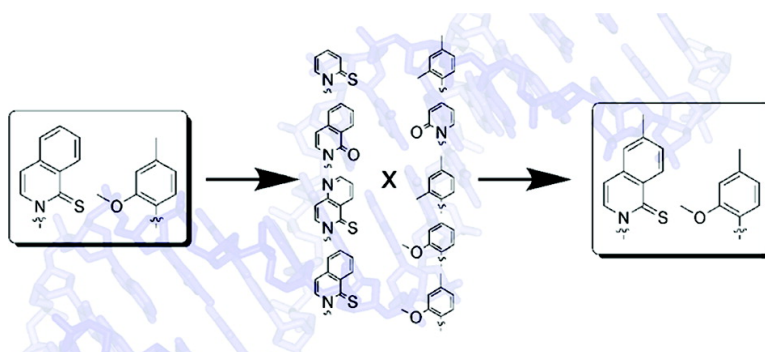


Discovery, Characterization, and Optimization of an Unnatural Base Pair for Expansion of the Genetic Alphabet

Aaron M. Leconte, Gil Tae Hwang, Shigeo Matsuda, Petr Capek, Yoshiyuki Hari, and Floyd E. Romesberg

J. Am. Chem. Soc., **2008**, 130 (7), 2336-2343 • DOI: 10.1021/ja078223d

Downloaded from <http://pubs.acs.org> on February 8, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 4 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)

Discovery, Characterization, and Optimization of an Unnatural Base Pair for Expansion of the Genetic Alphabet

Aaron M. Leconte, Gil Tae Hwang, Shigeo Matsuda, Petr Capek,
Yoshiyuki Hari, and Floyd E. Romesberg*

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

Received October 26, 2007; E-mail: floyd@scripps.edu

Abstract: DNA is inherently limited by its four natural nucleotides. Efforts to expand the genetic alphabet, by addition of an unnatural base pair, promise to expand the biotechnological applications available for DNA as well as to be an essential first step toward expansion of the genetic code. We have conducted two independent screens of hydrophobic unnatural nucleotides to identify novel candidate base pairs that are well recognized by a natural DNA polymerase. From a pool of 3600 candidate base pairs, both screens identified the same base pair, **dSICS:dMMO2**, which we report here. Using a series of related analogues, we performed a detailed structure–activity relationship analysis, which allowed us to identify the essential functional groups on each nucleobase. From the results of these studies, we designed an optimized base pair, **d5SICS:dMMO2**, which is efficiently and selectively synthesized by Kf within the context of natural DNA.

1. Introduction

DNA is an essential biomolecule which is responsible for encoding the complex information necessary for life. However, it is limited to a set of nucleobases that encode a finite number of three base codons; the expansion of the genetic alphabet to include additional, coding nucleobases would significantly increase the information potential of nucleic acids *in vitro*^{1,2} and, ultimately, *in vivo*. While significant progress has been made toward both developing and applying unnatural base pairs that form through unique hydrogen-bonding (H-bonding) patterns,^{3–7} pairing based on hydrophobic interactions has also emerged as a promising strategy for expansion of the genetic alphabet.^{8–12} Hydrophobic forces are capable of stabilizing the unnatural base pairs, as well as disfavoring mispairing with the natural nucleobases due to forced desolvation of the natural H-bonding functional groups. Indeed, several nucleotides bearing predominantly hydrophobic nucleobase analogues have been shown to pair stably and selectively in duplex DNA.^{13,14}

We,^{9,10,15,16} and others,^{8,12,17–19} have shown that hydrophobic forces are also sufficient for the enzymatic synthesis of an unnatural base pair by incorporation of an unnatural nucleoside triphosphate against a template unnatural nucleotide; however, synthesis beyond the unnatural base pair, i.e., extension, tends to be relatively inefficient and generally limits the utility of these base pairs. Recent efforts to modify either the nucleobases^{20–24} or the DNA polymerase²⁵ have significantly improved the rate of extension of unnatural base pairs and have demonstrated that efficient extension by the exonuclease deficient Klenow fragment of *E. coli* DNA Pol I (Kf) likely requires a minimally distorted primer terminus with a suitably positioned minor groove H-bond acceptor in the primer nucleobase. Unfortunately, these strategies have yet to yield a viable base pair candidate as the modifications that facilitate extension have also been

- (1) Joyce, G. F. *Ann. Rev. Biochem.* **2004**, *73*, 791–836.
- (2) Cropp, T. A.; Chin, J. W. *Curr. Opin. Chem. Biol.* **2006**, *10*, 601–6.
- (3) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322–8323.
- (4) Geyer, C. R.; Battersby, T. R.; Benner, S. A. *Structure (Camb.)* **2003**, *11*, 1485–1498.
- (5) Prudent, J. R. *Expert. Rev. Mol. Diagn.* **2006**, *6*, 245–52.
- (6) Yang, Z.; Sismour, A. M.; Sheng, P.; Puskar, N. L.; Benner, S. A. *Nucleic Acids Res.* **2007**, *35*, 4238–49.
- (7) Sismour, A. M.; Benner, S. A. *Nucleic Acids Res.* **2005**, *33*, 5640–6.
- (8) Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. *Nat. Methods* **2006**, *3*, 729–35.
- (9) Ogawa, A. K.; Wu, Y.; McMin, D. L.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 3274–3287.
- (10) Henry, A. A.; Olsen, A. G.; Matsuda, S.; Yu, C.; Geierstanger, B. H.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 6923–6931.
- (11) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727–733.
- (12) Hirao, I. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622–7.

- (13) Berger, M.; Wu, Y.; Ogawa, A. K.; McMin, D. L.; Schultz, P. G.; Romesberg, F. E. *Nucleic Acids Res.* **2000**, *28*, 2911–2914.
- (14) Matsuda, S.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 14419–14427.
- (15) McMin, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585–11586.
- (16) Ogawa, A. K.; Wu, Y.; Berger, M.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 8803–8804.
- (17) Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191–6192.
- (18) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704–708.
- (19) Chiaromonte, M.; Moore, C. L.; Kincaid, K.; Kuchta, R. D. *Biochemistry* **2003**, *42*, 10472–10481.
- (20) Leconte, A. M.; Matsuda, S.; Hwang, G. T.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2006**, *45*, (26), 4326–4329.
- (21) Hwang, G. T.; Leconte, A. M.; Romesberg, F. E. *ChemBioChem* **2007**, *8*, 1606–11.
- (22) Kim, Y.; Leconte, A. M.; Hari, Y.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7809–12.
- (23) Matsuda, S.; Leconte, A. M.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 5551–7.
- (24) Adelfinskaya, O.; Nashine, V. C.; Bergstrom, D. E.; Davisson, V. J. *J. Am. Chem. Soc.* **2005**, *127*, 16000–16001.
- (25) Leconte, A. M.; Chen, L.; Romesberg, F. E. *J. Am. Chem. Soc.* **2005**, *127*, 12470–1.

