Efforts toward the Expansion of the Genetic Alphabet: Information Storage and Replication with Unnatural Hydrophobic Base Pairs

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · MARCH 2000

Impact Factor: 12.11 · DOI: 10.1021/ja9940064

6 AUTHORS, INCLUDING:

Yiqin Wu
The Scripps Research Institute
25 PUBLICATIONS **1,260** CITATIONS
SEE PROFILE

David J Liu
Heliosense Biotechnologies Inc.
16 PUBLICATIONS **4,453** CITATIONS
SEE PROFILE

Efforts toward the Expansion of the Genetic Alphabet: Information Storage and Replication with Unnatural Hydrophobic Base Pairs

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

Contribution from the Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037

Received November 15, 1999

Abstract: The faithful recognition of the interstrand hydrogen bonds between complementary nucleobases forms the foundation of the genetic code. The ability to replicate DNA containing a stable third base pair would allow for an expansion of the information content of DNA by supplementing the existing two base pairs of the genetic alphabet with a third. We report the optimization of unnatural nucleobases whose pairing in duplex DNA is based on interbase hydrophobic interactions. We show that the stability and selectivity of such unnatural base pairs may be comparable to, or even exceed, that of native pairs. We also demonstrate that several unnatural base pairs are incorporated into DNA by Klenow fragment of *Escherichia coli* DNA polymerase I with an efficiency equivalent to that of native DNA synthesis. Moreover, the unnatural bases are orthogonal to the native bases, with correct pairing being favored by at least an order of magnitude relative to mispairing.

1. Introduction

Biological information storage and replication is based on a genetic alphabet encoded by the specific Watson−Crick hydrogen bonding (H-bonding) patterns of the adenine:thymine (A:T) and guanine:cytosine (G:C) base pairs. The faithful recognition of these H-bonds underscores virtually all nucleic acid biochemistry, including DNA structure and replication.[1] We are interested in increasing the information content of DNA with unnatural nucleic acids, where the pairing of the unnatural bases is driven by hydrophobic interactions instead of H-bonding. We have recently shown that hydrophobic interactions can thermodynamically compensate for the interstrand Watson−Crick H-bonds in duplex DNA.[2] Such bases would represent an expanded genetic alphabet, provided a polymerase is capable of efficient and high fidelity synthesis of DNA containing the unnatural pair. The ability to replicate DNA containing a third base pair would also allow for the enzymatic incorporation of additional functional moieties into DNA (e.g., catalytic groups, fluorophores, spin labels, etc.), allowing for increased functionality in in vitro selection experiments. In addition, the availability of a number of orthogonal base pairs would facilitate hybridization or encoding experiments in cases where natural sequences cross hybridize.

Essential thermodynamic requirements for third base pair candidates include stable pairing in duplex DNA and thermodynamic selectivity against mispairing with a natural base (thermodynamic orthogonality). The stability and selectivity of an unnatural base pair should be of the same order of magnitude as those for native base pairs. Essential kinetic properties for the third base pair candidates include efficient enzymatic incorporation into DNA by a polymerase, kinetic selectivity against enzymatic mispairing with a natural base (kinetic orthogonality), and the efficient continued primer extension after synthesis of the nascent unnatural base pair.

Previous efforts to generate orthogonal base pairs have relied on H-bonding patterns which are not found with the canonical Watson−Crick pairs. For example, the noncanonical H-bonding base pairs formed between dκ and dX[3−6] or d-*iso*-C and d-*iso*-G[7−11] adopt standard Watson−Crick geometry, but have an H-bonding pattern unlike that found in the natural base pairs, dG:dC and dA:dT. When incorporated into B-form DNA or DNA−RNA hybrid duplexes, these unnatural base pairs display moderate stability and selectivity.[4,5,12,13] Moreover, no DNA polymerase has been identified that can incorporate these base pairs into DNA with high efficiency and fidelity both in the template and as incoming triphosphate. In all cases, the modified bases were not kinetically orthogonal and instead competitively paired with natural bases during polymerase-catalyzed DNA synthesis.[3−5,9,11,14] Tautomeric isomerism, which would alter

* To whom correspondence should be adressed. 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 Company: New York, 1992.
(2) McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999** *121,* 11585−11586.
(3) Lutz, M. J.; Held, H. A.; Hottiger, M.; Hübscher, U.; Benner, S. A. *Nucleic Acids Res.* **1996**, *24*, 1308−1313.
(4) Horlacher, J.; Hottiger, M.; Podust, V. N.; Hübscher, U.; Benner, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6329−6333.
(5) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33−37.
(6) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537−539.
(7) Sugiyama, H.; Ikeda, S.; Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 9994−9995.
(8) Tor, Y.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 4461−4467.
(9) Lutz, M. J.; Horlacher, J.; Benner, S. A. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 499−504.
(10) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322−8323.
(11) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochem.* **1993**, *32*, 10489−10496.
(12) Horn, T.; Chang, C.-A.; Collins, M. L. *Nucleosides Nucleotides* **1995**, *14*, 1023−1026.
(13) Roberts, C.; Bandaru, R.; Switzer, C. *J. Am. Chem. Soc.* **1997**, *119*, 4640−4649.

10.1021/ja9940064 CCC: $19.00
Published on Web 03/24/2000

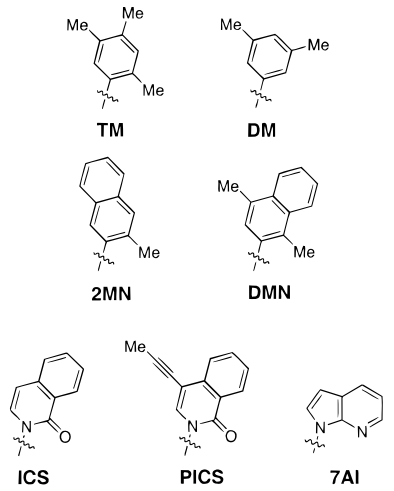


Figure 1.

H-bond donor and acceptor patterns, likely contributes to this kinetic infidelity.[13,15,16]

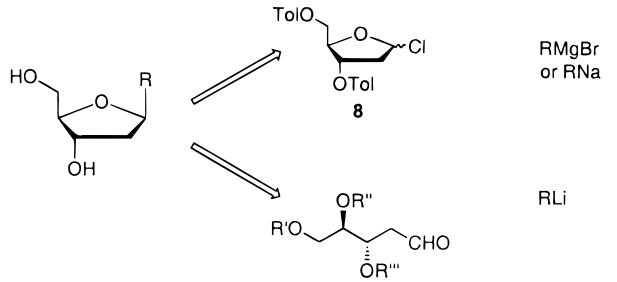
As an alternate approach we have been evaluating hydrophobic bases that are incapable of tautomerization.[2] Hydrophobicity should be a strong and selective force in the DNA duplex, based on the known importance of the hydrophobic effect in the folding and stability of proteins.[17] Desolvation of a native base when paired with an unnatural hydrophobic base should significantly disfavor mispairing. Moreover, the structure of the hydrophobic bases may be optimized to maximize interstrand packing interactions. We report here a detailed examination and optimization of interstrand hydrophobic packing in a series of hydrophobic base pairs with the intent of establishing a system of information storage, replication, and retrieval that is orthogonal to the natural H-bond-based system.

2. Results

1.1. Base Pair Design Considerations. In an effort to design two simple hydrophobic moieties that pack well when paired opposite one another in duplex DNA, several substituted phenyl rings were modeled as nucleobases[18] in B-form DNA.[19] Several bases were identified that could be accommodated without significant distortion of the helix. The trimethylphenyl nucleoside:dimethylphenyl nucleoside (**TM**:**DM**) pair was found to have the lowest calculated strain energy and was therefore chosen as the experimental starting point (Figure 1). To optimize the stability and selectivity of the unnatural pair, a variety of **TM** and **DM** derivatives were also synthesized. Modifications included increased aromatic surface area, alkyl group substituents, and the inclusion of minor groove H-bond acceptors (Figure 1). All of these bases have been evaluated with respect to their thermodynamic base pair stability in duplex DNA and the kinetics of enzymatic DNA replication.

2.2. Base Pair Synthesis. The syntheses of nucleosides **1**−**7**[20] involved two strategies for stereoselective glycosidic bond formation (Scheme 1). In the first, chloroglycoside **8**[21] served as a general electrophile for aryl Grignard additions and nitrogen nucleophiles.[22] Although Grignard addition to **8** generally favored the α-anomer, acid treatment of the corresponding benzylic ethers in refluxing xylene resulted in epimerization at C1′ to yield the *β*-anomer. This methodology, previously employed by Kool and co-workers,[22] was amenable to the large-scale nucleoside synthesis. The nucleophilic addition of indoles to **8**, by contrast, afforded the desired *β*-anomer, exclusively, which allowed the facile synthesis of azaindole nucleosides. The syntheses of aryl nucleosides **1** and **2** proceeded from condensation of in situ generated Grignard reagents with **8**, followed by epimerization and methoxide-mediated deprotection (Scheme 2).[22]

Scheme 1[a]



[a] Nucleobase is labeled R, and protecting groups are labeled R′, R″, and R‴ (see Experimental Section).

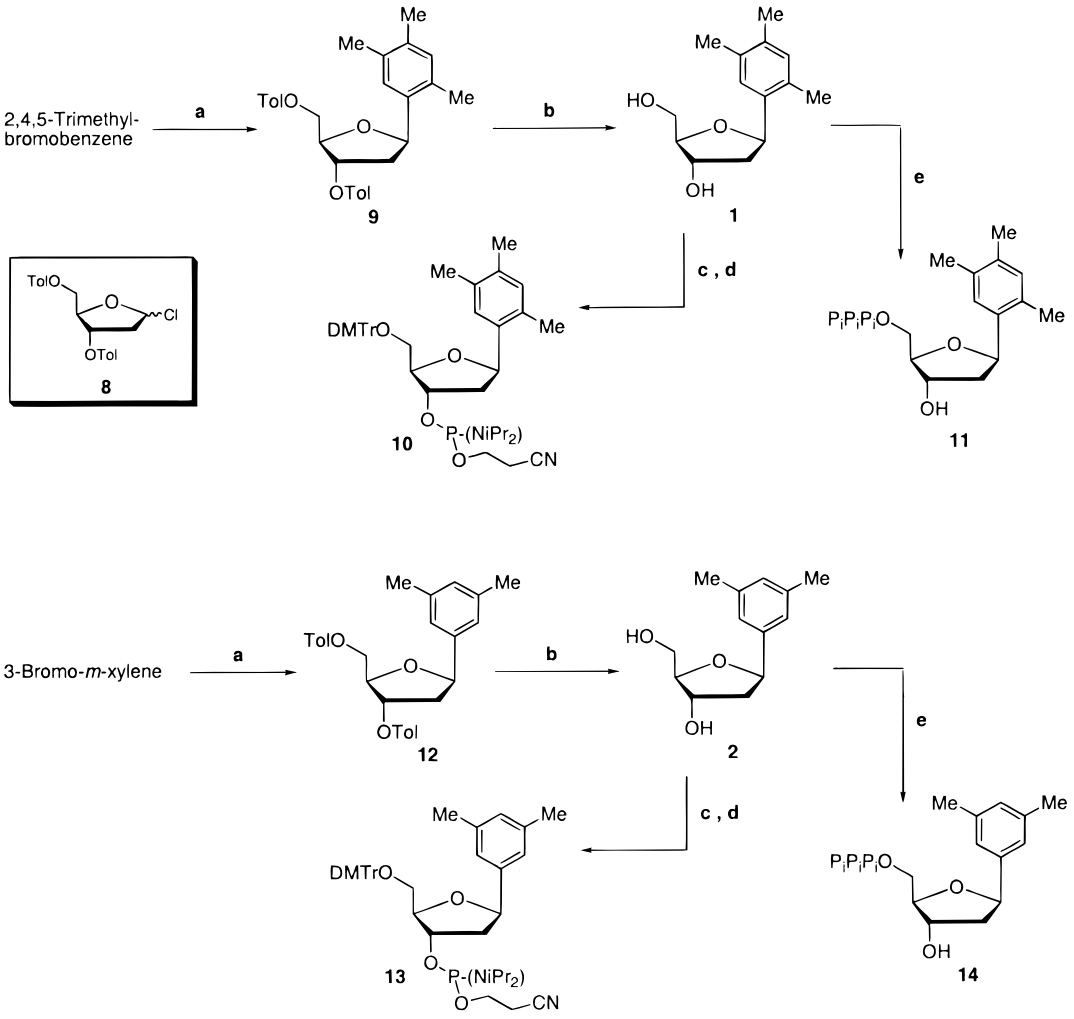
For some nucleosides, Grignard formation or condensation with **8** proved difficult. In these cases an alternate synthetic strategy was used that involved alkylation of an aldehyde substrate (**15** and **20**), followed by cyclization to the nucleoside (Scheme 1).[23,24] The seven-step synthetic route to **15** from D-ribose followed literature precedent and did not require the generation of any stereocenters beyond C1′.[24] A separate, five-step route route to *C*-nucleosides from aldehyde **20** again followed from literature precedent,[23] in which C3′ was generated from Felkin−Anh addition of an allylzinc nucleophile to isopropylidine-protected glyceraldehyde with high diastereoselectivity. Mesylation, followed by treatment with excess trifluoroacetic acid, afforded a diastereomeric mixture of separable *C*-nucleosides.

