the minor tautomer of iso-G in the template or because it prefers UTP as the natural nucleotide over the unnatural iso-CTP. In this context, it is worth noting that Sepiol *et al.* (1976) have shown that the amount of the phenolic tautomer of iso-G increases with decreasing solvent dielectric. Although the microscopic dielectric constant within the active site of a polymerase constructing a double-stranded helix is difficult to know, it is conceivable that the environment of iso-G when the recognition event occurs in T7 RNA polymerase is less polar than that of bulk water and that the amount of the O^2–H tautomer is correspondingly higher. In any case, as minor tautomeric forms of natural nucleosides are implicated in spontaneous mutation (Watson & Crick, 1953b; Topal & Fresco, 1976), this system offers intriguing possibilities for experimental studies on the mechanism by which polymerases prevent errors arising from minor tautomeric forms.

The tautomeric forms of iso-G suggest that it might have antiviral activity via a novel strategy. Provided that it can enter a cell and be phosphorylated, iso-G might be incorporated as its minor tautomer into an oligonucleotide opposite uracil by a reverse transcriptase with low fidelity but not into genomic DNA by polymerases with higher fidelity. Once incorporated, with the tautomeric equilibrium reestablished, iso-G might hinder or block entirely the biological function of the retroviral nucleic acid. Conversely, iso-G might be incorporated into a viral RNA genome with deleterious effects, while its incorporation into mRNA might have more transient effects.

It remains interesting to ask why natural oligonucleotides did not evolve to exploit the (iso-C)·(iso-G) base pair as the fifth and sixth "letters" in a genetic alphabet (Szathmary, 1991). In another, more whimsical form, we might ask whether the (iso-C)·(iso-G) base pair might be found as part of a coding system in extraterrestrial life.

The data reported here offer ample functional grounds (Benner & Ellington, 1988) to explain why the (iso-C)·(iso-G) base pair might be disfavored (over A, T, C, and G) as a component of a coding system. Deamination of iso-C appears to be faster than deamination of C [$t_{1/2}$ = 40 years for deamination of C in pure water (Frick, 1987)]. In both cases, the deamination reaction introduces a mutation. Further, the accessibility of an ambivalent O^2–H tautomeric form of iso-G might cause intrinsic infidelity in the replication of genetic information in a coding system also containing the

A·T (or A·U) base pair. Although a sophisticated enzyme might well enforce tautomeric fidelity upon iso-G, it would be more difficult to obtain such fidelity in primitive forms of life where the choice of nucleotide bases was presumably made.

While the intrinsic properties of iso-C and iso-G make it unlikely that they will be found (or previously existed) as components of a genetic system (at least in the presence of the A·T base pair), they do not render the base pair useless in the laboratory. As noted elsewhere (Switzer *et al.*, 1989; Bain *et al.*, 1992), much can be done with the (iso-C)·(iso-G) base pair in laboratory experiments *in vitro*, provided that the chemistry intrinsic to the components of this base pair is understood.

REFERENCES

Bain, J. D., Chamberlin, A. R., Switzer, C. Y., & Benner, S. A. (1992) *Nature 356*, 537–539.

Benner, S. A. (1988) *FEBS Lett. 233*, 225–228.

Benner, S. A., & Ellington, A. D. (1988) *Crit. Rev. Biochem. 23*, 369–426.

Benner, S. A., & Allemann, R. K. (1989) *Trends Biochem. Sci. 14*, 396–397.

Benner, S. A., Allemann, R. K., Ellington, A. D., Ge, L., Glasfeld, A., Leanz, G. F., Krauch, T., MacPherson, L. J., Moroney, S., Piccirilli, J. A., & Weinhold, E. (1987) in *Cold Spring Harbor Symp. Quant. Biol.*, Vol. 52, pp 53–63, Cold Spring Harbor Publications, Cold Spring Harbor, New York.

Cherbuliez, E., & Bernhard, K. (1932) *Helv. Chim. Acta 15*, 464–471.

Cobianchi, F., & Wilson, S. H. (1987) in *Methods in Enzymology*, (Berger, S. L., & Kimmel, A. R., Eds.) Vol. 152, pp 94–110, Academic Press, New York.

Elgureo, J., Marzin, C., Katritzky, A. R., & Linda, P. (1976) *Adv. Heterocycl. Chem. Suppl. 1*, 1.

Eritja, R., Horowitz, D. M., Walker, P. A., Ziehler-Martin, J. P., Boosalis, M. S., Goodman, M. F., Itakura, K., & Kaplan, B. E. (1986) *Nucleic Acids Res. 14*, 8135–8153.