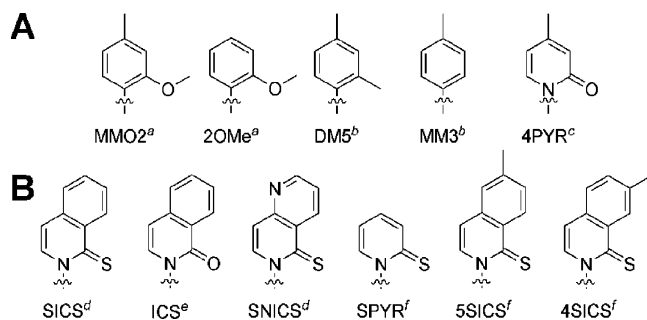


Figure 1. Unnatural nucleobases used in this study. (A) MMO2 and analogues: ^aref 23, ^bref 36, ^cref 20. (B) SICS and analogues: ^dref 28, ^eref 9, ^fsynthesized for this study.

found to limit synthesis and increase mispairing.^{20–22} Thus, rational approaches to base pair discovery have been limited by the conflicting demands of efficient synthesis and extension. In contrast to approaches based on rational design, screening-based approaches are not limited by our incomplete understanding of DNA stability and replication, and they can potentially identify both the determinants of efficient replication and unanticipated unnatural base pairs.^{26,27}

We now report the results of two independent screens of 3600 different unnatural base pairs, which both identified the same promising base pair formed between the dSICS²⁸ and dMMO2²³ nucleotide analogues (dSICS:dMMO2, Figure 1). Detailed steady-state kinetic studies with Kf confirmed that dSICS:dMMO2 is efficiently synthesized and extended; however, the utility of the base pair is limited by the facile formation of the dSICS:dSICS self-pair. To design improved analogues, as well as to better understand the determinants of efficient polymerase-mediated replication, we analyzed the DNA polymerase recognition of a family of related nucleobase derivatives. This, and previous data collected with other analogues, allowed us to determine the minimal requirements for polymerase recognition and to optimize the dSICS:dMMO2 base pair, culminating in the d5SICS:dMMO2 base pair (Figure 1). Comparison of the reported kinetic data with previously reported data for other unnatural base pair candidates suggests that d5SICS:dMMO2 is likely the most promising candidate for expansion of the genetic alphabet identified to date.

2. Results

2.1. Screening for Unnatural Base Pairs. We have previously reported the synthesis of a wide variety of unnatural nucleotide analogues bearing predominantly hydrophobic nucleobase analogues.^{8,9,11,14,15,20,22,28–34} From these, 60 were collected (Figure 2) and their phosphoramidites were incorporated into the 3' end of a 24-mer primer oligonucleotide and at

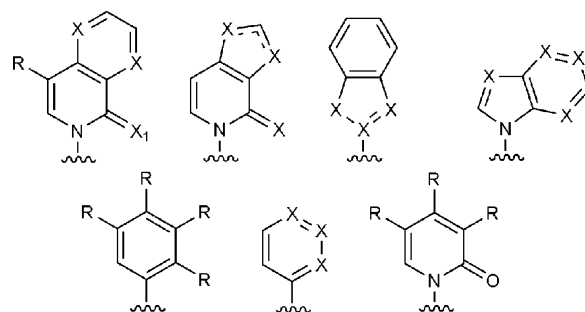


Figure 2. Representative nucleobase scaffolds and substitutions for unnatural nucleotides screened for functional heterobase pairs. X = heteroatom substitution; R = functional group. See Supporting Information for structures of the individual nucleotides.

the 24th position of a complementary 45-mer template oligonucleotide. Hybridization of any given primer strand with any template strand results in an unnatural primer terminus (dX:dY, where dX is the primer nucleobase and dY is the template nucleobase). The nucleotide in the template immediately 5' to the unnatural nucleotide is in all cases dG; thus, polymerase mediated dCTP incorporation results in extension of the unnatural terminus. 60 primers were divided into 10 groups of related unnatural nucleotides, and the primers were mixed in equal portions. Each group of primers was radiolabeled together and annealed to individual template oligonucleotide strands, creating a group of radiolabeled primers annealed to a single common template. Each “primer group:template” combination was challenged with dCTP and Kf, and the reaction products were quantified using gel electrophoresis. “Primer group:template” combinations that yielded the most 25-mer product were selected for further analysis of the individual component pairs. 274 single primer:template pairs were challenged with Kf and three different concentrations of dCTP. Reaction products were quantified using gel electrophoresis, and the Michaelis–Menten equation was fit to the data. From this rough estimate of the k_{cat}/K_M of extension of 274 individual pairs, the 40 most efficiently extended primer:template combinations were selected for full steady-state kinetic analysis using standard methods.³⁵ The steady-state rate constants of extension of these 40 pairs were examined in both the selected strand context (dX:dY) as well as the opposite strand context (dY:dX), resulting in six pairs that were efficiently extended in both contexts. We next measured the enzymatic synthesis of these pairs by Kf-mediated insertion of dXTP opposite dY and dYTP opposite dX. Of these six base pairs, the pair formed between dSICS and dMMO2 was the only identified pair that was efficiently incorporated in addition to being efficiently extended in both sequence contexts.

To increase our confidence that we had identified the most promising unnatural base pair, we ran a second screen. In this case, each of 3600 possible unnatural base pairs was screened simultaneously for efficient and high fidelity synthesis and extension using a fluorescence-based assay. Each 45-mer template with an unnatural nucleotide at position 24 was dissolved in polymerase reaction buffer and annealed to an 18-mer primer whose 3' nucleotide annealed to position 23 of the template, and the primer:template pairs were aliquoted into individual wells of a 384-well plate. To each well was added either all four natural dNTPs and one of the 60 unnatural dNTPs

(26) Kincaid, K.; Kuchta, R. D. *Nucleic Acids Res.* **2006**, *34*, e109.
 (27) Leconte, A. M.; Matsuda, S.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 6780–1.
 (28) Yu, C.; Henry, A. A.; Romesberg, F. E.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2002**, *41*, 3841–3844.
 (29) Wu, Y.; Ogawa, A. K.; Berger, M.; McMinn, D. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 7621–7632.
 (30) Tae, E. L.; Wu, Y. Q.; Xia, G.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 7439–7440.
 (31) Henry, A. A.; Yu, C.; Romesberg, F. E. *J. Am. Chem. Soc.* **2003**, *125*, 9638–9646.
 (32) Matsuda, S.; Henry, A. A.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2003**, *125*, 6134–6139.
 (33) Hwang, G. T.; Romesberg, F. E. *Nucleic Acids Res.* **2006**, *34*, 2037–45.
 (34) Matsuda, S.; Fillo, J. D.; Henry, A. A.; Rai, P.; Wilkens, S. J.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Schultz, P. G.; Spraggon, G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 10466–73.