Naphthalene-substituted nucleosides **3** and **4** were derived from the condensation of the aryllithium species onto aldehydes **15** and **20**, respectively (Schemes 3 and 4).[23,24] In situ cyclization of the mesylate afforded a ∼1:1 mixture of diastereomers from which **16** was isolated. Detritylation, followed by treatment with DDQ, afforded nucleoside **3** in 55% yield. An alternative method for deprotection of the PMB-group using trityl tetrafluoroborate proceeded in comparable yield. The synthesis of the 2-methylnaphthalene nucleoside **4** involved the two-step mesylation-acid treatment previously used by Hopkins and co-workers and afforded **4** in 24% yield.

Vorbruggen glycosylation methodology was used to synthesize nucleoside **23**, as it generally resulted in *N*-glycosylation to exclusively yield the *β*-anomer.[25] However, in the case of **23**, the reaction produced a 1:1 mixture of separable diastere-

(14) Lutz, M. J.; Horlacher, J.; Benner, S. A. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1149−1152.
(15) Roberts, C.; Chaput, J. C.; Switzer, C. *Chem. Biol.* **1997**, *4*, 899−908.
(16) Robinson, H.; Gao, Y.-G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H.-J. *Biochemistry* **1998**, *37*, 10897−10905.
(17) Dill, K. A. *Biochemistry* **1990**, *29*, 7133−7155.
(18) These are obviously not basic, h., we refer to all of them as "bases" or "nucleobases" by analogy to their natural counterparts.
(19) Modeling was done with Insight II and Discover from Biosym/MSI of San Diego, CA.
(20) C1′-stereochemistry assigned from 2D-COSY/NOESY assignments.
(21) Takeshita, M.; Chang, C.-N.; Johnson, F.; Will, S.; Grollman, A. P. *J. Biol. Chem.* **1987**, *262*, 10171−10179.
(22) Schweitzer, B. A.; Kool, E. T. *J. Org. Chem.* **1994**, *59*, 7238−7242.
(23) Solomon, M. S.; Hopkins, P. B. *J. Org. Chem.* **1993**, *58*, 2232−2243.
(24) Eaton, M. A. W.; Millican, T. A. *J. Chem. Soc., Perkin Trans. 1* **1988**, 545−547.
(25) Niedballa, U.; Vorbruggen, H. *J. Org. Chem.* **1974**, *39*, 3654−3660.

Scheme 2[a]



[a] **(a)** Mg^0, then **8**; **(b)** NaOMe; (c) DMTr-Cl, NEt_3; (d) Cl-P(N*i*Pr_2)(OCH_2CH_2CN), NEt_3; (e) $POCl_3$, 0 °C, then Bu_3N-PP_i.

omers (Scheme 5). Subsequent purification and deprotection yielded **5** in 89% overall yield. Our synthesis of hydrophobic purine analogue **7** again took advantage of the versatile chloroglycoside **8**. Following literature procedure,[26] the addition of the 7-azaindole sodium salt to **8** afforded a single diastereomer in 60% yield. Removal of the toluoyl groups with sodium methoxide subsequently produced the desired nucleoside **7**.

The assignment of β-stereochemistry at C1′ for each free nucleoside was based on NOESY data, in which H1′ showed cross-peaks with both H4′ and α-H2′.[27] In addition, 1-D ^{1}H NMR splitting patterns for H1′ generally conformed to literature precedent in which a doublet of doublets indicated β-stereochemistry for *C*-nucleosides, and a pseudo-triplet indicated likewise for *N*-nucleosides. In all cases, conversion of free nucleosides to the corresponding phosphoramidites utilized standard literature procedures.[28] In addition, nucleosides were converted to triphosphates following the methodology of Otvos (Schemes 6 and 7).[29]

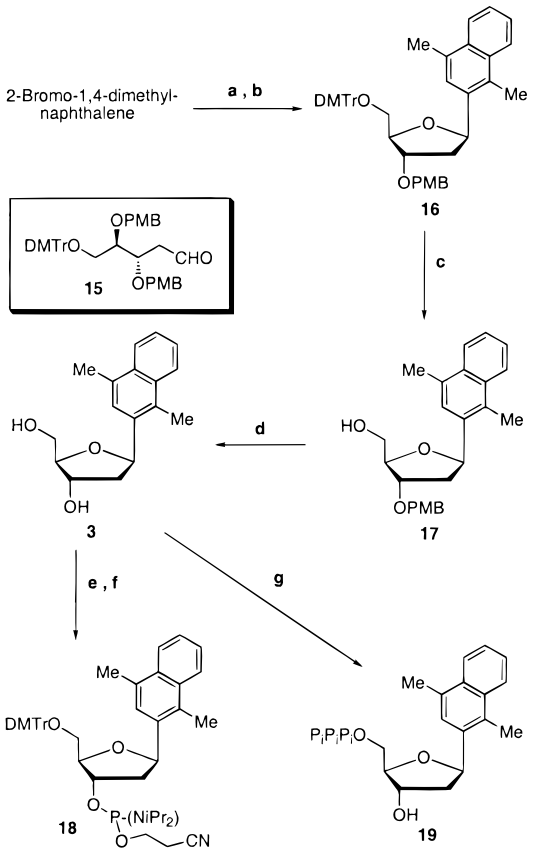
2.3. Stability of Unnatural Hydrophobic Base Pairs. To evaluate the thermodynamic stability of these novel hydrophobic bases they were incorporated into complementary oligonucleotides, 5′-GCGATGXGTAGCG-3′ and 5′-CGCTACYCATCGC-3′. The stability of the base pairs, as well as the mispairs with natural bases, was determined in this sequence context by duplex melting experiments. We use the terms "thermal stability" and "thermodynamic stability" interchangeably to refer only to the duplex stability as measured by the duplex melting temperature (T_m).

2.3.1. Stability and Thermodynamic Orthogonality of the DM:TM Unnatural Hydrophobic Base Pair. Initially, we evaluated the simple methyl-substituted phenyl bases, **TM** and **DM**. Duplex DNA containing the **TM**:**DM**, **TM**:**TM**, and **DM**:**DM** unnatural base pairs was less stable than that containing a dA:dT pair by 4.0−6.4 °C (Table 1). Despite this decrease in pairing stability, the hydrophobic bases displayed thermal selectivity against mispairing with any of the native bases (Table 1). For both **TM** and **DM**, the mispairs with dG, dA, and dT were all significantly less stable, and the mispair with dC was the least stable. The thermodynamic preference for the **TM**:**TM**, **DM**:**DM** and the **TM**:**DM** pairs relative to the most stable mispairs ranged from 2.8 to 5.2 °C. For comparison, in the same duplex the thermodynamic orthogonality of the dA:dT base pairs relative to mispairs with G, A, T, and C was 3.8−10.5 °C.

2.3.2. Stability and Thermodynamic Orthogonality of Unnatural Nucleobases with Increased Hydrophobic Surface Area. In an effort to increase the overall stability of the hydrophobic base pairs, while maintaining or increasing ther-

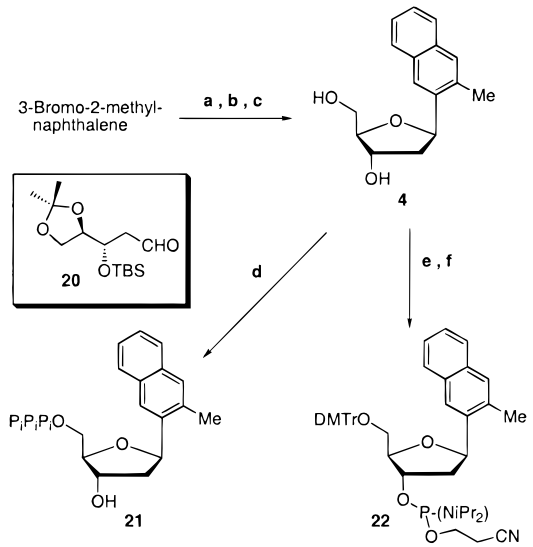
(26) Seela, F.; Gumbiowski, R. *Heterocycyles* **1989**, *29*, 795−805.
(27) Assignment of α-H2′ was based on a smaller corresponding cross-peak to H3′.
(28) Schweitzer, B. A.; Kool, E. T. *J. Am. Chem. Soc.* **1995**, *117*, 1863−1872.
(29) Kovacs, T.; Otvos, L. *Tetrahedron Lett.* **1988**, *29*.

Scheme 3[a]



[a] (a) *n*BuLi, −78 °C, then **15**; (b) MsCl, NEt_3, pyridine, 0 °C; (c) AcOH; (d) DDQ; (e) DMTr-Cl, NEt_3; (f) Cl-P(N*i*Pr$_2$)(OCH$_2$CH$_2$CN), NEt_3; (g) $POCl_3$, 0 °C, then Bu_3N-PP$_i$.

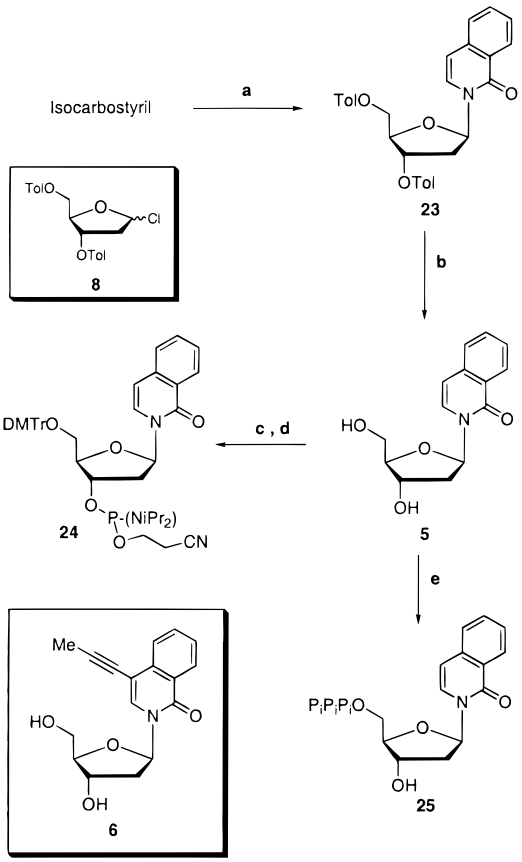
Scheme 4[a]



[a] (a) *n*BuLi, −78 °C, then **20**; (b) MsCl, 0 °C; (c) TFA; (d) $POCl_3$, 0 °C, then Bu_3N-PP$_i$; (e) DMTr-Cl, NEt_3; (f) Cl-P(N*i*Pr$_2$)(OCH$_2$CH$_2$CN), NEt_3.

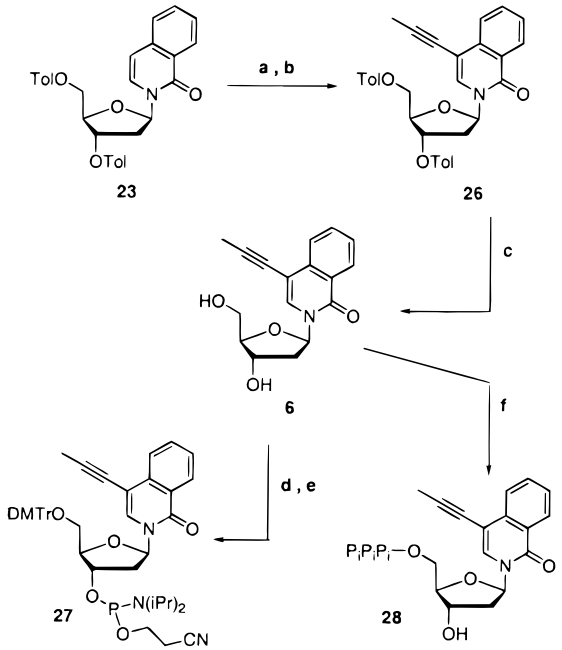
modynamic selectivity, we designed analogues of **DM** and **TM** with increased aromatic surface area. These analogues, dimethylnaphthyl nucleoside (**DMN**) and 2-methylnaphthyl nucleoside (**2MN**), were designed to probe the effect of increased aromatic surface area in different regions of the major groove. Molecular modeling studies suggested that the aromatic ring of **DMN** should project into the major groove toward the pairing

Scheme 5[a]



[a] (a) TMSN=C(OTMS)CH$_3$, then **8**, $SnCl_4$; (b) NaOMe; (c) DMTr-Cl, NEt_3; (d) Cl-P(N*i*Pr$_2$)(OCH$_2$CH$_2$CN), NEt_3; (e) $POCl_3$, 0 °C, then Bu_3N-PP$_i$.

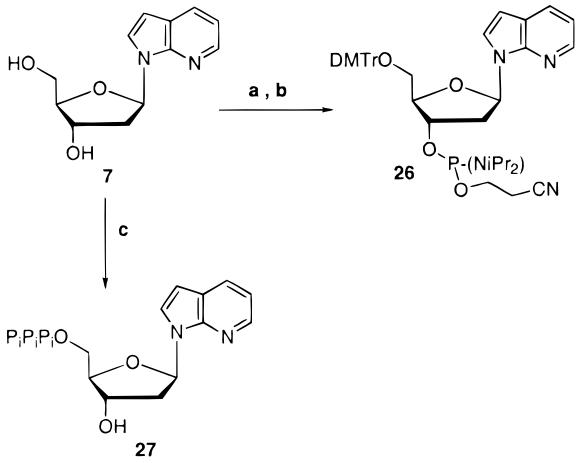
Scheme 6[a]



[a] (a) ICl; (b) CuI, Pd(II)(PPh$_3$)$_2$Cl$_2$, propyne, NEt_3; (c) NaOMe; (d) DMTr-Cl, NEt_3; (e) Cl-P(N*i*Pr$_2$)(OCH$_2$CH$_2$CN), NEt_3; (f) $POCl_3$, 0 °C, then Bu_3N-PP$_i$.

base, in a fashion similar to the methyl group of **DM**.[30] Additionally, to prevent syn−anti isomerization around the glycosidic bond, **DMN** has a methyl group at C2 that should

Scheme 7[a]



[a] (a) DMTr-Cl, NEt_3; (b) Cl-P(N*i*Pr$_2$)(OCH$_2$CH$_2$CN), NEt_3; (c) $POCl_3$, 0 °C, then Bu_3N-PP$_i$.