Eschenmoser, A., & Dobler, M. (1992) *Helv. Chim. Acta 75*, 218–259.

Frick, L. (1987) *Bioorg. Chem. 15*, 100.

Frohler & Mateucci (1983) *Nucleic Acids Res. 11*, 8031.

Hashimoto, H., & Switzer, C. (1992) *J. Am. Chem. Soc. 114*, 6255–6256.

Himmelsbach, F., Schulz, B. S., Trichtinger, T., Charubala, R., Pfleiderer, W. (1984) *Tetrahedron 40*, 59–72.

Huang, Z., Schneider, K. C., & Benner, S. A. (1991) *J. Org. Chem. 56*, 3869–3882.

Joyce, G. F., Schwartz, A. W., Miller, S. L., & Orgel, L. E. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4398–4402.

Kazimierczuk, Z., & Shugar, D. (1973) *Acta Biochim. Pol. 20*, 395–402.

Kazimierczuk, Z., Mertens, R., Kawczynski, W., & Seela, F. (1991) *Helv. Chim. Acta 74*, 1742–1748.

Kimura, J., Yagi, K., Suzuki, H., & Mitsunobu, O. (1980) *Bull. Chem. Soc. Jpn. 53*, 3670–3677.

Landauer, S. R., & Rydon, H. N. (1953) *J. Chem. Soc.* 2224.

Larson, K., Sahm, J., Shenkar, R., & Strauss, B. (1985) *Mutat. Res. 150*, 77–84.

Mantsch, H. H., Goia, I., Kezdi, M., Bârzu, O., Dânsoreanu, M., Jebeleanu, G., & Ty, N. G. (1975) *Biochemistry 14*, 5593–5601.

McBride, L. J., Kierzek, R., Beaucage, S. L., & Caruthers, M. H. (1986) *J. Am. Chem. Soc. 108*, 2040–2048.

Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783–8799.

Moore, M. F., & Beaucage, S. L. (1985) *J. Org. Chem. 50*, 2019–2025.

Moroney, S. E., & Piccirilli, J. A. (1991) *Biochemistry 30*, 10343–10349.

Ogilvie, K. K. (1973) *Can. J. Chem. 51*, 3799–3807.

Pettit, G. R., Ode, R. H., Coomes, R. M., & Ode, S. L. (1976) *Lloydia 39*, 363–367.

Piccirilli, J. A., Krauch, T., Moroney, S. E., & Benner, S. A. (1990) *Nature 343*, 33–37.

Piccirilli, J. A., Moroney, S. E., & Benner, S. A. (1991a) *Biochemistry 30*, 10350–10356.

Piccirilli, J. A., Krauch, T., MacPherson, L. J., & Benner, S. A. (1991b) *Helv. Chim. Acta 74*, 397–406.

Randerath, K., Gupta, R. C., & Randerath, E. (1980) *Methods Enzymol. 65*, 638–680.

Rich, A. (1962) in *Horizons in Biochemistry* (Kasha, M., & Pullman, B., Eds.), pp 103–126, Academic Press, New York.

Schneider, K. C., & Benner, S. A. (1990a) *J. Am. Chem. Soc. 112*, 453–455.

Schneider, K. C., & Benner, S. A. (1990b) *Tetrahedron Lett. 31*, 335–338.

Sepiol, J., Kazimierczuk, Z., & Shugar, D. (1976) *Z. Naturforsch 31C*, 361–370.

Sgaramella, V., & Khorana, H. G. (1972) *J. Mol. Biol. 72*, 427–444.

Switzer, C. Y., Moroney, S. E., & Benner, S. A. (1989) *J. Am. Chem. Soc. 111*, 8322–8323.

Szathmary, E. (1991) *Proc. R. Soc. London B 245*, 91–99.

Topal, M. D., & Fresco, J. R. (1976) *Nature 263*, 285–289.

Watanabe, K. A., Reichman, U., Chu, C. K., & Fox, J. J. (1978) in *Nucleic Acid Chemistry* (Tipson, R. S., & Townsend, L. B., Eds.) Part 1, pp 273–277, John Wiley and Sons, New York.

Watson, J. D., & Crick, F. H. (1953a) *Nature 171*, 737–738.

Watson, J. D., & Crick, F. H. (1953b) *Nature 171*, 964–967.

Wissler, J. H., Kiesewetter, S., Logemann, E., Sprinzl, M., & Heilmeyer, L. M. G., Jr. (1989) *Biol. Chem. Hoppe-Seyler 370*, 975.