(35) Creighton, S.; Bloom, L. B.; Goodman, M. F. *Methods Enzymol.* **1995**, *262*, 232–256.

Table 1. Steady-State Rate Constants of Kf-Mediated, dY-Templated, dXTP Incorporation for Synthesis of dSICS:dMMO2 and Mispairs^a

5' – dTAATACGACTCACTATAGGGAGA 3' – dATTATGCTGAGTGATATCCCTCT(Y)GCTAGGTTACGGCAGGATCGC				
dXTP	Y	k_{cat} (min ⁻¹)	K_M (μM)	k_{cat}/K_M (M ⁻¹ min ⁻¹)
SICS	MMO2	11.2 ± 0.2	0.8 ± 0.1	1.4 × 10 ⁷
MMO2	MMO2 ^b	5.1 ± 0.7	44 ± 4	1.2 × 10 ⁵
A	MMO2 ^b	3.3 ± 0.3	32 ± 4	1.0 × 10 ⁵
C, G, T	MMO2 ^b	nd ^d	nd ^d	< 1.0 × 10 ³
MMO2	SICS	6.2 ± 0.5	18 ± 3	3.4 × 10 ⁵
SICS	SICS ^c	5 ± 1	3.2 ± 3.1	1.6 × 10 ⁶
A	SICS ^c	1.6 ± 0.6	45 ± 2	3.7 × 10 ⁴
C	SICS ^c	nd ^d	nd ^d	1.3 × 10 ²
G	SICS ^c	13 ± 3	85 ± 9	1.5 × 10 ⁵
T	SICS ^c	2.7 ± 1.2	194 ± 66	1.4 × 10 ⁴

^a See Experimental Section for experimental details. ^b Reference 23. ^c Reference 28. ^d Rates too slow to independently determine k_{cat} and K_M .

(60 reactions) or only the four natural dNTPs (misincorporation control). Reactions were incubated with Kf and quenched with EDTA and SYBR Green I. SYBR Green I preferentially fluoresces in the presence of double stranded DNA, and the ratio of “reaction” fluorescence to “misincorporation” fluorescence was used as a measure of efficient and selective full length synthesis, simultaneously evaluating both incorporation and extension. The average ratio of fluorescence was 0.87, and 94 reactions were found in excess of 1.25. However, only one unnatural pair, dSICS:dMMO2, showed elevated values (1.9 and 1.5) in both possible sequence contexts. Thus, the same unnatural base pair was identified by both screens, allowing us to conclude with a high level of confidence that it is the most promising of the 3600 pairs screened. Thus, we chose the dSICS:dMMO2 base pair for further analysis.

2.2. Characterization of dSICS:dMMO2. To better understand the dSICS:dMMO2 pair, we next performed a detailed steady-state kinetic analysis. First, we characterized base pair synthesis by determining the rates at which each unnatural nucleoside triphosphate is inserted opposite its cognate base in the template (Table 1). For reference, in the analogous sequence context, dATP is inserted opposite dT with a second-order rate constant (k_{cat}/K_M , also referred to as efficiency) of approximately $3 \times 10^8 \text{ M}^{-1} \text{ min}^{-1}$. dSICSTP is inserted against template dMMO2 with a second-order rate constant of $1.4 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$, only approximately 20-fold less efficient than the synthesis of a natural pair. In the opposite context, dMMO2TP is incorporated against template dSICS with a second-order rate constant of $3.4 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$. Comparing these rates with those previously reported for dSICS and dMMO2 templated mispair synthesis,^{23,28} the correct unnatural base pair is synthesized with a minimum fidelity (e.g., the ratio of the second-order rate constants of correct synthesis to that for the most efficiently synthesized mispair) of 117-fold (insertion of dSICSTP opposite dMMO2) and 2-fold (insertion of dMMO2TP opposite dSICS). The relatively low, 2-fold selectivity of the dSICS:dMMO2 base pair results from the efficient insertion of dGTP opposite dMMO2. Although the dG:dMMO2 mispair is synthesized relatively efficiently, unlike the correct unnatural pair it is extended very inefficiently at a rate that is barely detectable (see below), and thus it is likely excised when the polymerase possesses exonuclease activity. In addition to fidelity against the natural nucleotides, the pair must be selective against

Table 2. Steady-State Rate Constants of Kf-Mediated Unnatural Terminus Extension by Insertion of dCTP for dSICS:dMMO2 and Mispairs^a

5' – dTAATACGACTCACTATAGGGAGA(X) 3' – dATTATGCTGAGTGATATCCCTCT(Y)GCTAGGTTACGGCAGGATCGC				
X	Y	k_{cat} (min ⁻¹)	K_M (μM)	k_{cat}/K_M (M ⁻¹ min ⁻¹)
SICS	MMO2	1.0 ± 0.1	0.9 ± 0.2	1.1 × 10 ⁶
A	MMO2 ^b	8.65 ± 0.89	187 ± 17	4.6 × 10 ⁴
MMO2	MMO2 ^b	0.87 ± 0.28	165 ± 47	5.3 × 10 ³
MMO2	SICS	3.8 ± 0.5	2.2 ± 0.3	1.7 × 10 ⁶
SICS	SICS ^c	1.5 ± 0.3	242 ± 78	6.3 × 10 ³
A	SICS	2.6 ± 1.0	85 ± 17	3.1 × 10 ⁴
C	SICS	0.7 ± 0.3	255 ± 84	2.6 × 10 ³
G	SICS	0.7 ± 0.3	191 ± 84	3.7 × 10 ³
T	SICS	12.0 ± 0.7	124 ± 26	9.6 × 10 ⁴

^a See Experimental Section for experimental details. ^b Reference 23. ^c Reference 28.

misincorporation of the incorrect unnatural nucleoside triphosphate. While dMMO2TP is inserted opposite dMMO2 in the template over 100-fold less efficiently than dSICSTP ($1.2 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$),²³ dSICSTP is inserted opposite dSICS at a rate of $1.6 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$ ²⁸ approximately 5-fold more efficiently than correct dMMO2TP incorporation against dSICS. Thus, dSICS:dMMO2 base pair synthesis is limited by the facile synthesis of the dSICS self-pair.