Table 1. T_m Values for Duplex Containing DM and TM[a]

5′-dGCGTACXCATGCG
3′-dCGCATGYGTACGC

X	Y	T_m (°C)	X	Y	T_m (°C)
TM	TM	55.2	DM	TM	53.8
	DM	52.8		DM	53.7
	A	51.7		A	48.5
	T	49.2		T	48.2
	G	50.0		G	48.7
	C	44.7		C	43.8
A	T	59.2	G	C	61.8

[a] See text for experimental details.

Table 2. T_m Values for Duplex Containing DMN and 2MN[a]

5′-dGCGTACXCATGCG
3′-dCGCATGYGTACGC

X	Y	T_m (°C)	X	Y	T_m (°C)
DMN	TM	55.0	2MN	TM	56.3
	DM	54.5		DM	54.7
	DMN	56.5		2MN	55.0
	2MN	54.5		ICS	55.4
	ICS	58.3		7AI	54.9
	7AI	55.8		A	52.3
	A	53.2		T	50.2
	T	51.5		G	52.5
	G	53.7		C	47.2
	C	49.2			

[a] See text for experimental details.

reside in the minor groove. The syn isomer should be destabilized due to eclipsing interactions between the methyl group and the ribose O4. In the case of **2MN**, the aromatic ring is expected to extend into the major groove, away from the pairing base, in a mode more analogous to the **TM**-substitution pattern.

Both the **DMN**:**DM** and **DMN**:**TM** base pairs were more stable than the **DM**:**DM** and **DM**:**TM** pairs by approximately 1 °C (Table 2). However, thermodynamic selectivity of the unnatural base pairs containing **DMN** was compromised relative to the monocyclic analogues. For example, the most stable mispair, **DMN**:dG, was only 1.3 °C less stable than the **DMN**:**TM** pair and only 0.8 °C less stable than the **DMN**:**DM** pair. Analysis of the 2-methylnaphthyl (**2MN**) **TM**-analogue also suggested that there is a correlation between increased aromatic surface area and duplex stability (Table 2). The **2MN**:**DM** and **2MN**:**TM** base pairs were more stable than the **TM**:**DM** and **TM**:**TM** pairs by approximately 1.9 and 1.1 °C, respectively. Unlike **DMN**, a comparison between the **2MN** pairs with other hydrophobic bases and mispairs with native bases showed that **2MN** retained the previously observed degree of thermodynamic selectivity ($\Delta T_m = 2-5$ °C), while improving the overall duplex stability.

2.3.3. Stability and Thermodynamic Orthogonality of Unnatural Nucleobases with Increased Polarizability. The design of the isocarbostyril nucleoside (**ICS**) paralleled that of **DMN**, with the notable exception of an *N*-glycosidic linkage and a minor groove carbonyl group provided by the endocyclic amide moiety.[30] Comparison of **ICS** with its carbocyclic analogue **DMN** suggests that the increased polarity provided by the amide functionality resulted in a marginal (opposite **DM**) to significant (opposite **TM** and **DMN**) increase in duplex stability (Table 3). The **ICS:ICS** self-pair had a T_m equal to that of a dA:dT pair (59.3 °C for **ICS**:**ICS** vs 59.2 °C for dA:dT). In comparison to the **ICS** self-pair, the hybrid base pair, **DMN**:**ICS**, was slightly less stable ($\Delta T_m = -1$ °C) and provided a measure of the thermodynamic contribution of the amide. Due to the stability of the **ICS**:dA mispair (55.1 °C), the thermodynamic orthogonality of the unnatural base pairs between **ICS** and **DM** and **TM** is compromised. However, the observed thermodynamic selectivity favoring the **ICS** pair with itself or with **DMN** was found to be 4.2 and 3.2 °C, respectively.

Our design of **PICS** capitalized on the stabilizing effect that C5-hydrophobic substituents are known to have with dC and dT.[31] However, the C5-propynyl group of **PICS** was found to have variable effects on duplex stability. When paired with the **DM**-like ring structures (**DM** and **DMN**), the propynyl group stabilized the resulting base pairs: the **PICS**:**DM** and the **PICS**:**DMN** pairs were each 1.1 °C more stable than their corresponding **ICS** pairs, **ICS**:**DM**, and **ICS**:**DMN**, respectively. However, when paired with the **TM**-like ring structures (**TM** and **2MN**), the propynyl group destabilized the resulting base pairs: the **PICS**:**TM** and **PICS**:**2MN** base pairs were 0.7 and 0.4 °C less stable than the **ICS**:**TM** and the **ICS**:**2MN** pairs, respectively. The propynyl substituent led to increased stability in the context of the **PICS**:**PICS** self-pair, which was 3.3 °C more stable than the corresponding **ICS**:**ICS** self-pair. Significantly, the **PICS**:**PICS** self-pair was more stable than both the dA:dT and dG:dC base pairs (T_m = **PICS**:**PICS**, 62.6 °C; dA:dT, 59.2 °C; dG:dC, 61.8 °C). Moreover, the selectivity of the **PICS** self-pair vs mispairing with the natural bases (ΔT_m = 7.1−11.2 °C) compares favorably to the selectivity of a dA:dT base pair in the same sequence context ($\Delta T_m = 3.8-10.5$ °C).[2] The **PICS**:**PICS** self-pair is the first stable and thermodynamically orthogonal unnatural base pair reported.

In addition to the bases derived from substituted phenyl rings, we initiated studies based on a hydrophobic purine scaffold, 7-azaindole nucleoside (**7AI**).[26] **7AI** is expected to pack in duplex DNA with only its hydrophobic surface interacting with its partner.[30] In a fashion analogous to that of the amide moiety in **ICS** and **PICS**, N7 in **7AI** provides for nucleobase polarizability as well as a minor groove H-bond acceptor. Thermodynamic evaluation of **7AI** paired with other hydrophobic bases, and itself, yielded melting temperatures that ranged from 55 to 57.6 °C (see Table 3). Of particular interest was the **7AI**:**ICS** heteropair which was only slightly less stable than dA:dT, providing an interesting lead for further optimization. Improve-

(30) Structural discussion of DNA containing unnatural base pairs is based on analogy to native base pairs. Structural studies of DNA containing unnatural base pairs is currently underway.

(31) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473−1487.

Table 3. T_m Values for Duplex Containing ICS, PICS, and 7AI[a]

5′-dGCGTACXCATGCG
3′-dCGCATGYGTACGC

X	Y	T_m (°C)	X	Y	T_m (°C)	X	Y	T_m (°C)
ICS	TM	56.8	PICS	TM	56.1	7AI	TM	55.8
	DM	54.7		DM	55.8		DM	55.0
	7AI	56.5		DMN	60.8		ICS	57.2
	ICS	59.3		2MN	57.4		7AI	55.5
	A	55.1		7AI	56.8		A	52.5
	T	53.0		ICS	60.0		T	50.5
	G	51.0		PICS	62.6		G	51.5
	C	52.2		C	48.5			

[a] See text for experimental details.

Table 4. Steady-State Kinetic Constants for KF-Mediated Synthesis of DNA Containing DM and TM in the Template[a]

5′-dTAATACGACTCACTATAGGGAGA
3′-dATTATGCTGAGTGATATCCCTCTXGTCA

template (X)	nucleoside triphosphate	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
DM	DM	1.0 ± 0.1	359 ± 88	2.8×10^3
	TM	29 ± 2	21 ± 5	1.4×10^6
	DMN	0.3 ± 0.1	6 ± 1	5.0×10^4
	2MN	131 ± 2	5.9 ± 0.3	2.2×10^7
	7AI	20.8 ± 0.5	66 ± 4	3.2×10^4
	ICS	17.1 ± 0.5	47 ± 4	3.6×10^5
	PICS	3.8 ± 0.2	7.4 ± 1.5	5.1×10^5
	A	1.1 ± 0.1	75 ± 15	1.5×10^4
	G	n.d.[b]	n.d.[b]	$<1.0 \times 10^3$
	C	0.68 ± 0.03	307 ± 28	2.2×10^3
	T	2.9 ± 0.5	182 ± 30	1.6×10^4
TM	DM	1.14 ± 0.04	315 ± 22	3.6×10^3
	TM	30.9 ± 1.8	14 ± 3	2.2×10^6
	DMN	3.3 ± 0.2	3.0 ± 0.5	1.1×10^6
	2MN	174 ± 10	5.0 ± 0.8	3.5×10^7
	7AI	27.8 ± 0.6	32 ± 2	8.7×10^5
	ICS	27.2 ± 0.8	24 ± 3	1.1×10^6
	PICS	6.8 ± 0.2	3.9 ± 0.4	1.7×10^6
	A	6.6 ± 0.2	26 ± 5	2.5×10^5
	G	0.07 ± 0.01	140 ± 25	5.0×10^2
	C	0.18 ± 0.04	381 ± 35	4.7×10^2
	T	6.9 ± 0.6	227 ± 42	3.0×10^4

[a] See text for experimental details. [b] Rates too slow for determination of k_{cat} and K_M independently.

ment in the heteropair stability would yield excellent thermal selectivity given the modest T_m values for the mispairs between **7AI** and the native bases (48.5−52.5 °C).

2.4. Enzymatic Incorporation of Single Unnatural Nucleotides into DNA. The unnatural nucleobases were evaluated as substrates for the exonuclease-deficient Klenow fragment of *E. coli* DNA polymerase I (KF). Initial velocities were determined during [γ-^{32}P]primer extension reactions by literature methods with varying concentrations of nucleoside triphosphates.[32] The reactions were analyzed by polyacrylamide gel electrophoresis and a PhosphorImager (Molecular Dynamics) was used to quantify gel band intensities corresponding to the extended primer. The measured velocities were plotted vs [dNTP] and fit to the Michaelis−Menten equation. Unnatural nucleobases were assayed both in template DNA and as incoming triphosphate nucleotides. Steady-state kinetic parameters for single nucleotide incorporation are reported in Tables 4−7.

2.4.1. Enzymatic Synthesis of the DM:TM Unnatural Base Pair. The rates of KF-dependent insertion of d**DM**TP and d**TM**TP were approximately the same opposite either **DM** or **TM** in the template. The triphosphate d**TM**TP bound to the enzyme−DNA complex more tightly than d**DM**TP ($K_{M(apparent)}$

Table 5. Steady-state Kinetic Constants for KF-Mediated Synthesis of DNA with DMN and 2MN in the Template[a]

5′-dTAATACGACTCACTATAGGGAGA
3′-dATTATGCTGAGTGATATCCCTCTXGTCA

template (X)	nucleoside triphosphate	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
DMN	DM	1.3 ± 0.2	78 ± 25	1.7×10^4
	TM	18.0 ± 1.2	25 ± 6	7.2×10^5
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	138 ± 2	5.4 ± 0.4	2.6×10^7
	7AI	53 ± 2	13 ± 2	4.1×10^6
	ICS	17 ± 1	24 ± 1	7.1×10^5
	PICS	5.2 ± 0.5	5.7 ± 2.2	9.1×10^5
	A	3.64 ± 0.15	29 ± 4	1.3×10^5
	G	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	C	0.24 ± 0.01	170 ± 18	1.4×10^3
	T	1.5 ± 0.1	72 ± 5	2.1×10^4
2MN	DM	1.66 ± 0.07	141 ± 15	1.2×10^4
	TM	78 ± 5	10 ± 3	7.8×10^6
	DMN	11.7 ± 0.2	2.8 ± 0.3	4.2×10^6
	2MN	175 ± 10	4.0 ± 0.5	4.4×10^7
	7AI	138 ± 4	14 ± 1	9.9×10^6
	ICS	140 ± 5	12 ± 1	1.2×10^7
	PICS	86 ± 8	4.8 ± 1.8	1.8×10^7
	A	104 ± 4	30 ± 4	3.5×10^6
	G	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	C	0.25 ± 0.05	30 ± 10	8×10^3
	T	2.2 ± 0.2	181 ± 31	1.2×10^4

[a] See text for experimental details. [b] Rates too slow for determination of k_{cat} and K_M independently.

