Efforts toward Expansion of the Genetic Alphabet: DNA Polymerase Recognition of a Highly Stable, Self-Pairing Hydrophobic Base

Dustin L. McMinn,† Anthony K. Ogawa,† Yiqin Wu, Jianquan Liu, Peter G. Schultz,* and Floyd E. Romesberg*

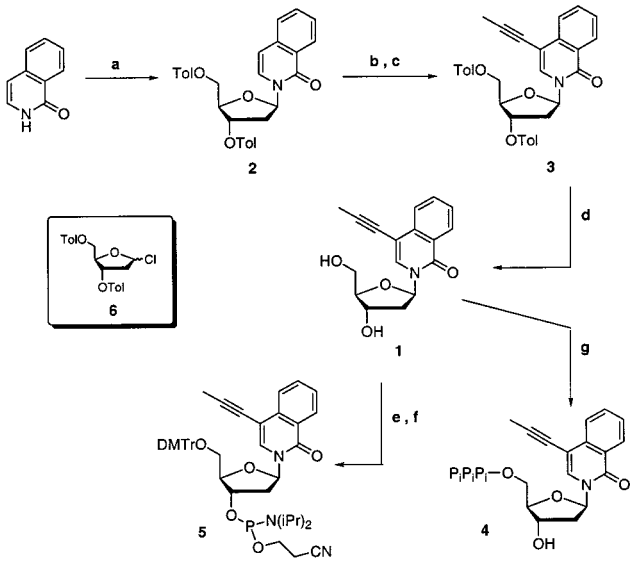
Department of Chemistry, The Scripps Research Institute 10550 North Torrey Pines Road La Jolla, California 92037

Received July 16, 1999

The storage and replication of biological information is based on the specific hydrogen-bonding patterns of the Watson−Crick A:T and G:C base pairs.[1] The generation of enzymatically replicable unnatural base pairs with thermal stability comparable to that of A:T and G:C would significantly expand the biological and chemical potential of DNA. One strategy for expanding the genetic alphabet has focused on the synthesis of bases with altered hydrogen-bonding patterns.[2−8] However, in light of the dominant role played by hydrophobicity in protein folding and stability,[9−11] we have explored the use of hydrophobic and van der Waals interactions to generate an orthogonal base pair for the storage and replication of genetic information. Previous experiments in which hydrophobic bases were incorporated into duplex DNA led to a destabilization of the duplex relative to the native pairs.[12−18] In the course of systematically evaluating a variety of predominantly hydrophobic nucleobases,[19] we have characterized a base, which cannot form H-bonds, whose self-pair stabilizes duplex DNA relative to an A:T or G:C base pair.

The 7-propynyl isocarbostyril nucleoside (**1**) was synthesized and converted into the nucleoside triphosphate (**4**) and phosphoramidite (**5**) as shown in Scheme 1. When incorporated into duplex DNA, the **1**:**1** interstrand interaction is expected to be completely hydrophobic in the absence of significant backbone deformation; the keto groups are designed to be in the minor groove and can H-bond to solvent. The thermal stability of the base pairs was evaluated by determining the melting temperature

Scheme 1[a]



[a] (a) Bis-TMS acetamide, **6**, then $SnCl_4$ (26%); (b) ICI, CH_2Cl_2 (87%); (c) propyne, $(Ph_3P)_2PdCl_2$, CuI, NEt_3 (77%); (d) 0.4 M NaOMe (87%); (e) DMTrCl, pyridine (80%); (f) 2-cyanoethyl diisopropylchlorophosphoramidite, DIEA, CH_2Cl_2 (65%); (g) $(Bu_3NH)_3(H_4P_2O_7)$, $POCl_3$, NBu_3, Proton Sponge, trimethylphosphate (7%).

Table 1. Denaturation Temperatures for **7**[a]

X	Y	T_M, °C	X	Y	T_M, °C	X	Y	T_M, °C
d**1**	d**1**	62.6 ± 0.52	dC	d**1**	51.4 ± 0.81	dT	dT	49.8 ± 0.18
dA	dT	59.2 ± 0.44	dG	d**1**	54.5 ± 0.56	dT	dC	45.0 ± 0.23
dC	dG	61.8 ± 0.09	dA	dA	52.7 ± 0.09	dT	dG	53.3 ± 0.87
dA	d**1**	55.5 ± 0.69	dA	dC	48.4 ± 0.01	dC	dC	44.8 ± 0.45
dT	d**1**	53.7 ± 0.46	dA	dG	55.4 ± 0.20	dG	dG	50.5 ± 0.12

[a] Conditions: 10 mM PIPES, 10 mM $MgCl_2$, 100 mM NaCl, pH 7.

(T_m) of the duplex **7** with X,Y = G:C, A:T, or **1**:**1** (Table 1). The **1**:**1** hydrophobic base pair stabilizes the DNA duplex relative to

```
5'−GCGATGXGTAGCG
3'−CGCTACYCATCGC
```

7

an A:T or G:C base pair by 3.4 and 0.8 °C, respectively, in 100 mM NaCl at pH 7. The specificity of the stabilizing interactions is demonstrated by the decreased stability of the mispairs of **1** with native bases. Relative to the **1**:**1** pair, the mispairs between **1** and dC, dT, dG, or dA are destabilized by 11.2, 8.9, 8.1, and 7.1 °C, respectively. This thermal selectivity is equivalent to that found with the native bases in this sequence context (Table 1).