We next examined the extension efficiency of the dSICS:dMMO2 unnatural base pair by characterizing the steady-state rate at which Kf extends a primer terminating with dSICS paired opposite dMMO2 by incorporation of dCTP opposite dG (Table 2). In this context, the pair is extended with a second-order rate constant of $1.7 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$, which is approximately equal to the rate of extension of the fastest unnatural termini reported to date²¹ and only approximately 100-fold less efficient than a natural base pair in the same context. We then examined the rate at which dMMO2 is extended when paired opposite dSICS. The unnatural pair is again extended efficiently, with a k_{cat}/K_M of $1.1 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$. Thus, the dSICS:dMMO2 is efficiently extended in both sequence contexts.

To examine the selectivity of unnatural base pair extension, we measured the rate at which the most problematic mispairs (i.e., those synthesized with detectable rates) are extended by insertion of dCTP opposite dG in the template (Table 2). With dMMO2 in the template, this includes only the mispair with dA (the other mispairs are all synthesized with rates below the detectable limit of $10^3 \text{ M}^{-1} \text{ min}^{-1}$). With dSICS in the template, this includes mispairs with each natural nucleotide (synthesized at rates between 1.3×10^2 and $1.5 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$). The dA:dMMO2 mispair is extended with a second-order rate constant of $4.6 \times 10^4 \text{ M}^{-1} \text{ min}^{-1}$, which is 24-fold less efficient than the rate at which dSICS:dMMO2 is extended. The most efficiently extended mispair with dSICS was that formed with dT, which was extended with a second-order rate constant of $9.6 \times 10^4 \text{ M}^{-1} \text{ min}^{-1}$. All other mispairs formed with dSICS are extended with a rate less than or equal to $3.1 \times 10^4 \text{ M}^{-1} \text{ min}^{-1}$. Importantly, the mispair formed between dSICS and dG, which is the most efficiently synthesized, is extended approximately 500-fold less efficiently than the correct pair. Additionally, the extension of incorrect unnatural base pairs (dSICS:dSICS or dMMO2:dMMO2) is inefficient, occurring at rates less than $10^4 \text{ M}^{-1} \text{ min}^{-1}$. Thus, the dSICS:dMMO2

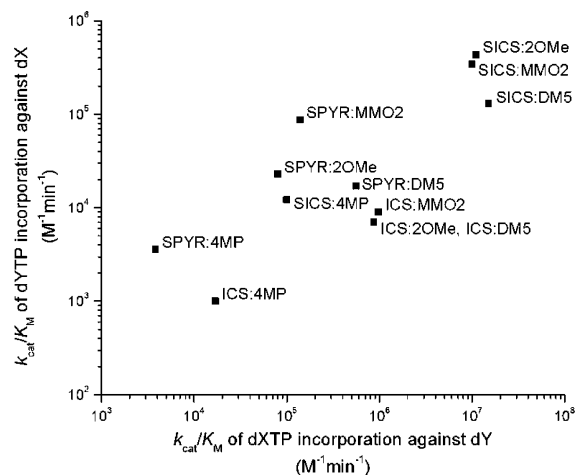


Figure 3. Second-order rate constants of Kf mediated synthesis of unnatural base pairs. The x-axis corresponds to incorporation of dXTP against dY while the y-axis corresponds to incorporation of dYTP against dX. Pairs are listed on the plot as X:Y. d4MPTP is incorporated against dICS at a rate below the limit of detection ($<1.0 \times 10^3 \text{ M}^{-1} \text{ min}^{-1}$). See Supporting Information (Table S1) for rate constants and sequence context.

base pair is extended with a minimum fidelity of 18 for dMMO2:dSICS and 24 for dSICS:dMMO2.

2.3. Structure–Activity Relationship Analysis. To understand the determinants of unnatural base pair synthesis and extension, as well as to further optimize the replication of the unnatural base pair, we measured the steady-state rate constants of both incorporation and extension of all possible pairs formed between several derivatives of dSICS and dMMO2 (Figure 1). The derivatives examined were designed to probe the role of nucleobase shape, electrostatics, minor groove H-bonding, and extended aromatic surface area.

We first examined the incorporation of different nucleoside triphosphate derivatives (dMMO2TP, d2OMeTP,²³ dDM5TP,³⁶ and d4MPTP²⁰) opposite dSICS (Figure 3, Table S1). In addition, to probe the role of the minor groove H-bond acceptor and aromatic surface area of the template nucleobase, we also examined the rates with which the same triphosphates are inserted opposite dICS⁹ or dSPYR. Interestingly, dMMO2TP, d2OMeTP, and dDM5TP are all incorporated against dSICS with similar rates, ranging from 1.3×10^5 to $4.3 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$; however, d4MPTP is incorporated 10- to 30-fold less efficiently than the other analogues. The data suggest that a more hydrophobic minor groove substituent in the dNTP facilitates unnatural base pair synthesis. The same general trend is observed when dICS is in the template since dMMO2TP, d2OMeTP, and dDM5TP are all incorporated more efficiently than d4MPTP. However, relative to dSICS, the rates are uniformly decreased by 12–60-fold with dICS in the template, demonstrating that the thiopyridone is more efficient at templating unnatural triphosphate insertion than the more natural-like pyridone. In addition, the ability to template unnatural base pair synthesis is moderately dependent on the bicyclic aromatic scaffold, as dSPYR templates the incorporation of all four MMO2 analogues less efficiently than dSICS, but still more efficiently than dICS.

We next examined the incorporation of dSICSTP and its analogues against dMMO2 and its analogues in the template

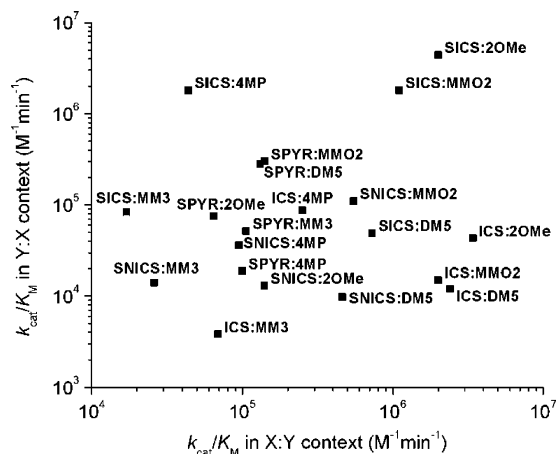


Figure 4. Second-order rate constants of Kf mediated extension of unnatural base pairs. The x-axis corresponds to the extension of the dX:dY context while the y-axis corresponds to the extension of the dY:dX context. Pairs are listed on the plot as X:Y. See Supporting Information (Table S2) for rate constants and sequence context.