Table 6. Steady-state Kinetic Constants for KF-Mediated Synthesis of DNA with ICS, PICS, and 7AI in the Template[a]

5′-dTAATACGACTCACTATAGGGAGA
3′-dATTATGCTGAGTGATATCCCTCTXGTCA

template (X)	nucleoside triphosphate	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
ICS	DM	0.92 ± 0.1	358 ± 57	2.6×10^3
	TM	0.99 ± 0.03	34 ± 4	2.9×10^4
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	46 ± 2	87 ± 10	5.3×10^5
	7AI	11.3 ± 1.4	45 ± 17	2.5×10^5
	ICS	3.6 ± 0.1	57 ± 6	6.3×10^4
	PICS	2.6 ± 0.4	33 ± 11	7.8×10^4
	A	0.15 ± 0.01	73 ± 10	2.1×10^3
	G	0.28 ± 0.02	83 ± 12	3.4×10^3
	C	0.03 ± 0.01	34 ± 19	8.8×10^2
	T	1.2 ± 0.1	234 ± 83	5.1×10^3
PICS	DM	0.45 ± 0.08	40 ± 20	1.1×10^4
	TM	0.67 ± 0.03	9.6 ± 1.5	7.0×10^4
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	30 ± 1	64 ± 7	4.7×10^5
	7AI	4.0 ± 0.12	23 ± 2	1.7×10^5
	ICS	0.7 ± 0.03	23 ± 3	3.0×10^4
7AI	DM	0.61 ± 0.03	346 ± 38	1.7×10^3
	TM	3.1 ± 0.1	33 ± 4	9.4×10^4
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	98 ± 3	21 ± 3	4.7×10^6
	7AI	8.7 ± 0.5	40 ± 5	2.2×10^5
	ICS	7.9 ± 0.3	21 ± 3	3.8×10^5
	PICS	2.9 ± 0.2	7.5 ± 2.0	3.9×10^5
	A	0.32 ± 0.02	54 ± 10	5.9×10^3
	G	0.08 ± 0.01	49 ± 13	1.6×10^3
	C	0.17 ± 0.02	31 ± 11	5.5×10^3
	T	0.23 ± 0.05	146 ± 40	1.6×10^3

[a] See text for experimental details. [b] Rates too slow for determination of k_{cat} and K_M independently.

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

Table 7. Steady-state Kinetic Constants for KF-Mediated Synthesis of DNA with Natural Templates and Unnatural Nucleoside Triphosphates[a]

5′-dTAATACGACTCACTATAGGGAGA
3′-dATTATGCTGAGTGATATCCCTCT**X**GTCA

template (X)	nucleoside triphosphate	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
A	T	163 ± 7	3.5 ± 1.0	4.7×10^7
	DM	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	TM	6.1 ± 0.4	23 ± 5	2.7×10^5
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	144 ± 4	14 ± 2	1.0×10^7
	7AI	7.9 ± 0.8	40 ± 12	2.0×10^5
	ICS	3.0 ± 0.3	57 ± 16	5.3×10^4
	PICS	0.40 ± 0.10	20 ± 10	2.0×10^4
G	DM	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	TM	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	7AI	0.59 ± 0.02	24 ± 3	2.5×10^4
	ICS	0.43 ± 0.02	26 ± 3	1.7×10^4
	PICS	0.09 ± 0.01	9.3 ± 3.6	9.7×10^3
T	DM	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	TM	0.31 ± 0.05	151 ± 55	2.0×10^3
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	11.9 ± 0.3	142 ± 9	8.4×10^4
	7AI	3.9 ± 0.1	120 ± 7	2.8×10^4
	ICS	4.1 ± 0.2	101 ± 13	4.1×10^4
	PICS	2.80 ± 0.10	9.4 ± 2.0	3.0×10^5
C	DM	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	TM	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	DMN	n.d.[b]	n.d.[b]	$\leq 1.0 \times 10^3$
	2MN	1.2 ± 0.1	72 ± 12	1.7×10^4
	7AI	1.4 ± 0.1	96 ± 4	1.5×10^4
	ICS	1.5 ± 0.2	83 ± 27	1.8×10^4
	PICS	1.05 ± 0.04	2.6 ± 0.7	4.0×10^5

[a] See text for experimental details. [b] Rates too slow for determination of k_{cat} and K_M independently.

= 20 vs 300 μM) and was also incorporated into the growing strand more quickly ($k_{cat(apparent)}$ = 30 vs 1 min^{-1}). Tight binding and fast turnover result in the efficient insertion of d**TM**TP opposite either **DM** or **TM** in the template (10^6 M^{-1} min^{-1}).

KF did not efficiently incorporate a natural triphosphate opposite **DM** in the template (apparent $k_{cat}/K_M < 10^4$ M^{-1} min^{-1}). However, dATP was incorporated with a catalytic efficiency of 10^5 M^{-1} min^{-1} opposite **TM** in the template (10-fold less efficiently than the incorporation of d**TM**TP opposite **DM** or itself). The triphosphate of thymine was incorporated opposite **TM** with a catalytic efficiency of 10^4 M^{-1} min^{-1}, while the triphosphates of both dG and dC were only inefficiently incorporated opposite **TM** in the template.

The increased orthogonality of **DM** relative to **TM** was also evident when these unnatural bases were present as triphosphates. The triphosphate of **TM** was incorporated reasonably efficiently only opposite dA in the template (10^5 M^{-1} min^{-1}); however, d**DM**TP was not incorporated opposite any natural template. The relatively poor recognition of d**DM**TP and the reasonable efficiency of KF mediated synthesis of the mispair between **TM** and dA, in either template or triphosphate combination, limits the kinetic orthogonality of a base pair composed of **DM** and **TM**. The **TM**:**TM** self-pair is marginally kinetically orthogonal, being synthesized at least 10-fold faster than all possible mispairs.

2.4.2. Enzymatic Synthesis with Unnatural Nucleobases with Increased Hydrophobic Surface. Enzymatic synthesis with **DMN** and **2MN** was examined, to evaluate the effect of increased aromatic surface area on the recognition of hydrophobic bases by KF. A comparison of templates containing **DM** and **DMN** reveals that increased aromatic surface area of bases in the template increases the absolute rate for incorporation of d**DM**TP (d**DM**TP is incorporated an order of magnitude more efficiently opposite the bulkier template). However, the opposite was true for d**TM**TP, which was incorporated with a slightly reduced efficiency opposite **DMN** relative to **DM** in the template. A comparison of d**DM**TP and d**DMN**TP incorporation reactions reveals that increased aromatic surface area in the triphosphate favors the insertion opposite **TM** in template by 3 orders of magnitude (10^3 M^{-1} min^{-1} and 10^6 M^{-1} min^{-1} for d**DM**TP and d**DMN**TP, respectively) while not affecting the efficiency for incorporation opposite **DM** in the template. It is interesting to note that with the **DM**-like ring structures, increased steric bulk in the template (**DM** to **DMN**) is favorable for insertion of d**DM**TP and disfavorable for d**TM**TP, while the opposite is true when the increase in steric bulk is in the triphosphate. This behavior may result from the aromatic surface area of **DMN** as well as the C2-methyl group.

The rates for the KF-mediated synthesis of mispairs between **DMN** and a natural base were nearly identical to those found with **DM**, regardless of whether the unnatural base was present in the template or as the triphosphate. No mispairs were synthesized with rates greater than 10^4 M^{-1} min^{-1}, with the single exception of incorporation of dATP opposite **DMN** in the template, which proceeded with an efficiency of 10^5 M^{-1} min^{-1} (10-fold faster than the incorporation of dATP opposite **DM** in the template). The increased rate for insertion of dATP opposite **DMN** relative to opposite **DM** in the template, resulted from changes in both the apparent K_M and k_{cat}. The increased efficiency of pairing dATP with **DMN** may result from a specific interaction between adenine and the minor groove methyl group of **DMN**, not present in **DM** (vide infra).

Unlike the **DM**-ring structure discussed above, an increase in the surface area of the **TM**-like ring structure (**TM** to **2MN**) in template or triphosphate, resulted in more efficient synthesis of the unnatural base pair. **2MN** in the template strand directed the incorporation of d**DM**TP and d**TM**TP with 3-fold increased efficiency relative to **TM** in the template. The triphosphate d**2MN**TP was inserted with greater than 10-fold increased efficiency opposite **DM** and **TM** in the template, relative to d**TM**TP. The increased rates resulted from both apparent k_{cat} and K_M effects. (The data in Table 5 shows the same trend with the base pairs between **2MN** and all of the hydrophobic bases described in this manuscript.) The increased k_{cat}/K_M for the **2MN** nucleobase makes the synthesis of the **2MN**:**DM**, **2MN**:**TM**, and **2MN**:**2MN** unnatural base pairs as efficient as the synthesis of a native base pair.

Mispairs resulting from the insertion of dATP or dCTP opposite **2MN** were formed faster than opposite **TM**, by approximately 1 order of magnitude. The increased aromatic surface area of **2MN** also resulted in the more efficient incorporation of d**2MN**TP opposite natural bases in the template, relative to d**TM**TP. The efficiency of incorporation of d**2MN**TP opposite dA and dT in the template increased 30−40-fold, predominantly from an increase in the apparent k_{cat}. In fact, the incorporation of d**2MN**TP opposite dA in the template is virtually as efficient as the insertion of dTTP opposite dA in the same sequence context (Table 7). Like the **TM** self-pair described above, the **2MN**:**2MN** self-pair was marginally kinetically orthogonal, with the correct pair being synthesized 5-fold faster than all possible mispairs. However, the **2MN** self-pair is synthesized with a native-like catalytic efficiency (4.4

$\times 10^7$ M^{-1} min^{-1} for the self-pair and 4.6×10^7 M^{-1} min^{-1} for a dA:dT pair). This efficiency and orthogonality, albeit marginal, make **2MN** an interesting hydrophobic nucleobase ring structure for further optimization.

2.4.3. Enzymatic Synthesis with Unnatural Bases with Increased Polaraizability. The **ICS** and **PICS** nucleobases were designed to examine the effect of an *N*-glycosidic linkage and the presence of a hydrophilic minor groove carbonyl group, while retaining the hydrophobic packing surface found in **DMN**. The triphosphates of both **DM** and **TM** were inserted roughly an order of magnitude more slowly opposite **ICS** than they were opposite **DMN** in the template, due to both k_{cat} and K_M effects. A comparison of the rates for incorporation of d**ICS**TP or d**DMN**TP opposite **TM** in the template shows that the increase in polarizability had little effect—both bases were incorporated with an efficiency of 10^6 M^{-1} min^{-1}. However, in the case of **DM** in the template, the insertion of d**ICS**TP was significantly more efficient than the insertion of d**DMN**TP.

The native triphosphates were incorporated by KF opposite **ICS** in the template with reduced efficiency relative to **DMN** in the template. Most dramatically, incorporation of dATP was reduced 4000-fold relative to **DMN** in the template. However, several unnatural triphosphates were incorporated with reasonable efficiencies. For example, KF incorporated d**2MN**TP with an efficiency of 5.3×10^5 M^{-1} min^{-1}, which is more than two-orders of magnitude faster than the most efficient insertion of a native triphospahate, (dTTP, 5.1×10^3 M^{-1} min^{-1}). As a triphosphate, d**ICS**TP was incorporated with greater efficiency opposite the native bases in the template, relative to d**DMN**TP, but never with a catalytic efficiency greater than 5×10^4 M^{-1} min^{-1}.

PICS behaved similarly to **ICS** when in template. There were small differences arising from the slightly tighter binding and slightly smaller apparent k_{cat} for the incorporation of an unnatural triphosphates opposite **PICS**, relative to opposite **ICS**. Comparison of d**ICS**TP and d**PICS**TP shows that the addition of the propynyl group has very little effect on the efficiency of incorporation, with small but compensating changes in the apparent k_{cat} and K_M values. Like **ICS**, the insertion of d**PICS**TP opposite **2MN** is approximately as efficient as the insertion of dTTP opposite dA.

With respect to the incorporation of a natural triphosphate, **PICS** again behaved in a manner similar to that of **ICS**. No natural triphosphate was incorporated opposite **PICS** with an efficiency greater than 10^4 M^{-1} min^{-1}. When present as a triphosphate, d**PICS**TP was incorporated with an order of magnitude greater catalytic efficiency opposite dC and dT, relative to d**ICS**TP. However, the rate of misincorporation of d**PICS**TP opposite a natural template did not exceed 5×10^5 M^{-1} min^{-1}, which is 3 orders of magnitude slower than the rate of insertion of a correct, natural base. Therefore, these rates do not compromise the orthogonality of the unnatural nucleobase.

The hydrophobic purine analogue, **7AI**, was also examined as a substrate for KF. The triphosphate of **TM** was incorporated two-orders of magnitude more efficiently than **DM** opposite **7AI** in the template. The triphosphate of **7AI** was inserted more efficiently opposite **TM** relative to **DM** in all cases, by 1 to 2 orders of magnitude. The increased aromatic surface area of **DMN** relative to **DM** did not result in more efficient synthesis with the unnatural base present as either triphosphate or in the template. However, in comparing **2MN** to **TM**, the increased aromatic surface area resulted in more efficient synthesis of the unnatural base pair. The triphosphate of **2MN** was inserted 50-fold more efficiently opposite **7AI** than was d**TM**TP (4.7×10^6 M^{-1} min^{-1} vs 9.4×10^4 M^{-1} min^{-1}); and d**7AI**TP was inserted opposite **2MN** 10-fold more efficiently than opposite **TM** in the template (9.9×10^6 vs 8.7×10^5 M^{-1} min^{-1}).

The increased polarizability of d**ICS**TP, relative to that of d**DMN**TP, resulted in a significant increase in the efficiency of incorporation opposite **7AI** in the template. However, the same increase in polarizabilty of the nucleobase in the template strand resulted in less efficient insertion of d**7AI**TP. The efficiency of incorporation of d**7AI**TP decreased from 4.1×10^6 to 2.5×10^5 M^{-1} min^{-1} upon changing the unnatural base in the template from **DMN** to **ICS**. Comparison of the rate data for **ICS** and **PICS** shows that the propynyl group of **PICS** has no significant effect on the synthesis of unnatural base pairs with **7AI**.