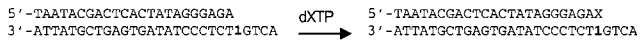
Several hydrophobic nucleobases, incapable of forming H-bonds, have been shown to be surprisingly good substrates for the Klenow fragment of *Escherichia coli* DNA polymerase I (KF).[20−22] These hydrophobic bases were designed as shape analogues of the natural bases and were efficiently incorporated opposite the corresponding native base. However, it is unclear to what extent the efficiency of incorporation was dependent on shape complementarity with a natural nucleobase.[23] In addition to being hydrophobic, **1** does not have shape complementarity to any native base, but it is incorporated opposite itself with

* To whom correspondence should be addressed. Telephone: (858) 784-7290. Fax: (858) 784-7472. E-mail: floyd@scripps.edu.

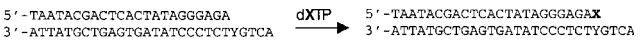
(1) Kornberg, A.; Baker, T. A. *DNA Replication*, 2nd ed.; W. H. Freeman and Co.: New York, 1992.
(2) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489−10496.
(3) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537−539.
(4) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322−8323.
(5) Lutz, M. J.; Horlacher, J.; Benner, S. A. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1149−1152.
(6) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33−37.
(7) Horlacher, J.; Hottiger, M.; Podust, V. N.; Hübscher, U.; Benner, S. A. *Proc. Natl. Acad Sci. U.S.A.* **1995**, *92*, 6329−6333.
(8) Lutz, M. J.; Held, H. A.; Hottiger, M.; Hübscher, U.; Benner, S. A. *Nucleic Acids Res.* **1996**, *24*, 1308−1313.
(9) Privalov, P. L.; Gill, S. J. *Pure Appl. Chem.* **1989**, *61*, 1097−1104.
(10) Kellis, J. T. J.; Nyberg, K.; Sali, D. *Nature* **1988**, *333*, 784−186.
(11) Dill, K. A. *Biochemistry* **1990**, *29*, 7133−7155.
(12) Schweitzer, B. A.; Kool, E. T. *J. Am. Chem. Soc.* **1995**, *117*, 1863−1872.
(13) Zhang, P.; Johnson, W. T.; Klewer, D.; Paul, N.; Hoops, G.; Davisson, V. J.; Bergstrom, D. E. *Nucleic Acids Res.* **1998**, *26*, 2208−2215.
(14) Lin, K.-Y.; Jones, R. J.; Matteucci, M. *J. Am. Chem. Soc.* **1995**, *117*, 3873−3874.
(15) Guckian, K. M.; Krugh, T. R.; Kool, E. T. *Nature Struct. Biol.* **1998**, *5*, 954−959.
(16) Bischofberger, N.; Matteucci, M. D. *J. Am. Chem. Soc.* **1989**, *111*, 3041−3046.
(17) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473−1487.
(18) Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191−6192.
(19) These hydrophobic "bases" are not actually basic, but we refer to them as "bases" or "nucleobases" by analogy to their natural counterparts.
(20) Moran, S.; Ren, R. X.-F.; Rumney, S. I.; Kool, E. T. *J. Am. Chem. Soc.* **1997**, *119*, 2056−2057.
(21) Morales, J. C.; Kool, E. T. *Nature Struct. Biol.* **1998**, *5*, 950−954.
(22) Moran, S.; Ren, R. X.-F.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10506−10511.
(23) Kool, E. T. *Biopolymers* **1998**, *48*, 3−17.

10.1021/ja9925150 CCC: $18.00
Published on Web 11/25/1999

Table 2. Steady-State Kinetic Constants for exo$^-$ KF-Mediated Synthesis of DNA Containing **1**[a]



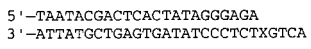
dXTP	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
d**1**TP	0.90 ± 0.04	3.7 ± 0.7	(2.4 ± 0.7) × 10^5
dATP	0.17 ± 0.07	92 ± 30	(1.8 ± 0.4) × 10^3
dGTP	0.12 ± 0.04	69 ± 22	(1.7 ± 0.4) × 10^3
dCTP	0.22 ± 0.12	123 ± 81	(1.8 ± 0.5) × 10^3
dTTP	0.89 ± 0.15	76 ± 36	(1.2 ± 0.2) × 10^4



X	Y	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})[b]
1	A	0.40 ± 0.10	20 ± 10	(2.0 ± 1.0) × 10^4
1	G	0.09 ± 0.01	9.3 ± 3.6	(9.7 ± 3.7) × 10^3
1	C	1.05 ± 0.04	2.6 ± 0.7	(4.0 ± 1.0) × 10^5
1	T	2.80 ± 0.10	9.4 ± 2.0	(3.0 ± 0.6) × 10^5
A	T	163 ± 7	3.5 ± 1.0	4.7 × 10^7