(Figure 3, Table S1). In the template, dMMO2, d2OMe, and dDM5 direct dSICSTP incorporation at similar rates ($1.1 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$ to $1.7 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$); these rates are approximately 10-fold greater than the rate of dICSTP incorporation ($8.7 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$ to $1.6 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$). d4MP templates unnatural base pair incorporation poorly, with rates that are 50- to 150-fold lower than those for the other three phenyl analogues. The rates of incorporation of dSPYRTP are significantly less than those for either dICSTP or dSICSTP, confirming previous results indicating that the incorporation of hydrophobic base pairs tends to correlate strongly with the total aromatic surface area of the incoming triphosphate.

Nucleobase derivatization also has significant effects on the rate of unnatural pair extension (Figure 4, Table S2). In addition to the aforementioned analogues, we also examined MM3³⁶ to observe the effect of removing one of the minor groove H-bond acceptors, and SNICS,²⁸ to observe the effect of heteroatom substitution on the dSICS scaffold. When in the primer, the four dSICS analogues act similarly when paired with common templates. Each is extended most efficiently when paired with dMMO2, d2OMe, or dDM5 in the template. These rates are highest with dSICS and dICS, which are extended with rates greater than or equal to $1 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$. Pairs formed with dSNICS or dSPYR in the primer and dMMO2, d2OMe, or dDM5 in the template are extended less efficiently than their dSICS or dICS counterparts, but still more efficiently than those with either dMM3 or d4MP in the template. Pairs formed with dSICS, dICS, dSNICS, or dSPYR in the primer and d4MP in the template are extended with only moderate efficiency ($k_{\text{cat}}/K_{\text{M}}$ ranging from 2×10^4 to $3 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$), and pairs formed with dMM3 are generally extended the least efficiently ($k_{\text{cat}}/K_{\text{M}}$ ranging from 2×10^4 to $7 \times 10^4 \text{ M}^{-1} \text{ min}^{-1}$).

The results are significantly different in the opposite context, with the dMMO2 analogues in the primer and the dSICS analogues in the template (Figure 4, Table S2). Consistent with previous studies,^{20–24} primers bearing a hydrogen bond acceptor in the minor groove (i.e., d2OMe, dMMO2, and d4MP) were extended more efficiently than those bearing either hydrogen or a methyl group in the minor groove (dMM3 or dDM5), regardless of the templating nucleobase. The only exception was

(36) Matsuda, S.; Henry, A. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 6369–75.

Table 3. Steady-State Rate Constants of Self-Pair Synthesis by Incorporation of dXTP against dX for **SICS** and Analogues^a

5' – dTAATACGACTCACTATAGGGAGA 3' – dATTATGCTGAGTGATATCCCTCT(X)GCTAGGTTACGGCAGGATCGC				
X	k_{cat} (min ⁻¹)	K_M (μM)	k_{cat}/K_M (M ⁻¹ min ⁻¹)	
SICS ^b	5 ± 1	3.2 ± 3.1	1.6 × 10 ⁶	
4SICS	2.6 ± 0.8	8.6 ± 1.7	3.0 × 10 ⁵	
5SICS	1.7 ± 0.5	63 ± 4	2.7 × 10 ⁴	

^a See Experimental Section for experimental details. ^b Reference 23.

with **dSPYR** in the template, where primer extension was largely independent of the primer nucleotide. When paired with **dSICS** as the template, primers terminating with **d2OMe**, **dMMO2**, or **d4MP** are extended with rates greater than 1 × 10⁶ M⁻¹ min⁻¹. When paired with **dICS** or **dSNICS**, these same pairs are extended less efficiently (1.3 × 10⁴ to 1.5 × 10⁵ M⁻¹ min⁻¹), but still generally more efficiently than the pairs without the primer H-bond acceptor, which were extended with rates between 3.8 × 10³ and 1.4 × 10⁴ M⁻¹ min⁻¹. Surprisingly, the identity of the hydrogen bond acceptor was relatively unimportant, especially when paired with **dSICS**; in these cases, primers bearing methoxy groups (**d2OMe** or **dMMO2**) were extended essentially as efficiently as those bearing more natural, stronger H-bond accepting carbonyl groups (**d4MP**).

The rate of extension is also highly dependent on the template nucleobase. Pairs formed with **dSICS** in the template are consistently more efficiently extended than those formed with **dICS**, **dSNICS**, or **dSPYR**. The effects are the greatest when the primer nucleobase bears a methoxy group (either **d2OMe** or **dMMO2**), as the **dMMO2:dSICS** and **d2OMe:dSICS** pairs are extended greater than 100-fold more efficiently than **dMMO2:dICS** or **d2OMe:dICS**. The **dMMO2:dSICS** and **d2OMe:dSICS** pairs are also among the most efficiently extended pairs overall, with rates greater than 1 × 10⁶ M⁻¹ min⁻¹. Relative to **dSICS**, **dSNICS** is also a poorer template for all of the pairs examined; thus, the heteroatom located away from the interbase interface alters the extension properties of these pairs. In contrast, primer termini possessing **dSPYR** as the template nucleobase were extended at approximately the same rate (ranging from 6.4 × 10⁴ to 1.6 × 10⁵ M⁻¹ min⁻¹) regardless of primer nucleobase.

2.4. Optimization of dSICS:dMMO2: The d5SICS:dMMO2 Unnatural Base Pair. While the derivatizations described above help to elucidate the determinants of unnatural base pair replication, none resulted in significantly improved synthesis or extension. Also, as described above, the most significant limitation of the **dSICS:dMMO2** heteropair is the lack of selectivity against the **dSICS** self-pair, which is formed 5-fold more efficiently than the heteropair. Interestingly, previous studies with the isocarbostryl framework suggested that self-pair formation may be minimized by methyl substitution at the 4- or 5-position.²⁹ Thus, we synthesized and analyzed the **d4SICS** or **d5SICS** analogues (Figure 1B). When we measured the rates of self-pair synthesis for **d4SICS** and **d5SICS**, we found that the rates are decreased 5-fold and 60-fold, respectively, relative to **dSICS** (Table 3). Considering the substantial decrease in the self-pair synthesis of **d5SICS**, we examined whether this modification would modify the efficiency or fidelity of heteropair synthesis. Since **dMMO2** and **d2OMe** acted similarly at all steps in our structure–activity relationship,