When **7AI** was present in the template, KF did not incorporate native triphosphates with a catalytic efficiency greater than 6×10^3 M^{-1} min^{-1}. The triphosphate, d**7AI**TP, was inserted only slowly opposite dG, dC, and dT in the template (10^4 M^{-1} min^{-1}) and only slightly faster opposite dA in the template (2.0×10^5 M^{-1} min^{-1}).

3. Discussion

Efforts to expand the genetic alphabet rely on the specific thermodynamic and kinetic parameters of nucleosides containing unnatural bases. The unnatural nucleobases should form stable pairs in B-form DNA, with high selectivity relative to mispairing with the native bases. Furthermore, a DNA polymerase must be capable of efficiently synthesizing DNA containing the third base pair, and must do so with high fidelity. We have evaluated hydrophobicity, as opposed to H-bonding, as a driving force for selective information storage and replication. Previous efforts to expand the genetic alphabet by Benner and co-workers[3,5,9,11,14] were predicated on the use of unnatural H-bonding nucleobases. Such nucleobases were found to be marginally stable and themodynamically orthogonal to the native bases when incorporated into duplex DNA.[4,5,12,13] However, none of these bases were found to be kinetically orthogonal. Significant mispairing was believed to be caused, at least in part, by tautomeric forms of the bases.[3−5,9,11,14] Carbocyclic hydrophobic nucleobases cannot adopt the tautomeric structures that would compromise their orthogonality relative to the native bases. This is especially important, considering our incomplete understanding of the tautomeric equilibria of nucleobases in duplex DNA.

Kool and co-workers demonstrated that several hydrophobic nucleobases are surprisingly good substrates for KF.[33−35] These hydrophobic bases were designed as shape-analogues of the natural bases and were incorporated opposite the corresponding native base with reasonable efficiency. However, it is unclear to what extent the efficiency of incorporation was dependent on shape-complementarity with a natural nucleobase (vide infra). Kool also showed that pyrene paired opposite an abasic site in duplex DNA results in only a relatively small destabilization relative to a dA:dT pair.[36] Moreover, a technique to sequence abasic sites in DNA was developed, based on the fact that pyrene is inserted efficiently and specifically opposite sites that lack bases.[36]

We have reported the design, synthesis, and characterization of the triphosphates and phosphoramidites of seven hydrophobic nucleobases. These hydrophobic bases have been characterized thermodynamically in duplex DNA and kinetically during DNA

(33) Morales, J. C.; Kool, E. T. *Nat. Struct. Biol.* **1998**, *5*, 950−954.
(34) Moran, S.; Ren, R. X.-F.; Rumney, S. I.; Kool, E. T. *J. Am. Chem. Soc.* **1997**, *119*, 2056−2057.
(35) Moran, S.; Ren, R. X.-F.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10506−10511.
(36) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704−708.

synthesis with KF. Generally, they have been found to be both thermodynamically and kinetically orthogonal to the natural bases.

3.1. Stability and Orthogonality of Unnatural Hydrophobic Bases. To date there has not been a report of an unnatural nucleobase without H-bonds that forms a base pair as stable as native base pairs.[2] A variety of the unnatural bases described in this study, the pairing of which is based exclusively on hydrophobicity, form base pairs that are virtually as stable as native pairs. Several are actually more stable than native pairs. This demonstrates that interstrand hydrophobic packing interactions may effectively compensate for the complete removal of Watson−Crick H-bonds. The initially designed methyl-substituted phenyl bases, **DM** and **TM,** generally formed base pairs that were significantly less stable than native base pairs. The **TM**:**TM** base pair was the most stable, but was still significantly less stable than a dA:dT pair (55.2 vs 59.2 °C). Increasing the aromatic surface area of **DM** and **TM**, resulting in **DMN** and **2MN**, respectively, resulted in an increase in unnatural base pair stability in all cases. The inclusion of an amide linkage in **ICS**, did not significantly affect the stability of the base pair with **DM** (54.5 and 54.7 °C for **DMN**:**DM** and **ICS**:**DM**, respectively) but did stabilize the pair with **TM** (55.0 and 56.8 °C for **DMN**:**TM** and **ICS**:**TM**, respectively). The propinyl group of **PICS** stabilized base pairs with **DM**-like ring structures (**DM** and **DMN**), but destabilized base pairs with **TM**-like ring structure nucleobases (**TM** and **2MN**). The hydrophobic purine analogue **7AI** paired with greater stability opposite **TM** relative to opposite **DM** (55.8 and 55.0 °C, respectively).

The unnatural hydrophobic bases are also generally found to be thermodynamically orthogonal to the natural nucleobases. This is demonstrated by the consistent increased stability of the hydrophobic base pairs relative to that of the mispairs with native bases. In fact, the most stable unnatural pairs were virtually as thermodynamically selective as the native bases. This presumably results, at least in part, from the desolvation of the native bases that occurs upon their pairing with a hydrophobic base; waters of solvation are lost and are not replaced by compensating Watson−Crick H-bonds. This desolvation model is supported by the consistent decreased stability of the mispairs between the hydrophobic nucleobase and cytosine. Cytosine and guanine are the most hydrophilic of the natural nucleobases, but the hydrophilicity of guanine may be partly compensated for by its increased aromatic surface area.[37] Cytosine does not have a large aromatic surface area, and the instability of the hydrophobic base mispairs with cytosine may be a result of desolvation in conjunction with poor base-stacking.

The most important result concerning the stability of the unnatural base pairs reported in this manuscript is that H-bonds are not required for the stable and selective pairing of bases in duplex DNA. The base pairs **ICS**:**DMN**, **ICS**:**ICS**, **PICS**:**DMN**, and **PICS**:**ICS** and **PICS**:**PICS** are each at least as stable and selective as a dA:dT base pair in the same sequence context. Despite the demonstrated selectivity possessed by the unnatural bases against mispairing with the native bases, pairing between the unnatural bases is not as selective as would be desirable for an unnatural heteropair. Our efforts to understand the hydrophobic determinants of selectivity are only in their infancy, and in fact, preliminary results indicate that judicious placement of a methyl group on either the **ICS** or **7AI** ring structures allows for dramatically increased selectivity among the hydrophobic bases. However, the use of hydrophobic self-pairs avoids the problem altogether, without compromising the ability to store increased information. The **PICS** self-pair is particularly attractive and is discussed further below.

3.2. Kinetic Efficiency and Orthogonality of Unnatural Hydrophobic Bases as Polymerase Substrates. Klenow fragment of DNA Pol I is able to efficiently recognize a large number of unnatural hydrophobic bases and incorporate them into DNA. This is remarkable, considering that the ring structures and functionalities of the unnatural bases are significantly different from those of the native bases. The **TM**-like ring structure was especially well-suited for KF-catalyzed DNA synthesis. When in template, **TM** directed the incorporation of five of the seven unnatural nucleobases, d**TM**TP, d**DMN**TP, d**2MN**TP, d**ICS**TP, and d**PICS**TP with efficiencies (apparent k_{cat}/K_M) greater than or equal to 10^6 M^{-1} min^{-1}. The nucleoside triphosphate, d**TM**TP, was efficiently inserted opposite **DM**, **TM**, and **2MN** with rates greater than or equal to 10^6 M^{-1} min^{-1}. These rates are only 10-fold lower than those for the synthesis of native DNA.

Opposite all seven unnatural bases in the template, d**2MN**TP was the most efficiently incorporated unnatural triphosphate. Indeed, d**2MN**TP was incorporated opposite **DM**, **TM**, **DMN**, or **2MN** with efficiencies equal to those characteristic of native base pair synthesis (10^7 M^{-1} min^{-1}). The incorporation of d**2MN**TP opposite **7AI** is marginally less efficient (10^6 M^{-1} min^{-1}), and its incorporation opposite the isocarbostyril templates, **ICS** and **PICS**, is reduced 100-fold (10^5 M^{-1} min^{-1}). When in the template **2MN** directs the incorporation of **ICS**, **PICS**, and **7AI** with efficiencies that are within a factor of 5 of that for dA:dT synthesis. There is a segregation of the unnatural bases with respect to their KF-mediated pairing with **2MN**, depending on whether the **2MN** is present in the template or as the nucleoside triphosphate. When **2MN** is present in the template, the pairs with **ICS**, **PICS**, and **7AI** are most efficiently synthesized. However, when **2MN** is present as the triphosphate, the pairs with **DM**, **TM**, and **DMN** are more efficiently synthesized.

The base pairs composed of hydrophobic bases other than **2MN** or **TM** were generally poorer substrates for KF. The only base pair synthesized with an efficiency greater than 10^6 M^{-1} min^{-1}, which did not involve **2MN** or **TM**, was the insertion of d**7AI**TP opposite **DMN**.

Kinetic orthogonality is evident from the increased efficiency with which the unnatural hydrophobic nucleobases are paired opposite other unnatural hydrophobic bases, relative to the natural bases. With any natural base in the template, efficiencies of unnatural hydrophobic base insertion were typically less than or equal to 10^4 M^{-1} min^{-1}. Mispairs resulting from the insertion of an unnatural triphosphate are therefore formed at least 3 orders of magnitude less efficiently than native base pairs, allowing for good kinetic orthogonality with the unnatural bases as triphosphates. However, there are five exceptions where mispairs between native and hydrophobic bases were synthesized with catalytic efficiencies greater than 10^4 M^{-1} min^{-1}; d**TM**TP and d**7AI**TP were inserted opposite dA, and d**PICS**TP was inserted opposite dC or dT, with catalytic efficiencies of approximately 10^4 M^{-1} min^{-1}. The insertion of d**2MN**TP opposite dA was catalyzed with a remarkable efficiency of 10^7 M^{-1} min^{-1}.

Insertion of a native triphosphate opposite an unnatural hydrophobic bases was generally found to be less efficient than insertion of a hydrophobic base opposite another hydrophobic base. Unlike the case where the native base is in the template, when the native base is present as the nucleoside triphosphate,

(37) Shih, P.; Pedersen, L. G.; Gibbs, P. R.; Wolfenden, R. *J. Mol. Biol.* **1998**, *280*, 421−430.

there is an approximate correlation between the hydrophobicity of the native base and the efficiency of its insertion opposite the unnatural base in the template. Recall that this same trend was observed in the stability of the mispairs between hydrophobic and native bases. Correspondingly, the triphosphate of adenine, the most hydrophobic of the natural nucleobases,[37] tends to be inserted opposite the hydrophobic bases with the greatest efficiency (up to 10^5-10^6 M^{-1} min^{-1}), followed by dTTP, dGTP, and dCTP. The most efficient insertion of dATP occurred opposite **2MN**, **DMN**, and **TM** with efficiencies of 10^5-10^6 M^{-1} min^{-1}. However, hydrophobic bases are generally incorporated at least 1 order of magnitude more efficiently than are natural triphosphates opposite hydrophobic bases in the template.

The diversity of unnatural hydrophobic bases with which KF is capable of efficiently synthesizing DNA is remarkable. In addition to the efficient synthesis of unnatural DNA containing the **TM** and the **2MN** unnatural nucleobases, a variety of reactions catalyzed by KF are remarkably efficient. The insertion of d**7AI**TP opposite **DMN** was catalyzed by KF with an efficiency only 10-fold reduced relative to that for native synthesis (4.1×10^6 M^{-1} min^{-1}). The facile incorporation of dATP opposite **TM**, **DMN** and **2MN** is also fast (10^5-10^6 M^{-1} min^{-1}). The presence of a templating minor groove methyl group and a hydrophobic region of a purine are common to each of these unnatural bases. A comparison of **DMN** and **ICS** shows that the specific interactions between the incoming dATP and the templating minor groove methyl group, as opposed to minor groove amide carbonyl group, results in a 60-fold reduction in the efficiency of dATP incorporation. It may seem surprising that the substitution would be so detrimental to the enzymatic incorporation of dATP, considering that thymine possesses an equivalent carbonyl group at the analogous position. However, it is possible that a component of the transition state stabilization that favors dATP insertion opposite **DMN**, relative to **ICS**, may result from specific hydrophobic interactions between the minor groove methyl group of the unnatural base and adenine. Adenine is the only natural nucleobase that has a hydrophobic C−H group which is capable of interstand hydrophobic packing in the minor groove.

The hypothesis of a specific interaction, favoring dATP incorporation, should not be confused with the "A-rule" which hypothesizes that polymerases tend to insert dATP opposite a non-instructive site.[38,39] However, as argued by Kool and coworkers, the high efficiency of the incorporation makes the A-rule less tenable in this case.[35] The hypothesis of a specific hydrophobic interaction should also not be taken as evidence supporting a shape-complementarity mechanism for polymerase specificity.[34,35] In the majority of the data described in this manuscript, such shape-complementarity appears to play a less important role than hydrophobicity. For example, **TM** could be viewed as a shape mimic of T, and KF does exhibit a higher selectivity for the insertion of d**TM**TP opposite A (100-fold), relative to the insertion of d**TM**TP opposite other natural bases. However, d**7AI**TP is inserted with virtually the same efficiency as that with d**TM**TP opposite adenine, and d**ICS**TP is inserted only 5-fold less efficiently. There is also only a slight preference (less than 10-fold) for dATP insertion opposite **TM**, relative to the insertion of the other natural triphosphates. Moreover, dATP is inserted opposite **DMN** in the template with only 5-fold reduced efficiency relative to its insertion opposite **TM** in the template. Other unnatural hydrophobic bases, which have no shape resemblance to adenine, were incorporated more efficiently opposite **TM** than was dATP. Therefore, although shape may contribute to KF specificity, it is apparent that hydrophobicity plays a more dominant role.

