Non-Enzymatic Transcription of an IsoG·IsoC Base Pair

John C. Chaput and Christopher Switzer*

Department of Chemistry
University of California
Riverside, California 92521

Received June 23, 2000

During the course of molecular evolution, the fitness of nongenomic nucleobases such as iso-G and iso-C may have been explored.[1−3] The iso-G·iso-C base pair in particular may be predicted to share several physical attributes with its isomeric parent, G·C, that could be advantageous in the context of abiotic reactions. For example, apart from the same hydrogen-bond donor/acceptor groups,[4] dipole moments[5] of iso-G and iso-C nucleobases should be similar to those of G and C. Reflecting these factors, the ab initio derived interaction energy of iso-G and iso-C in a Watson−Crick geometry compares to that of G and C, and is 2−3-fold greater than that of (2-amino-)A and T (Figure 1).

Incorporation of iso-G·iso-C into an expanded genetic system has been constrained by iso-G infidelity, seen in its ability to code for both T and iso-C. In vitro polymerase studies,[6] ribosomal translation,[7] and *rec*A-promoted DNA strand-exchange[8] experiments, all suffer from this phenomenon, which is consistently explained by the coexistence of N1−H and O2−H iso-G tautomers,[9,10] and the formation of iso-$G_{(N1-H)}$·iso-C and iso-$G_{(O2-H)}$·T Watson−Crick pairs (Figure 2). Crystallographic observation of both Watson−Crick and wobble iso-G·T geometries in a single double helix provides the most direct evidence to date in support of a multiple tautomer mechanism for iso-G infidelity.[11] In this structure, three water molecules are located along the major and minor groove edges of the iso-G·T wobble pair, but are absent from the iso-G·T Watson−Crick pair, owing to the proximity of a Hoechst minor groove binding agent (Figure 2, panel B).[11] These data suggest that aqueous conditions should favor the wobble over the Watson−Crick geometry of iso-G·T, and thereby provide a suitable environment for replication of iso-G and iso-C. We describe below nonenzymatic, template-directed experiments consistent with this prediction.

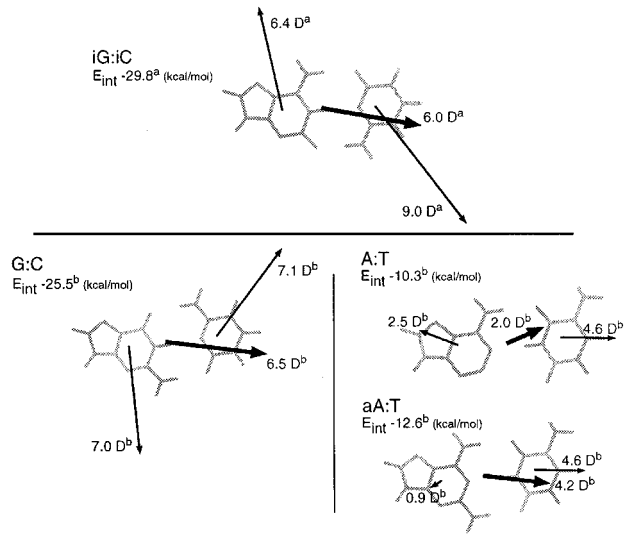

Figure 1. Summary of the Hartree−Fock 6-31G** dipole moments and interaction energies. Thin arrows refer to dipole moments for individual bases; thick arrows correspond to dipole moments for base pairs. Interaction energies have been corrected for basis set superposition error. [a]This work. [b]Reference 5b.

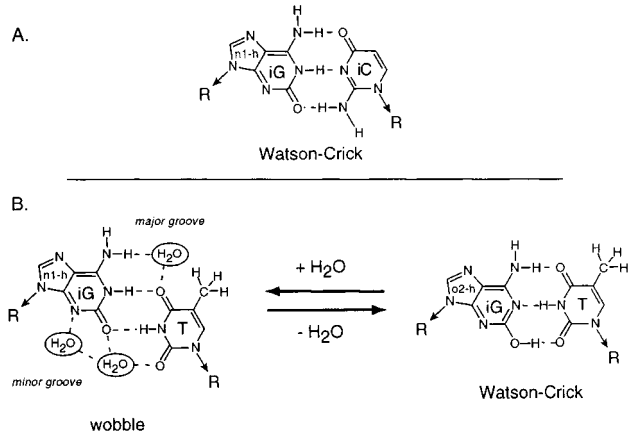

Figure 2. Iso-G base pairs relevent to the current study. Top, iso-G(N1−H):iso-C Watson−Crick base pair. Bottom, two iso-G:T base-pair geometries according to a dodecamer duplex crystal structure: Robinson, H.; Gao, Y.-G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H.-J. *Biochemistry* **1998**, *37*, 10897−10905.