Table 4. Steady-State Rate Constants of Kf-Mediated, dY-Templated, dXTP Incorporation for Synthesis of **d5SICS:dMMO2** and Mispairs^a

5' – dTAATACGACTCACTATAGGGAGA 3' – dATTATGCTGAGTGATATCCCTCT(Y)GCTAGGTTACGGCAGGATCGC				
dXTP	Y	k_{cat} (min ⁻¹)	K_M (μM)	k_{cat}/K_M (M ⁻¹ min ⁻¹)
5SICS	MMO2	8.5 ± 0.1	0.18 ± 0.01	4.7 × 10 ⁷
MMO2	5SICS	12 ± 1	33 ± 6	3.6 × 10 ⁵
5SICS	5SICS	1.7 ± 0.5	63 ± 4	2.7 × 10 ⁴
A	5SICS	1.2 ± 0.3	54 ± 6	2.2 × 10 ⁴
C	5SICS	nd ^b	nd ^b	<1.0 × 10 ³
G	5SICS	4.6 ± 0.6	36 ± 4	1.3 × 10 ⁵
T	5SICS	1.7 ± 0.2	130 ± 19	1.3 × 10 ⁴

^a See Experimental Section for experimental details. ^b Rates are too slow to determine k_{cat} and K_M independently.

Table 5. Steady-State Rate Constants of Kf-Mediated Unnatural Terminus Extension by Insertion of dCTP for **d5SICS:dMMO2** and Mispairs^a

5' – dTAATACGACTCACTATAGGGAGA(X) 3' – dATTATGCTGAGTGATATCCCTCT(Y)GCTAGGTTACGGCAGGATCGC				
X	Y	k_{cat} (min ⁻¹)	K_M (μM)	k_{cat}/K_M (M ⁻¹ min ⁻¹)
MMO2	5SICS	6.4 ± 1.1	3.4 ± 0.3	1.9 × 10 ⁶
5SICS	MMO2	3.8 ± 0.3	5.7 ± 1.2	6.7 × 10 ⁵
5SICS	5SICS	nd ^b	nd ^b	<1.0 × 10 ³
A	5SICS	0.79 ± 0.01	76 ± 11	1.0 × 10 ⁴
C	5SICS	0.18 ± 0.01	44 ± 9	4.2 × 10 ³
G	5SICS	0.62 ± 0.02	127 ± 8	4.9 × 10 ³
T	5SICS	2.0 ± 0.1	4.9 ± 0.3	4.0 × 10 ⁵

^a See Experimental Section for experimental details. ^b Rates are too slow to determine k_{cat} and K_M independently.

we examined the pairing of **d5SICS** with both **dMMO2** and **d2OMe**.

For **d5SICS**, **dMMO2** serves as a slightly better pairing partner than **d2OMe** (compare Tables 4 and 5 to Tables S3 and S4), and although the difference is small, we focused on pairing with **dMMO2**. Importantly, addition of the methyl group at the 5-position of the **dSICS** scaffold has virtually no effect on mispair formation with natural nucleotides but does increase the rate of pairing with **dMMO2** approximately 3-fold to a rate within 10-fold of natural synthesis. We then examined extension of the **d5SICS:dMMO2** heteropair (Table 5). The rates of extension are virtually unchanged in either context by methyl substitution of the isocarbostryl scaffold. Thus, it appears that methyl substitution at the 5-position results in a selective decrease of self-pair synthesis. In either strand context, the **d5SICS:dMMO2** unnatural base pair is synthesized and extended with rates that are reasonable and with a minimal overall fidelity of 130- and 6300-fold considering incorporation and extension (Table 6). This data compare favorably to those for any unnatural base pair reported to date.

3. Discussion

Typically, the design and development of unnatural base pairs have relied on the prevalent theories of molecular recognition, especially within the context of the DNA polymerase.^{3,37} Here, we have used two screens for efficient Kf-mediated synthesis, one which probes extension and another which probes synthesis and continued primer extension, to identify novel candidate base

(37) Ohtsuki, T.; Kimoto, M.; Ishikawa, M.; Mitsui, T.; Hirao, I.; Yokoyama, S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4922–4925.

Table 6. Fidelities of d5SICS:dMMO2 Synthesis and Extension

X*	Y	incorporation fidelity ^b	extension fidelity ^c	total fidelity ^d
5SICS	5SICS	13	1900	2.5×10^4
A	5SICS	16	190	3.0×10^3
C	5SICS	> 360	452	$> 1.6 \times 10^5$
G	5SICS	2.8	388	6.3×10^3
T	5SICS	28	4.8	1.3×10^2
MMO2	MMO2	392	126	4.9×10^4
A	MMO2^a	470	13	6.3×10^3
C, G, T	MMO2^a	>47000	<i>e</i>	<i>f</i>

^a Reference 23. ^b Relative rate of correct nucleoside triphosphate incorporation to incorrect dX*TP incorporation against template dY. ^c relative rate of correct pair extension to incorrect dX*:dY extension. ^d Total fidelity is the product of incorporation and extension fidelity. ^e Extension was not measured since incorporation product was not observed. ^f Total fidelity was not measured since extension was not measured.

pairs. From 3600 potential unnatural base pairs, the pair formed between dSICS and dMMO2 is clearly the strongest candidate base pair. Detailed kinetic analysis revealed that the pair is limited by facile synthesis of the dSICS:dSICS self-pair, which is likely driven by favorable packing interactions between the large aromatic rings.³⁴ Interestingly, the self-pairing problem appears to be a common problem for hydrophobic base pairs.^{8,16} Notably, the dDs:dPa base pair developed by Hirao and co-workers is, like dSICS:dMMO2, limited by self-pair formation of the larger nucleobase. To solve this problem, Hirao and co-workers modified the gamma phosphate of the large nucleobase to nonspecifically decrease the rate of dDsTP incorporation and, thus, self-pair formation.⁸ While this is a clever solution, it is likely to limit at least some of the potential applications in which the base pair might be used. To solve the same problem, we used synthetic derivatization of the pair, designed using a thorough structure–activity relationship as well as previous studies, to produce the candidate base pair d5SICS:dMMO2. Importantly, here, the decreased self-pair synthesis is highly selective; the synthetic modification has little to no effect on any other step of replication.