(38) Drew, H. R.; Travers, A. A. *Cell* **1988**, *37*.
(39) Randall, S. K.; Eritja, R.; Kalplan, B. E.; Petruska, J.; Goodman, M. F. *J. Biol. Chem.* **1987**, *262*, 6864−6870.

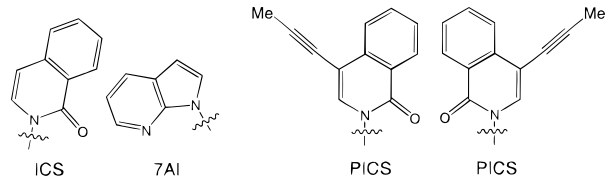


Figure 2.

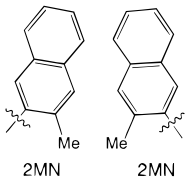


Figure 3.

3.3. Unnatural Base Pairs. On the basis of the above analysis, a number of combinations of these unnatural hydrophobic bases were found to possess thermodynamic and kinetic properties that make them interesting leads for the development of third base pair.

3.3.1. The ICS:7AI Unnatural Base Pair. Two candidate bases discussed above are **ICS** and the related **PICS**. These nucleobases have been found to be thermodynamically orthogonal to the native nucleobases. In duplex DNA they form stable base pairs with a variety of other hydrophobic bases, and some of these pairs have a stability and selectivity that rivals or even exceeds the stability of native DNA. During KF-mediated DNA synthesis, the nucleobases most efficiently paired with **ICS** and **PICS** were all hydrophobic. The incorporation of native triphosphates opposite **ICS** all proceeded with very low efficiencies, and as a triphosphate, d**ICS**TP insertion opposite a native base in the template was not competitive with the incorporation of a native triphosphate.

The unnatural nucleobase **7AI** was found to be an attractive partner found for **ICS** (Figure 2). As a pair in duplex DNA, **ICS**:**7AI** was only slightly less stable than a dA:dT pair. Misinsertion of native bases opposite **7AI** were all also unfavorable, all proceeding with a catalytic efficiency less than 6×10^3 M^{-1} min^{-1}. d**7AI**TP incorporation opposite native bases was not competitive with native DNA synthesis, and in template, **7AI** directed the incorporation of a native base with catalytic efficiencies less than 7×10^3 M^{-1} min^{-1}. The catalytic efficiencies for DNA synthesis were 4×10^5 and 3×10^5 M^{-1} min^{-1} for pairing d**ICS**TP opposite **7AI** and d**7AI**TP opposite **ICS**, respectively. Moreover, preliminary results indicate that, after synthesis of the unnatural base pair, continued DNA synthesis is reasonably efficient. The major fidelity problem for the **ICS**:**7AI** base pair arises from the synthesis of **ICS**:**ICS** and **7AI**:**7AI**. However, these base pairs are not extended by KF (data not shown).

3.3.2. The PICS:PICS Self-Pair. The use of a hydrophobic base that self-pairs eliminates half of the possible thermodynamic and kinetic mispairs, thus facilitating orthogonality to the native bases. As reported previously,[2] the **PICS**:**PICS** self-base pair (Figure 3) stabilizes the DNA duplex used in this study relative to a dA:dT or dG:dC base pair by 3.4 and 0.8 °C, respectively. The specificity of the stabilizing interactions is demonstrated by the decreased stability of the mispairs of **PICS** with native bases. The mispairs between **PICS** and the native

bases are all more than 7.1 °C destabilized, a thermodynamic selectivity equivalent to that found with the native bases. The kinetic selectivity of **PICS** for self-pairing is evident from the 20−1000-fold more efficient incorporation of d**PICS**TP opposite **PICS**, when compared to the rate for incorporation of the natural triphosphates opposite **PICS**. Moreover, d**PICS**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 **PICS**, is favored over all possible mispairs by at least a factor 20.

3.3.3. Base Pairs Containing 2MN. Despite the reasonable orthogonality demonstrated by **ICS** and **PICS**, the absolute efficiency of KF-mediated synthesis of DNA containing either unnatural base is several orders of magnitude reduced from the efficiency for the synthesis of DNA containing only native bases. However, several unnatural base pairs involving **2MN** are synthesized with efficiencies approaching, and in some cases, virtually identical to those characteristic of native bases. However, the utility of this base is compromised by the efficiency with which its triphosphate is incorporated opposite dA. As discused above, this may result from specific interactions between adenine and the minor groove methyl group. If the interactions between **2MN** and adenine could be selectively destabilized, **2MN** would be very attractive nucleobase. Experiments directed toward this end are in progress.

3.4. Efforts to Expand the Genetic Alphabet. Hydrophobicity is a strong and selective force in nucleic acid biochemistry. Hydrophobic bases are expected to avoid the issues of tautomerization that have been an obstacle in past efforts to identify new base pairs.[13,15,16] Here we have summarized our initial efforts to optimize interstrand hydrophobic packing to control information storage and replication. Unnatural nucleobases have been found which use hydrophobicity to drive correct pairing in duplex DNA as well as during enzymatic DNA synthesis. The stability and selectivity of the unnatural base pairs is comparable with native pairs. Several unnatural base pairs are synthesized by KF with a kinetic efficiency equivalent to that of native DNA synthesis. Moreover, the unnatural bases are orthogonal to the native bases, with at least an order of magnitude selectivity favoring correct pairing. We are currently examining derivatives of these nucleobases reported above, with the goal of optimizing both function and orthogonality. Structural studies are also in progress to identify any perturbations of the duplex which result from the unnatural bases. Polymerase-mediated extension of primers containing the unnatural base is also critical for the efficient synthesis of DNA containing three base pairs, and experiments addressing this issue are also in progress.

4. Experimental Section

General Methods. All reactions were carried out in oven-dried glassware under inert atmosphere unless otherwise stated. All solvents were dried over 4 Å molecular sieves except dichloromethane (distilled from CaH_2), tetrahydrofuran (distilled from sodium and potassium metal), and diethyl ether (distilled from $LiAlH_4$). Oligonucleotides were synthesized using an Applied Biosystems Inc. 392 DNA/RNA synthesizer. DNA synthesis reagents were purchased from Glen Research, Sterling, VA. All other reagents were purchased from Sigma-Aldrich. High-resolution mass spectroscopic data were obtained from the facilities at The Scripps Research Institute and the University of California-Riverside.

General Melting Temperature Procedure. Duplex oligonucleotide denaturation temperature measurements were made in buffer containing 10 mM PIPES, 10 mM $MgCl_2$, 100 mM NaCl, and 3 μM oligonucleotide using a Cary 200 Bio UV−visible spectrometer. Measurements were taken over a range of 16−80 °C at 0.5 °C/min intervals. Melting temperatures were obtained from the derivative method utilizing the Cary Win UV thermal application software.

General Polymerase Kinetic Assay Protocol. Primer was 5′-end ^{32}P-labeled with polynucleotide kinase. Primer-template duplexes were annealed by mixing in the reaction buffer, heating to 90 °C, and slow cooling to room temperature. Assay conditions include: 40 nM template−primer duplex, 0.11−1.34 nM enzyme, 50 mM Tris buffer (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, and 50 μg/mL BSA. The reactions were initiated by adding the DNA−enzyme mixture to an equal volume (5 μL) of a 2× dNTP 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). The reaction mixture, (5 μL) was then analyzed by 15% polyacrylamide gel electrophoresis. Radioactivity was quantified using a PhosphorImager (Molecular Dynamics), with overnight exposures and the ImageQuant program. The kinetic data were fit to the Michaelis−Menten equation using the program Kaleidograph (Synergy Software). The data presented are averages of duplicates or triplicates.

Compound 9. The data are presented for comparison to literature values.[40] ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.98 (4H, m), 7.20−7.30 (5H, m), 6.91 (1H, s), 5.62 (1H, m), 5.39 (1H, dd, J = 10.9, 5.0 Hz), 4.73 (1H, dd, J = 11.7, 3.8 Hz), 4.67 (1H, dd, J = 11.8, 3.6 Hz), 4.51 (1H, m), 2.52 (1H, dd, J = 13.8, 5.0 Hz), 2.43 (3H, s), 2.40 (3H, s), 2.28 (3H, s), 2.19 (3H, s), 2.15 (1H, m), 2.04 (3H, s). ^{13}C NMR (150 MHz, $CDCl_3$) δ 166.4, 166.2, 144.1, 143.8, 136.0, 135.6, 134.4, 131.7, 131.6, 129.7, 129.2, 129.1, 127.1, 127.0, 126.2, 82.5, 77.6, 64.7, 40.6, 21.7, 21.7, 19.2, 18.5. HRMS calcd for $C_{30}H_{33}O_5Na$ (MH^+): 473.2328; found: 473.2339.

Compound 1. The data are presented for comparison to literature values.[40] ^{1}H NMR (500 MHz, CD_3OD) δ 7.25 (1H, s), 6.86 (1H, s), 5.27 (1H, dd, J = 10.5, 5.5 Hz), 4.29 (1H, m), 3.90 (1H, m), 3.71 (1H, dd, J = 11.8, 5.5 Hz), 3.67 (1H, dd, J = 11.8, 5.5 Hz), 2.23 (3H, s), 2.19 (3H, s), 2.17 (3H, s), 2.16 (1H, m), 1.80 (1H, ddd, J = 13.0, 10.5, 6.0 Hz). ^{13}C NMR (150 MHz, CD_3OD) δ 138.1, 136.2, 135.0, 133.0, 132.4, 127.3, 88.7, 78.3, 74.4, 64.0, 43.4, 19.4, 19.3, 18.6.

Compound 10. To a solution of nucleoside **1** (60 mg, 0.254 mmol) in pyridine (1.5 mL) and CH_2Cl_2 (1.5 mL) was added triethylamine (300 μL, 2.152 mmol), followed by DMTr-Cl (115 mg, 0.339 mmol). After stirring at room temperature for 30 min, the reaction mixture was concentrated and purified by column chromatography on silica gel (50−80% ethyl acetate in hexane) to afford the tritylated nucleoside (127 mg), which was dissolved in CH_2Cl_2 (2.5 mL). To the purified nucleoside was added a catalytic amount of DMAP (∼2 mg), followed by triethylamine (250 μL, 1.794 mmol) and 2-cyanoethyl diisopropylaminochlorophosphoramidite (105 μL, 0.471 mmol). After 15 min, the reaction mixture was partitioned between CH_2Cl_2 (20 mL) and saturated aqueous $NaHCO_3$ (20 mL). The layers were separated, and the aqueous layer was extracted with 2 × 20 mL CH_2Cl_2. The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification by column chromatography on silica gel (10−30% ethyl acetate in 5% triethylamine/hexane) afforded phosphoramidite **10** (151 mg, 80.5% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.25−7.54 (10H, m), 6.94 (1H, s), 6.85 (4H, m), 5.33 (1H, m), 4.55 (1H, m), 4.22 (1H, m), 3.25−4.85 (10H, m), 2.30−2.65 (5H, m), 2.29 (3H, m), 2.23 (3H, s), 2.17 (3H, s), 2.06 (1H, m), 1.05−1.20 (12H, m). ^{31}P NMR (140 MHz, $CDCl_3$) δ 148.9, 148.2. HRMS calcd for $C_{44}H_{55}N_2O_6PCs$ (MCs^+): 871.2852; found: 871.2833.

Compound 11. Proton sponge (12 mg, 0.057 mmol) and nucleoside **1** (9 mg, 0.038 mmol) were dissolved in trimethyl phosphate (0.19 mL) and cooled to 0 °C. $POCl_3$ (4 μL, 0.045 mmol) was added, and the lavender slurry was stirred at 0 °C for 2 h. Tributylamine (56 μL, 0.235 mmol) was added, followed by a 0.4 M solution of tributylammonium pyrophosphate (33 mg) in DMF. After 1 min, the reaction was quenched by addition of 1 M aqueous triethylammonium bicarbonate (4 mL). The resulting crude solution was lyophilized; purification via reverse phase (C18) HPLC (4−30% CH_3CN in 0.1M TEA-bicarbonate, pH 7.5) afforded triphosphate **11** as a white solid.

(40) Schweitzer, B. A.; Kool, E. T. *J. Org. Chem.* **1995**, *60*, 8326.

Compound 12. To a suspension of magnesium metal (52 mg, 2.167 mmol) in THF (2 mL) was added 5-bromo-*m*-xylene (293 μL, 2.164 mmol). The resulting suspension was heated to ~50 °C. After 1 h, 400 μL of this solution was added to chloroglycoside **8** (126 mg, 0.324 mmol) in THF (1 mL). After 14 h, an additional 100 μL of the aforementioned Grignard solution was added. After 15 h, the reaction was partitioned between ethyl acetate (10 mL) and saturated aqueous NH_4Cl (10 mL). The layers were separated, and the aqueous layer was extracted with ethyl acetate (2 × 15 mL). The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification by column chromatography on silica gel (5−15% ethyl acetate in hexane) afforded nucleoside **12** (13 mg, 9%) and its α-anomer (48 mg, 32%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 (4H, m), 7.28 (2H, d, J = 8.1 Hz), 7.24 (2H, d, J = 8.1 Hz), 7.02 (2H, s), 6.92 (1H, s), 5.62 (1H, m), 5.20 (1H, dd, J = 10.9, 5.0 Hz), 4.71 (1H, dd, J = 11.8, 4.0 Hz), 4.66 (1H, dd, J = 11.8, 3.6 Hz), 4.53 (1H, m), 2.50 (1H, m), 2.44 (3H, s), 2.41 (3H, s), 2.26 (7H, m). HRMS calcd for $C_{29}H_{31}O_5$ (MH^+): 459.2171; found: 459.2179.