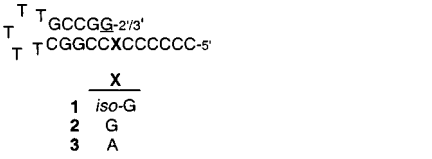

Figure 3. DNA hairpin templates used in nonenzymatic oligomerization reactions. The 3′-terminal·riboG residue is underscored.

Hairpin templates used in this study are shown in Figure 3. These templates were modeled after those first reported by Wu and Orgel.[12] We have previously used this system to demonstrate monomer oligomerization on templates bearing 2′,5′-linked RNA.[13] Deoxyribo-iso-G was incorporated into template **1** using methodology previously described.[14] Each hairpin template is composed of DNA except for a single ribo-G residue at the 3′-

* To whom correspondence should be addressed. Telephone: (909) 787-7266. Fax: (909) 787-4713. E-mail: switzer@citrus.ucr.edu.

(1) Rich, A. In *Horizons in Biochemistry*; Kasha, M., Pullman, B., Eds.; Academic Press: New York, 1962; pp 103−126.
(2) Eschenmoser, A. *Science* **1999**, *284*, 2118−2124.
(3) Benner, S. A.; Switzer, C. In *Simplicity and Complexity in Proteins and Nucleic Acids. Dahlem Workshop Report*; Frauenfelder, H., Deisenhofer, J., Wolynes, P. G., Eds.; Dahlem University Press: Berlin, 1999; pp 335−359.
(4) Pranata, J.; Wierschke, S. C.; Jorgensen, W. L. *J. Am. Chem. Soc*. **1991**, *113*, 2810.
(5) (a) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn*. **1996**, *14*, 117−135. (b) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem*. **1996**, *100*, 1965−1974.
(6) (a) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc*. **1989**, *111*, 8322−8323. (b) Switzer, C.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489-10496. (c) Tor, Y.; Dervan, P. B. *J. Am. Chem. Soc*. **1993**, *115*, 4461−4467.
(7) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537−539.
(8) Rice, K. P.; Chaput, J. C.; Cox, M. M.; Switzer, C. *Biochemistry* **2000**, *39*, 10177−10188.
(9) (a) Sepiol, J.; Kazimierczuk, Z.; Shugar, D. *Z. Naturforsch*. **1976**, *31*, 361−370. (b) Seela, F.; Wei, C.; Kazimierczuk, Z. *Helv. Chim. Acta* **1995**, *78*, 1843−54.
(10) See also: Pitsch, S.; Krishnamurthy, R.; Bolli, M.; Wendeborn, S.; Holzener, A.; Minton, M.; Lesueur, C.; Schlonvogt, I.; Jaun, B.; Eschenmoser, A. *Helv. Chim. Acta* **1995**, *78*, 1621−35.
(11) Robinson, H.; Gao, Y.-G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H.-J. *Biochemistry* **1998**, *37*, 10897−10905.
(12) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc*. **1992**, *114*, 317−322.
(13) Prakash, T. P.; Roberts, C.; Switzer, C. *Angew. Chem., Int. Ed. Engl*. **1997**, *36*, 1522−1523.

10.1021/ja002260m CCC: $19.00
Published on Web 12/27/2000

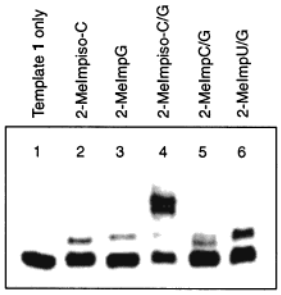

Figure 4. Autoradiogram of 20% PAGE-analyzed template-directed oligomerization reactions on template **1** after 10 days.

terminus to facilitate oligomerization. Oligomerizations on hairpin templates were performed in the presence of 200 mM $MgCl_2$, 1 M NaCl, 0.2 M lutidine·HCl, pH 8, and 100 mM activated monomer[15] at 0 °C. After 10 days, template-directed reactions were subjected to 20% PAGE analysis and visualized by autoradiography.

Results from template-directed oligomerization on iso-GC_6 bearing template **1** are given in Figure 4. Full-length product, determined by comparison to the results from all natural templates **2** and **3** (vide infra), was observed only in the presence of both activated iso-cytidylate and guanylate monomers as seen in lane 4 versus lanes 2 and 3. Apart from indicating successful incorporation of iso-C opposite template iso-G, these data also support the fidelity of iso-C relative to C, and fidelity of G relative to iso-G. Further, cooperative iso-cytidylate and guanylate residue stacking along the template prior to covalent bond formation is implied by formation of disproportionate quantities of products in lane 2 as $n+1$-mer (iso-C alone) and lane 4 as $\geq n+1$-mers (both iso-C and G present). A similar effect of downstream G-residues on the incorporation of upstream A, C, or U residues has been noted.[16]

To assess the fidelity of iso-G relative to the natural pyrimidines under nonenzymatic conditions, template **1** was incubated in separate experiments with 2-MeImpC and 2-MeImpU along with 2-MeImpG. As evident in lanes 5 and 6 of Figure 4, in neither case does full-length product form. The absence of full length product in lane 6 is consistent with iso-G recognition of U via a wobble rather than Watson−Crick geometry.