Steady-state kinetics demonstrated that all steps of enzymatic DNA synthesis of the unnatural base pair are within 1000-fold of the efficiency of a natural base pair, and considering the fidelity of base pair synthesis and extension, the overall selectivity against the most efficiently formed mispair is 130- and 6300-fold, with d5SICS and dMMO2 in the template, respectively (Table 6). Previously we reported the characterization of the d3FB self-pair,^{10,34} which was until now our best candidate for the expansion of the genetic alphabet. The d3FB self-pair is synthesized 20-fold slower than d5SICSTP is inserted opposite dMMO2, but 8-fold faster than dMMO2TP is inserted opposite d5SICS. However, in either sequence context, the d5SICS:dMMO2 pair is extended significantly faster than the d3FB self-pair. In addition, the overall fidelity of the d5SICS:dMMO2 pair is significantly greater than that for the d3FB self-pair, which is 20. Thus, relative to the d3FB self-pair, the d5SICS:dMMO2 pair is more efficiently synthesized and extended with overall greater fidelity. Although direct comparison with base pairs developed by other groups is difficult due to differing sequence contexts, the fidelity and efficiency of the d5SICS:dMMO2 pair appear to compare favorably with those for previously reported dDs:dPa,⁸ dc:dX,³⁸ and disoC:disoG^{3,39} unnatural base pairs.

We have made initial progress toward understanding the origins of the remarkable properties of the dSICS:dMMO2 and d5SICS:dMMO2 base pairs by analyzing the relationship between nucleobase shape, electronics, or H-bonding potential and Kf-mediated synthesis and extension (summarized in Table 7). During incorporation of the unnatural nucleoside, hydrophobic substituents at the position *ortho* to the glycosidic linkage, which are most likely oriented into the minor groove, are required for efficient synthesis. When the analogue in either the template or triphosphate bears a hydrophilic group at this position, on either scaffold, the rates of unnatural dNTP incorporation decrease significantly. In a manner similar to natural mispairing, uncompensated desolvation of hydrophilic functional groups likely renders dNTP incorporation energetically unfavorable. Thus, in agreement with the previously reported characterization of other hydrophobic analogues, efficient unnatural base pair synthesis is strongly facilitated by hydrophobic groups disposed in the developing minor groove.^{9,16,36} Interestingly, effects in the major groove tend to be more variable; methyl substitution at the 4-position of dMMO2 has only small effects while methyl substitution at the 5-position of the dSICS scaffolds profoundly decreases the rate of self-pair formation. The effects on self-pair synthesis may result from cross-strand eclipsing interactions manifest only in the intercalated mode of pairing that is unique to the self-pair.³⁴

The most efficiently extended termini all had a minor groove H-bond acceptor in the primer and a relatively hydrophobic minor groove substituent in the template nucleobase. Biochemical studies of natural DNA have clearly shown that H-bond acceptors in the minor groove of the primer nucleobase facilitate extension;^{40–43} interactions between these functional groups and Kf, particularly H-bonds formed between the minor groove H-bond acceptor of the primer nucleobase, Arg668 of Kf, and the ribosyl oxygen of the incoming dNTP, are necessary for the polymerase to properly orient the primer terminus and the incoming triphosphate for catalysis.⁴² We have repeatedly observed the same phenomenon in unnatural scaffolds.^{18–22} Here, this idea is further supported by the significantly decreased efficiency of extension of primers terminating in either dMM3 or dDM5. Despite the importance of this H-bond, the strength of the interaction appears to be less important; within the dSICS scaffold, the oxygen and sulfur substituted analogues are extended essentially equivalently; within the dMMO2 scaffold, the methoxy and carbonyl variants are extended similarly as well.

In addition to a minor groove hydrogen bond acceptor on the primer nucleobase, efficient extension appears to require a hydrophobic substituent oriented into the minor groove of the template nucleobase, although the origins of this effect are less clear. Examination of the various crystal structures of Pol I polymerases shows evolutionarily conserved H-bonding contacts between the polymerase and minor groove H-bond acceptors of the template nucleobase.^{44–47} However, in stark contrast to

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

(39) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489–10496.

(40) Morales, J. C.; Kool, E. T. *J. Am. Chem. Soc.* **1999**, *121*, 2323–2324.

(41) Morales, J. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, (6), 1001–1007.

(42) Meyer, A. S.; Blandino, M.; Spratt, T. E. *J. Biol. Chem.* **2004**, *279*, 33043–6.

(43) Spratt, T. E. *Biochemistry* **2001**, *40*, 2647–52.

(44) Li, Y.; Korolev, S.; Waksman, G. *EMBO J.* **1998**, *17*, 7514–7525.

Table 7. Relative Rates of Incorporation and Extension for dSICS and dMMO2 Analogues^a

	X	Y	Relative Rate of Incorporation ^b		Relative Rate of Extension ^c	
			dXTP: dY	dYTP: dX	X:Y	Y:X
	SICS	MMO2	1	1	1	1
	ICS	MMO2	0.10	0.03	1.82	0.01
	SPYR	MMO2	0.01	0.26	0.28	0.08
	SNICS	MMO2	nd ^d	nd ^d	0.50	0.06
	5SICS	MMO2	3.13	1.06	0.6	1.1
	SICS	2OMe	1.10	1.26	1.82	2.44
	SICS	DM5	1.50	0.38	0.66	0.03
	SICS	4MP	0.01	0.03	0.04	1.00
	SICS	MM3	nd ^d	nd ^d	0.05	0.02

^a Relative rates are defined as the ratio of the second-order rate constants of the X:Y pair to the SICS:MMO2 pair. See Supporting Information for experimental details. ^b Incorporation of dXTP against dY or dYTP against dX. ^c Extension of X:Y or Y:X. Bases are listed primer:template. ^d Not determined.

the interactions between the polymerase and the primer nucleobase, biochemical disruption of this contact, by atomic substitution of the H-bonding group in the nucleobase, results in only small or insignificant changes in the rate of enzymatic synthesis of natural DNA.^{41,43} Thus, this template H-bond appears to play a smaller role in the synthesis of natural DNA. Interestingly, with hydrophobic unnatural base pairs, it appears that increased hydrophobicity in the template position of the developing minor groove actually facilitates extension.²³ We observed a large reduction in extension associated with substitution of the sulfur atom of dSICS with an oxygen, or the substitution of the C-glycosidic methoxy analogue of dMMO2 with an N-glycosidic amide. Although the carbonyl functional group is more similar to a natural nucleobase and should H-bond with the polymerase at the conserved residues, it significantly decreases the rate of extension in either scaffold. Additionally, within the dMMO2 scaffold, substituting the methoxy group with a hydrogen (dMM3) results in a significant decrease in extension while substituting it with a methyl group (dDM5) has little effect.