Compound 2. To a solution of nucleoside **12** in CH_3OH (10 mL) was added 1 M NaOMe (2 mL). After 45 min, the excess NaOMe was quenched with NH_4Cl (~100 mg). The resulting slurry was concentrated, and purified by column chromatography on silica gel (1−5% CH_3OH in CH_2Cl_2) to afford nucleoside **2** (105 mg, 82%). ^{1}H NMR (400 MHz, CD_3OD) δ 6.98 (2H, s), 6.78 (1H, s), 5.03 (1H, dd, J = 10.6, 5.3 Hz), 4.29 (1H, m), 3.92 (1H, ddd, J = 7.6, 5.2, 2.4 Hz), 3.66 (1H, m), 2.27 (6H, s), 2.14 (1H, ddd, J = 13.1, 5.4, 1.6 Hz), 1.91 (1H, ddd, J = 13.1, 10.6, 5.9 Hz). ^{13}C NMR (150 MHz, $CDCl_3$) δ 140.8, 138.1, 129.5, 129.3, 123.8, 123.4, 87.2, 80.1, 63.4, 43.9, 21.3. HRMS calcd for $C_{13}H_{19}O_3$ (MH^+): 223.1334; found: 223.1326.

Compound 13. To a solution of nucleoside **1** (90 mg, 0.405 mmol) in pyridine (2 mL) and CH_2Cl_2 (2 mL) was added triethylamine (300 μL, 2.152 mmol), followed by DMTr-Cl (175 mg, 0.516 mmol) in two portions over 30 min. After stirring at room temperature for 2 h, the reaction mixture was concentrated and purified by column chromatography on silica gel (50−80% ethyl acetate in hexane) to afford the tritylated nucleoside (185 mg), which was dissolved in CH_2Cl_2 (3.5 mL). A catalytic amount of DMAP (~2 mg) was added, followed by triethylamine (350 μL, 2.512 mmol) and 2-cyanoethyl diisopropylaminochloro phosphoramidite (160 μL, 0.718 mmol). After 30 min, the reaction mixture was partitioned between ethyl acetate (20 mL) and saturated aqueous $NaHCO_3$ (20 mL). The layers were separated, and the aqueous layer was extracted with 2 × 20 mL ethyl acetate. The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification by column chromatography on silica gel (10−30% ethyl acetate in 5% triethylamine/hexane) afforded phosphoramidite **13** (238 mg, 81% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.20−7.55 (9H, m), 7.09 (2H, s), 6.93 (1H, s), 6.83 (4H, m), 5.12 (1H, m), 4.25 (1H, m), 4.14 (1H, m), 3.60−3.90 (8H, m), 3.20−3.40 (2H, m), 2.40−2.80 (5H, m), 2.30 (6H, m), 2.07 (1H, m), 1.03−1.29 (12H, m). ^{31}P NMR (140 MHz, $CDCl_3$) δ 148.5, 148.3.

Compound 14. Proton sponge (23 mg, 0.107 mmol) and nucleoside **2** (16 mg, 0.072 mmol) were dissolved in trimethyl phosphate (0.36 mL) and cooled to 0 °C. $POCl_3$ (8 μL, 0.090 mmol) was added dropwise, and the lavender slurry was stirred at 0 °C for 2 h. Tributylamine (105 μL, 0.441 mmol) was added, followed by a solution of tributylammonium pyrophosphate (62 mg) in DMF (0.8 mL). After 1 min, the reaction was quenched via addition of 1 M triethylammonium bicarbonate (7 mL). The resulting crude solution was lyophilized, and purification via reverse phase (C18) HPLC (4−30% CH_3CN in 0.1 M TEA-bicarbonate, pH 7.5) afforded triphosphate **14** as a white solid.

Compound 16. To a stirred solution of 2-bromo-1,4-dimethylnaphthylene (640 mg, 2.72 mmol) in THF (15 mL) at −78 °C was added *n*BuLi (2 mL, 2 M in cyclohexane) slowly dropwise. The lime-green mixture was stirred for 15 min, at which time a solution of aldehyde **15** (1.214 g, 1.801 mmol) in THF (5 mL) was added dropwise down the side of the flask. After 1 h, the bath was removed, and the reaction mixture was warmed to room temperature. After a total of 2.5 h, the reaction mixture was partitioned between ethyl acetate (100 mL) and saturated aqueous $NaHCO_3$ (100 mL). The layers were separated, and the aqueous layer was extracted with 2 × 100 mL ethyl acetate. The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification via column chromatography on silica gel (10−20% ethyl acetate in hexane) afforded the desired *S*-diastereomer (712 mg, 48%; the *R*-diastereomer was isolated in 42% yield), which was dissolved in pyridine (9 mL) and cooled to 0 °C. Triethylamine (570 μL, 4.09 mmol) was added, followed by the slow dropwise addition of mesyl chloride (90 μL, 1.17 mmol). The reaction mixture was warmed slowly to room temperature over 1.5 h, at which time the reaction was quenched by addition of saturated aqueous $NaHCO_3$ (~2 mL). Concentration, followed by purification by column chromatography on silica gel (5−15% ethyl acetate in hexane), afforded the protected nucleoside **16** (176 mg, 17% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.17 (1H, m), 8.07 (1H, m), 7.83 (1H, s), 7.65 (2H, d, J = 7.9 Hz), 7.31−7.59 (11H, m), 6.92 (6H, m), 5.74 (1H, dd, J = 10.6, 5.2 Hz), 4.58 (2H, s), 4.39 (1H, m), 4.32 (1H, m), 3.86 (3H, s), 3.84 (6H, s), 3.58 (1H, dd, J = 9.9, 4.5 Hz), 3.47 (1H, dd, J = 9.9, 3.7 Hz), 2.72 (3H, s), 2.67 (3H, s), 2.53 (1H, dd, J = 13.1, 5.3 Hz), 2.06 (1H, m). ^{13}C NMR (100 MHz, $CDCl_3$) δ 159.1, 158.4, 145.0, 136.2, 136.1, 132.7, 132.3, 132.0, 130.1, 130.1, 129.2, 128.2, 127.9, 127.7, 126.7, 125.4, 125.0, 124.5, 124.4, 124.1, 113.8, 113.0, 86.1, 84.1, 81.1, 77.5, 70.7, 64.3, 55.1, 55.0, 40.7, 19.4, 13.7. HRMS calcd for $C_{46}H_{46}O_6Cs$ (MCs^+): 827.2349; found: 827.2378.

Compound 17. To a stirred solution of nucleoside **16** (190 mg, 0.274 mmol) in acetic acid (6 mL) and methanol (1 mL) was added trifluoroacetic acid (10 drops). After stirring at room temperature for 20 min, the orange reaction mixture was concentrated. Purification via column chromatography on silica gel (20−50% ethyl acetate in hexane) afforded nucleoside **17** (93 mg, 86%). ^{1}H NMR (400 MHz, CDCl3) δ 8.09 (1H, m), 8.01 (1H, m), 7.54 (2H, m), 7.32 (2H, d, J = 8.5 Hz), 6.92 (2H, d, J = 8.6 Hz), 5.61 (1H, dd, J = 10.6, 5.4 Hz), 4.53 (2H, m), 4.21 (1H, m), 4.16 (1H, m), 3.91 (1H, dd, J = 11.7, 3.7 Hz), 3.82 (6H, s), 3.80 (1H, m), 2.43 (1H, dd, J = 13.4, 5.4 Hz), 1.89 (1H, ddd, J = 13.4, 10.6, 6.7 Hz). 13C NMR (100 MHz, CDCl3) δ 159.3, 135.1, 132.7, 132.4, 132.0, 130.0, 129.3, 128.5, 125.6, 125.2, 124.6, 124.5, 123.4, 113.9, 113.8, 85.1, 80.4, 77.6, 63.6, 55.2, 40.4, 19.6, 13.8. HRMS calcd for $C_{25}H_{28}O_4Na$ (MNa^+): 415.1885; found: 415.1882.

Compound 3. To stirred a solution of **17** (8 mg, 0.020 mmol) in CH_2Cl_2 (1 mL) and H_2O (1 drop) was added DDQ (7 mg, 0.031 mmol). After 1 h, the reaction mixture was partitioned between ethyl acetate (15 mL) and saturated aqueous $NaHCO_3$ (15 mL). The layers were separated, and the aqueous layer was extracted with 2 × 15 mL ethyl acetate. The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification by column chromatography on silica gel (1−5% CH_3OH in CH_2Cl_2) afforded nucleoside **3** (3 mg, 55%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.07 (1H, m), 7.98 (1H, m), 7.50 (2H, m), 7.45 (1H, s), 5.24 (1H, dd, J = 11.5, 1.7 Hz), 4.23 (1H, m), 3.96 (2H, m), 3.81 (1H, m), 2.67 (3H, s), 2.62 (3H, s), 2.10 (1H, m), 1.88 (1H, ddd, J = 14.2, 11.6, 2.3 Hz). ^{13}C NMR (150 MHz, $CDCl_3$) δ 135.5, 132.8, 132.7, 132.1, 127.6, 125.7, 125.3, 124.7, 124.5, 123.8, 76.0, 70.9, 69.2, 68.0, 36.5, 19.5, 13.8. HRMS calcd For $C_{13}H_{24}NO_3$ (MNH_4^+): 290.1756; found: 290.1764.

Compound 18. To a solution of nucleoside **3** (28 mg, 0.103 mmol) in pyridine (1 mL) was added triethylamine (75 μL, 0.535 mmol), followed by DMTr-Cl (70 mg, 0.206 mmol). After stirring at room temperature for 30 min, the reaction mixture was concentrated; purification via column chromatography on silica gel (50−80% ethyl acetate in hexane) afforded the tritylated nucleoside (45 mg), which was dissolved in CH_2Cl_2 (1 mL). A catalytic amount of DMAP (~2 mg) was added, followed by triethylamine (65 μL, 0.470 mmol), and 2-cyanoethyl diisopropylaminochloro phosphoramidite (35 μL, 0.157 mmol). After 15 min, the reaction mixture was partitioned between CH_2Cl_2 (20 mL) and saturated aqueous $NaHCO_3$ (20 mL). The layers were separated, and the aqueous layer was extracted with 2 × 20 mL CH_2Cl_2. The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification by column chromatography on silica gel (10−30% ethyl acetate in 5% triethylamine/hexane) afforded phosphoramidite **18** (51 mg, 64% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.09 (1H, m), 7.98 (1H, m), 7.75 (1H, d, J = 1.8 Hz), 7.53 (4H, m), 7.43 (4H, m), 7.20−7.30 (3H, m), 6.83 (4H, m), 5.66 (1H, m), 4.56 (1H, m), 4.26 (1H, m), 3.83 (1H, m), 3.78 (3H, s), 3.77 (3H, s), 3.45−3.70 (4H, m), 3.32 (1H, m), 2.64 (3H, s), 2.62 (1H, m), 2.56 (3H, s), 2.46 (1H, dd, J = 6.6, 6.5 Hz), 2.42 (1H, m), 2.08 (1H, m), 1.19 (9H,

m), 1.08 (3H, d, $J = 6.8$ Hz). ^{31}P NMR (140 MHz, $CDCl_3$) δ 148.9, 148.3. HRMS calcd for $C_{47}H_{55}N_2O_6PCs$ (MCs^+): 907.2852; found: 907.2837.

Compound 19. Proton sponge (22 mg, 0.105 mmol) and nucleoside **3** (19 mg, 0.070 mmol) were dissolved in trimethyl phosphate (0.34 mL) and cooled to 0 °C. $POCl_3$ (7 μL, 0.077 mmol) was added dropwise, and the lavender slurry was stirred at 0 °C for 2 h. Tributylamine (90 μL, 0.378 mmol) was added, followed by a 0.5 M solution of tributylammonium pyrophosphate in DMF (0.35 mL). After 1 min, the reaction was quenched via addition of 1 M triethylammonium bicarbonate (7 mL). The resulting crude solution was lyophilized; purification via reverse phase (C18) HPLC (4−35% CH_3CN in 0.1 M TEA-bicarbonate, pH 7.5) afforded triphosphate **19** as a white solid.