Templates **2** and **3** are designed to address the fidelity of iso-C relative to the natural purine bases in the context of nonenzymatic template reactions. As shown in Figure 5, template-directed oligomerization of 2-MeImpC/G and 2-MeImpU/G on their respective natural templates GC_6 (panel A, lane 4) and AC_6 (panel B, lane 4) occurs efficiently to afford full-length product in 10 days. Control experiments with pyrimidine monomer alone for a given cognate purine template (lane 2, panels A and B) showed very little extension product. In contrast, control experiments where templates **2** and **3** were incubated with G monomer alone showed visible products arising from multiple incorporations (Figure 5, lane 3, panel A (trace) and panel B). Misincorporations at the first template position are consistent with an equilibrium of *syn* and *anti* conformers of the template A/G residues and/or the incoming G residue, and the formation of purine(syn)·purine(anti) base pairs. The circumstances where the greatest

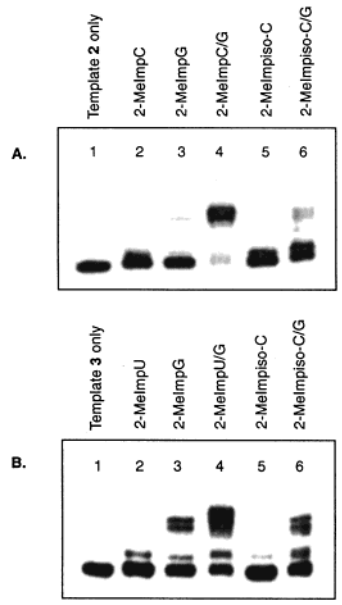

Figure 5. Autoradiogram of 20% PAGE-analyzed template-directed oligomerization reactions on templates **2** (panel A) and **3** (panel B) after 10 days.

extent of misincorporation is observed, Figure 5, panel B, lanes 3 and 6, may be explained by invoking an A(syn)·G(anti) pair. Products of this kind were notably absent from experiments with the iso-G bearing template **1** (Figure 4, lane 3), possibly because iso-G avoids a syn conformation. Incubation of templates **2** and **3** with 2-MeImp-iso-C/G does not lead to any appreciable difference in the quantity of full length product as compared to the incubation with 2-MeImpG alone (Figure 5, panels A and B, lanes 6 versus 3). These results indicate the lack of a pronounced tendency for iso-C to misincorporate opposite A or G in a template and be extended to longer products. However, lanes 5 and 6 of Figure 5, panel A, show that iso-C can yield single residue misincorporations opposite G in a template in the absence of competing C monomer.

The regiochemical preference for phosphodiester bond formation between the terminal ribo-G of the primer and the 5′-iso-CMP residue incorporated into full-length product from template **1** was examined by RNase T1 cleavage. RNase T1 specifically cleaves 3′,5′-phosphodiester ribonucleoside bonds located 3′ to a *ribo*-G residue.[12,13] The product resulting from digestion of the oligomerization product from template **1** is a single faster moving band with similar mobility to the original template (see Supporting Information). This result is consistent with iso-G directing the formation of 3′,5′-phosphodiester linkages in the product strand.

In summary, in contrast to previous enzymatic experiments, nonenzymatic conditions in an aqueous environment promote faithful transcription of the iso-G·iso-C base pair. This finding is in keeping with iso-G recognition of U(T) via a wobble rather than Watson−Crick geometry due to structural water molecules in the helical grooves and suggests that limitations surrounding incorporation of the iso-G·iso-C pair into an expanded genetic system may be avoided by appropriate engineering.

Acknowledgment. This work was funded by grants from the NASA Astrobiology Institute and the NASA exobiology program.

Supporting Information Available: A figure representing the autoradiogram from RNase T1 cleavage of the full-length product from template **1** (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA002260M

(14) Roberts, C.; Bandaru, R.; Switzer, C. *J. Am. Chem. Soc*. **1997**, *119*, 4640−4649.

(15) Isocytidine was synthesized according to the method of Mitsunobu: Kimura, J.; Yagi, K.; Suzuki, H.; Mitsunobu, O. *Bull. Chem. Soc*. **1980**, *53*, 3670−3677, and phosphorylated under standard conditions: Sowa, T.; Ouchi, S. *Bull. Chem. Soc. Jpn*. **1975**, *48*, 2084−2090. All ribonucleotide 5′-phosphoro-2-methylimidazolide (2-MeImpN) monomers were prepared as described in Joyce, G. F.; Inoue, T.; Orgel, L. E. *J. Mol. Biol*. **1984**, *176*, 279−306.

(16) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc*. **1992**, *114*, 5496−5501.