[a] Assay conditions were as follows: 40 nM template−primer duplex, 0.11−1.34 nM enzyme, 50 mM Tris buffer (pH 7.5), 10 mM $MgCl_2$, 1 μM DTT, 50 μg/mL BSA. The reactions were initiated by adding the DNA−enzyme mixture to an equal volume (5 μL) of a 2× dXTP stock solution, incubated at room temperature for 1−10 min, and quenched by the addition of 20 μL of loading buffer (95% formamide, 20 mM EDTA). A 5-μL portion of the reaction mixture was then analyzed by 15% polyacrylamide gel electrophoresis containing 8 M urea. Radioactivity was quantified using a PhosphorImager (Molecular Dynamics), with overnight exposures and the ImageQuant program. The data were obtained using a nonlinear fit of the Michaelis−Menten equation. The data presented here are an average of four independent determinations. [b] For comparison, the catalytic efficiency for synthesis of A:T and G:C base pairs is 10^7−10^8 M^{-1} min^{-1}, while misincorporation of G, C, A, and T is catalyzed at a rate of approximately 10^2 M^{-1} min^{-1}.[25,28]

reasonable efficiency by KF. Steady-state kinetic experiments were conducted with exonuclease deficient KF.[24] Kinetic constants for the single nucleotide extension of primer/duplex **8** with X = G, A, T, C, or **1** are reported in Table 2. The kinetic selectivity



8

of base **1** for self-pairing in this sequence context is evident from the relative efficiencies (k_{cat}/K_M) for incorporation of d**1**TP opposite **1** in the template compared to that for the natural triphosphates opposite **1** (Table 2). The insertion of d**1**TP opposite C or G is competitive with the insertion of d**1**TP opposite itself. However, the important comparisons are the relative efficiencies for the synthesis of the correct base pairs (i.e., dATP opposite T, Table 2) compared to the synthesis of the mispairs between d**1**TP and a native base. These relative efficiencies indicate that d**1**TP incorporation is not competitive (within a factor of 10^3) with the insertion of any natural nucleoside triphosphate opposite its natural Watson−Crick partner. Therefore, faithful replication of DNA, containing native pairs in addition to the self-pair of **1**, is favored over all possible mispairs by a factor of 20−1000. While not of the order of magnitude of the selectivity observed among the natural bases, we emphasize that **1**:**1** is the most "orthogonal" base pair, relative to native bases, yet reported.

The ability of KF to continue DNA synthesis after incorporation of **1** was also examined. Although KF inserts d**1**TP opposite **1** with reasonable efficiency, continued synthesis proceeds inefficiently (data not shown). It seems likely that this results from the unnatural pair assuming a geometry that inappropriately positions the 3′-hydroxyl group of the growing strands for nucleophilic reaction with the incoming nucleoside triphosphate. It has been suggested that polymerases,[24] including KF,[25] more efficiently extend purine:pyrimidine mispairs relative to purine:purine or pyrimidine:pyrimidine mispairs, and the inability of KF to efficiently extend the **1**:**1** base pair may be related to the pyrimidine-like ring architecture of isocarbostyril.

This study demonstrates that hydrophobicity is a sufficient driving force for the stable and selective pairing of **1** in duplex DNA, as well as for the selective enzymatic incorporation of **1** against itself during DNA synthesis. We are currently exploring other natural and in vitro evolved polymerases, as well as further modifications of the hydrophobic nucleobase, which may allow for efficient continued synthesis of DNA after incorporation of the hydrophobic base. This hydrophobicity-based approach may aid efforts to expand the genetic code by circumventing base tautomerization, which has been a major obstacle in expanding the genetic alphabet[7,26,27] (but which may naturally provide a mechanism for mutation). The use of an unnatural homobase pair does not compromise an expanded genetic alphabet, and in fact, it reduces by half the number of thermodynamic and kinetic parameters that need to be orthogonal to the native bases. We have initiated NMR experiments in an effort to determine the geometry of the **1**:**1** base pair, as well as the effects, if any, that these unnatural bases have on the structure of the DNA helix.

Acknowledgment. D.L.M. and A.K.O. contributed equally to this work. Funding was provided by the Skaggs Institute for Chemical Biology (P.G.S. and F.E.R.) and a National Institutes of Health postdoctoral fellowship (F32 GM19833-01, A.K.O.).

Supporting Information Available: Experimental procedures and characterizations (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA9925150

(24) Goodman, M. F.; Creighton, S.; Bloom, L. B.; Petruska, J. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 83−126.

(25) Kuchta, R. D.; Benkovic, P.; Benkovic, S. J. *Biochemistry* **1988**, *27*, 6716−6725.

(26) Robinson, H.; Gao, Y.-G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H.-J. *Biochemistry* **1998**, *37*, 10897−10905.

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

(28) Kuchta, R. D.; Mizrahi, V.; Benkovic, P. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 8410−8417.