Compound 4. To a solution of 2-bromo-3-methylnaphthalene (0.595 g, 2.70 mmol) in THF (13.3 mL) at −78 °C was added *n*BuLi (2.0 M solution in cyclohexane) dropwise over 10 min. After 15 min, a solution of aldehyde **20** was added dropwise down the side of the flask over 5 min. The reaction was stirred for 30 min and then brought to room temperature during which time the reaction color turned from deep green to brown. The reaction was partitioned between ethyl acetate (100 mL) and saturated aqueous $NaHCO_3$ (100 mL). The layers were separated, and the aqueous layer was extracted with 2 × 100 mL ethyl acetate. The combined organics were dried over Na_2SO_4, filtered, and concentrated. The material was subjected to a short plug of silica gel (20% ethyl acetate in hexane) to afford a crude mixture of both diasteromers. The crude product was dissolved in CH_2Cl_2 (32 mL) and the solution cooled to 0 °C. To the solution was added triethylamine (0.241 mL) followed by methanesulfonyl chloride (0.120 mL). After 30 min, the reaction was quenched with saturated aqueous $NaHCO_3$ (30 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated. To the crude product was added a 4:1 solution of trifluoroacetic acid:methanol (56 mL). After 20 min of stirring, the reaction was concentrated. Residual trifluoroacetic acid was neutralized with minimal saturated aqueous $NaHCO_3$ and extracted with dichloromethane (4 × 20 mL). The combined organics were dried over Na_2SO_4, filtered, and concentrated. Column chromatography (3% MeOH in CH_2Cl_2) afforded the desired β-anomer **4** (0.108 g, 0.4181 mmol) in 24% yield over 3 steps. ^{1}H NMR (600 MHz, MeOH) δ 8.03 (1H, s), 7.78 (1H, d, $J = 7.1$ Hz), 7.71 (1H, $J = 7.1$ Hz), 7.59 (1H, s), 7.37 (2H, m), 5.44 (1H, dd, $J = 10.1$, 5.3 Hz), 4.35 (1H, m), 3.98 (1H, m), 3.76 (2H, m), 2.47 (3H, s), 2.36 (1H, ddd, $J = 13.15$, 5.7, 1.8 Hz), 1.90 (1H, m). ^{13}C NMR (150 MHz, $CDCl_3$) δ 139.8, 133.8, 133.6, 133.1, 128.6, 128.2, 127.3, 126.1, 125.6, 124.1, 88.2, 78.2, 73.8, 63.5, 42.8, 19.1. HRMS calcd for $C_{16}H_{18}O_3$ (MNa^+): 281.1154; found: 281.1158.

Compound 21. Proton sponge (13 mg, 0.062 mmol) and nucleoside **3** (10 mg, 0.038 mmol) were dissolved in trimethyl phosphate (0.2 mL) and cooled to 0 °C. $POCl_3$ (4 μL, 0.044 mmol) was added dropwise, and the lavender slurry was stirred at 0 °C for 2 h. Tributylamine (60 μL, 0.252 mmol) was added, followed by a 0.5 M solution of tributylammonium pyrophosphate in DMF (0.16 mL). After 1 min, the reaction was quenched via addition of 1 M triethylammonium bicarbonate (7 mL). The resulting crude solution was lyophilized, and purification via reverse phase (C18) HPLC (4−35% CH_3CN in 0.1 M TEA-bicarbonate, pH 7.5) afforded triphosphate **21** as a white solid. ^{31}P NMR (140 MHz, 50mM Tris, 2mM EDTA, pH 7.5 in D2O) δ −5.77 (d, $J = 18.5$), −10.4. (d, $J = 17.1$ Hz), −21.94 (t, $J = 49.0$ Hz).

Compound 22. To **4** (0.093 g, 0.36 mmol), azeotroped from pyridine (2 × 0.1 mL), was added pyridine (1.5 mL). To the resulting solution was added a solution of DMTr-Cl (0.18 g, 0.54 mmol) in pyridine (0.67 mL) over 20 min. The reaction solution was stirred at room temperature for an additional 20 min, at which point ethyl acetate (20 mL) was added. The organics were extracted with saturated $NaHCO_3$ (5 mL) and brine (5 mL), dried over Na_2SO_4, and concentrated in vacuo. Flash chromatography on silica gel (50−80% ethyl acetate in hexane) yielded 130 mg (66%) of 5′-protected nucleoside as a white foam. To the protected nucleoside in CH_2Cl_2 (1.8 mL) was added triethylamine (0.11 mL, 0.80 mmol). The solution was cooled to 0 °C, and cyanoethyl diisopropylchlorophosphoramidite (0.068 mL, 0.26 mmol) was added. The reaction was allowed to reach rt over 15 min. The solution was transferred to ethyl acetate (15 mL) and extracted with saturated $NaHCO_3$ (10 mL) and brine (10 mL), dried over Na_2SO_4 and concentrated in vacuo. Flash chromatography (25% ethyl acetate in hexane, 2% NEt_3) yielded 120 mg (78%) of **22** as a white foam. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.16 (1H, d, $J = 5.6$ Hz), 7.75 (1H, d, $J = 7.9$ Hz), 7.66 (1H, d, $J = 7.9$ Hz), 7.62 (1H, s), 7.56 (2H, m), 7.45−7.20 (9H, m), 6.83 (4H, m), 5.49 (1H, m), 4.56 (1H, m), 4.28 (1H, m), 3.84 (1H, m), 3.78 (6H, s), 3.7−3.3 (5H, m), 2.70−2.30 (3H, m), 2.03 (1H, m), 1.18 (9H, m), 1.09 (3H, d, $J = 9.2$ Hz). ESMS calcd for $C_{46}H_{53}N_2O_6P$ ($MH^=$): 761.3; found: 761.

Compound 23. To isocarbostyril (0.50 g, 3.4 mmol) was added acetonitrile (11 mL) and bis(trimethylsilyl)acetamide (0.85 mL, 3.4 mmol). After 30 min, an additional 12 mL of acetonitrile was added to the reaction solution followed by addition of **8** (1.1 g, 2.8 mmol). The reaction was brought to 0 °C and $SnCl_4$ (0.059 mL, 0.69 mmol) was added dropwise. After 30 min, complete dissolution of the ribofuranoside had occurred. Ethyl acetate (250 mL) and the resulting solution was successively extracted with saturated $NaHCO_3$ (3 × 100 mL) and brine (1 × 100 mL). The organics were dried over Na_2SO_4 and solvents were removed in vacuo. Flash chromatography (EtOAc:hexanes; 1:4) yielded 0.39 g (23%) of β-isomer **23** (faster migrating anomer) as a white foam. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.40 (1H, m), 7.98 (2H, m), 7.92 (2H, m), 7.63 (1H, ddd, $J = 8.9$, 7.3, 1.3 Hz), 7.47 (2H, m), 7.44 (1H, d, $J = 7.6$ Hz), 7.20−7.28 (4H, m), 6.86 (1H, dd, $J = 8.6$, 5.5 Hz), 6.43 (1H, d, $J = 7.7$ Hz), 5.63 (1H, m), 4.72 (2H, m), 4.59 (1H, m), 2.87 (1H, m), 2.43 (3H, s), 2.39 (3H, s), 2.32 (1H, m). HRMS (MALDI) calcd for $C_{30}H_{27}NO_6Na$ (MNa^+): 520.1736; found: 520.1748.

Compound 5. To a solution of **23** (205 mg, 0.412 mmol) in methanol (4 mL) was added 1 M NaOMe (1 mL). After 30 min, the reaction was quenched by addition of NH_4Cl (∼100 mg) and concentrated. Purification via column chromatography on silica gel (1−10% CH_3OH:CH_2Cl_2) afforded 96 mg (89%) of **5** as a white foam. ^{1}H NMR (400 MHz, CD_3OD) δ 8.28 (1H, dd, $J = 8.1$, 0.6 Hz), 7.75 (1H, d, $J = 7.6$ Hz), 7.68 (1H, m), 7.59 (1H, d, $J = 7.7$ Hz), 7.50 (1H, m), 6.69 (1H, ddd, $J = 6.7$, 4.2, 2.7 Hz), 4.42 (1H, ddd, $J = 6.5$, 3.3, 3.2 Hz), 3.99 (1H, dd, $J = 7.3$, 3.7 Hz), 3.83 (1H, dd, $J = 12.0$, 3.5 Hz), 3.77 (1H, dd $J = 12.0$, 4.1 Hz), 2.42 (1H, ddd, $J = 13.5$, 6.1, 3.4 Hz), 2.18 (1H, ddd, $J = 13.6$, 7.0, 6.7 Hz). ^{13}C NMR (100 MHz, CD_3OD) δ 163.5, 138.6, 134.0, 128.3, 128.0, 127.9, 127.3, 126.5, 108.0, 88.9, 88.5, 72.4, 63.0, 42.0. HRMS (MALDI) calcd for $C_{14}H_{15}NO_4Na$ (MNa^+): 284.0899; found: 284.0897.

Compound 24. To a solution of nucleoside **5** (85 mg, 0.325 mmol) in pyridine (3.5 mL) was added triethylamine (230 μL, 1.649 mmol), followed by DMTr-Cl (229 mg, 0.678 mmol). After stirring for 20 min at room temperature, the reaction mixture was concentrated. Purification via column chromatography on silica gel (50−80% ethyl acetate in hexane) afforded 151 mg of tritylated product, which was dissolved in CH_2Cl_2 (4 mL). A catalytic amount of DMAP (∼2 mg) was added, followed by triethylamine (224 μL, 1.607 mmol) and 2-cyanoethyl diisopropylaminochloro phosphoramidite (120 μL, 0.536 mmol). After 15 min, the reaction mixture was partitioned between CH_2Cl_2 (20 mL) and saturated aqueous $NaHCO_3$ (20 mL). The layers were separated, and the aqueous layer was extracted with 2 × 20 mL CH_2Cl_2. The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification via column chromatography on silica gel (15−50% ethyl acetate in 5% triethylamine/hexane) afforded phosphoramidite **24** (151 mg, 61% over two steps). ^{31}P NMR (140 MHz, $CDCl_3$) δ 149.8, 149.0. HRMS calcd for $C_{44}H_{50}N_2O_7PCs$ (MCs+): 896.2441; found: 896.2412.

Compound 25. Proton sponge (28 mg, 0.132 mmol) and nucleoside **5** (23 mg, 0.088 mmol) were dissolved in trimethyl phosphate (0.9 mL) and cooled to 0 °C. $POCl_3$ (9 μL, 0.101 mmol) was added dropwise, and the lavender slurry was stirred at 0 °C for 2 h. Tributylamine (130 μL, 0.546 mmol) was added, followed by a solution of tributylammonium pyrophosphate (71 mg) in DMF (1 mL). After 1 min, the reaction was quenched via addition of 1 M triethylammonium bicarbonate (15 mL). The resulting crude solution was lyophilized, and purification via reverse phase (C18) HPLC (4−35% CH_3CN in 0.1 M TEA-bicarbonate, pH 7.5) afforded triphosphate **25** as a white solid. ^{31}P NMR (140 MHz, 50mM Tris, 2mM EDTA, pH 7.5 in D_2O) δ −7.62 (d, $J = 16.8$), −10.61 (d, $J = 18.2$ Hz), −22.04 (t, $J = 16.8$ Hz).

Compound 26. To a solution of nucleoside **7** (101 mg, 0.431 mmol) in pyridine (4 mL) was added triethylamine (400 μL, 2.875 mmol), followed by DMTr-Cl (310 mg, 0.915 mmol). After stirring at room temperature for 15 min, the reaction mixture was concentrated; purification by column chromatography on silica gel (30−50% ethyl acetate in hexane) afforded the tritylated nucleoside (162 mg), which was dissolved in CH_2Cl_2 (3.5 mL). A catalytic amount of DMAP (~2 mg) was added, followed by triethylamine (250 μL, 1.797 mmol) and 2-cyanoethyl diisopropylaminochloro phosphoramidite (130 μL, 0.582 mmol). After 15 min, the reaction mixture was partitioned between CH_2Cl_2 (20 mL) and saturated aqueous $NaHCO_3$ (20 mL). The layers were separated, and the aqueous layer was extracted with 1 × 30 mL CH_2Cl_2. The combined organics were dried over Na_2SO_4, filtered, and concentrated. Purification via column chromatography on silica gel (10−30% ethyl acetate in 5% triethylamine/hexane) afforded phosphoramidite **26** (194 mg, 61% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (1H, m), 7.87 (1H, d, J = 7.8 Hz), 7.15−7.50 (9H, m), 7.07 (1H, dd, J = 7.8, 4.7 Hz), 6.91 (1H, m), 6.78 (4H, m), 6.44 (1H, dd, J = 3.6, 1.8 Hz), 4.76 (1H, m), 4.24 (1H, m), 3.55−3.85 (12H, m), 3.38 (1H, m), 3.28 (1H, m), 2.72 (1H, m), 2.60 (1H, t, J = 6.4 Hz), 2.58 (1H, m), 2.44 (1H, t, J = 6.4 Hz), 1.09−1.20 (12H, m). ^{31}P NMR (140 MHz, $CDCl_3$) δ 149.1, 148.9.

Compound 27. Proton sponge (15 mg, 0.071 mmol) and nucleoside **7** (11 mg, 0.045 mmol) were dissolved in trimethyl phosphate (0.22 mL) and cooled to 0 °C. $POCl_3$ (5 μL, 0.054 mmol) was added dropwise, and the lavender slurry was stirred at 0 °C for 2 h. Tributylamine (55 μL, 0.225 mmol) was added, followed by a solution of tributylammonium pyrophosphate (40 mg) in DMF (0.5 mL). After 1 min, the reaction was quenched via addition of 1 M triethylammonium bicarbonate (5 mL). The resulting crude solution was lyophilized, and purification via reverse phase (C18) HPLC (4−35% CH_3CN in 0.1 M TEA-bicarbonate, pH 7.5) afforded triphosphate **27** as a white solid.

Acknowledgment. Funding was provided by the National Institutes of Health (GM 60005 to F.E.R.) and the Skaggs Institute for Chemical Biology (F.E.R. and P.G.S.) 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.

JA9940064