The structure–activity relationship data described above suggest that functionality within the developing minor groove is generally most important for polymerase recognition and replication of hydrophobic base pairs. However, at first glance, these requirements appear to be mutually exclusive: efficient synthesis requires a hydrophobic substituent in both the primer and template analogue while subsequent extension requires an

H-bond acceptor in the primer nucleobase and a hydrophobic substituent in the template nucleobase. These conflicting requirements for polymerase mediated recognition appear to explain one of the most conspicuous and, according to our structure activity relationship studies, essential aspects of the dSICS:dMMO2 and d5SICS:dMMO2 unnatural base pairs—the unusual substituents at the position *ortho* to the glycosidic linkage. dMMO2 possesses an *ortho* methoxy substituent which provides both a hydrophobic methyl group and a potential H-bond acceptor; dSICS and d5SICS possess an *ortho* sulfur atom which, relative to an oxygen atom, is more hydrophobic but retains the ability to act as an H-bond acceptor. Thus, the dSICS:dMMO2 and d5SICS:dMMO2 pairs appear to simultaneously satisfy the apparently contradictory demands of efficient synthesis and extension. It is unlikely that such a subtle balance of properties would have been designed rationally.

Our characterization of the d5SICS:dMMO2 heteropair demonstrates that its synthesis and extension are both relatively efficient and selective. The pair compares favorably with all previously characterized unnatural base pair candidates for which relevant kinetic data are available. Additionally, our structure–activity relationship, along with the growing literature regarding hydrophobic base pairs, begins to suggest discreet design principles which will aid in the development of future unnatural base pairs. Indeed, we continue to search for modifications that optimize the d5SICS:dMMO2 heteropair as well as new scaffolds with promising properties with the expectation that this work will not only elucidate fundamental principles underlying the storage and retrieval of genetic

(45) Li, Y.; Waksman, G. *Protein Sci.* **2001**, *10*, 1225–1233.

(46) Beese, L. S.; Derbyshire, V.; Steitz, T. A. *Science* **1993**, *260*, 352–355.

(47) Keifer, J. R.; Mao, C.; Hansen, C. J.; Basehore, S. L.; Hogrefe, H. H.; Braman, J. C.; Beese, L. *Structure* **1997**, *5*, 95–108.

information but also ultimately culminate in a fully functional unnatural base pair.

Experimental Methods

Extension Screen - Primer Group:Template Screening. Primers, varying at position X of the sequence 5'-dTAA TAC GAC TCA CTA TAG GGA GAX-3', were grouped into 10 groups based loosely on their functional groups and general shape. Primer pools consisting of (n) nucleobases were constructed by equally representing all primers (final concentration of each individual primer = 1 $\mu\text{M}/n$) into a pool with a total DNA concentration of 1 μM . Primer pools were 5' radiolabeled with [γ - ^{33}P]-ATP (GE Healthcare) and T4 polynucleotide kinase (New England Biolabs). Primer pools were annealed with 2-fold excess template of the complementary sequence, 3'-dATT ATG CTG AGT GAT ATC CCT CTY GCT AGG TTA CGG CAG GAT CGC-5', in the reaction buffer by heating to 90 °C and slow cooling to room temperature. Assay conditions include the following: 40 nM total primer–template duplex, 0.3 nM enzyme, 50 mM Tris buffer (pH 7.5), 10 mM MgCl_2 , 1 mM DTT, 50 $\mu\text{g}/\text{mL}$ BSA, 400 μM dCTP. Prior to reaction, dCTP was mixed with enzyme, and the reactions were initiated by adding the dCTP–enzyme mixture to an equal volume (5 μL) of a 2 \times DNA stock solution, incubated at 25 °C for 5 min, and quenched with 20 μL of loading buffer (95% formamide, 20 mM EDTA). The reaction mixture (3 μL) was then analyzed by 15% polyacrylamide gel electrophoresis. Radioactivity was quantified using a Phosphorimager (Molecular Dynamics) with overnight exposures and the ImageQuant program. Percent conversion was defined as the ratio of singly extended product to the sum of the singly extended product and the unextended product. The top 7% of the pairs were carried on to individual primer:template analysis.

Extension Screen - Individual Primer:Template. Individual primer, in the same sequence context as above, was 5' radiolabeled with [γ - ^{33}P]-ATP and T4 polynucleotide kinase. Primer–template duplexes were annealed in the reaction buffer by heating to 90 °C and slow cooling to room temperature. Assay conditions were as follows: 40 nM template–primer duplex, 0.3 nM enzyme, 50 mM Tris buffer (pH 7.5), 10 mM MgCl_2 , 1 mM DTT, and 50 $\mu\text{g}/\text{mL}$ BSA. The reactions were initiated by adding the DNA–enzyme mixture to an equal volume (5 μL) of a 2 \times dNTP stock solution resulting in a final concentration of 0, 10, 100, or 1000 μM dCTP, incubated at 25 °C for 5 min, and quenched with 20 μL of loading buffer (95% formamide, 20 mM EDTA). The reaction mixture (8 μL) was then analyzed by 15% polyacrylamide gel electrophoresis. Radioactivity was quantified using a Phosphorimager (Molecular Dynamics) with overnight expo-

sure and the ImageQuant program. The Michaelis–Menten equation was fit to the data using the program Kaleidagraph (Synergy software). The data used were from single experiments.

Full Length Synthesis Screen. Primer (5'-d(CGACTCACTAT-AGGGAGA)) and templates (5'-d(CGCTAGGACGGCATTGGATCGX-TCTCCCTATAGTGAGTCGTATTA)) were annealed in reaction buffer by heating to 95 °C followed by slow cooling to room temperature. A mixture of the four natural dNTPs was added, and 5 μL of the solution were aliquoted into a 384-well plate. To each well was added either a 2.5 μL aliquot of an unnatural dNTP (60 reactions, final concentration of 100 μM) or water (misincorporation control) followed by addition of 2.5 μL of Kf (exo-) in reaction buffer. Reaction conditions were as follows: 80 nM primer, 40 nM template, 0.4 nM Kf, 10 μM natural dNTPs, 50 mM Tris-HCl (pH 7.5), 10 mM MgCl_2 , 1 mM DTT, and 50 $\mu\text{g}/\text{mL}$ acetylated BSA. A negative control, in the absence of dNTPs, and a positive control, where dX = dA, were run simultaneously to all experiments. The reaction was incubated at 25 °C for 1 h and quenched by addition of 10 μL of quenching solution (25 mM Tris-HCl (pH 7.5), 5 mM EDTA, 2 \times SYBR Green I (Invitrogen)) and then incubated for 4 min in the dark, and the fluorescence intensities were quantified using a fluorescence plate reader (Spectra MAX Gemini, Molecular Devices) with excitation at 485 nm and emission at 538 nm. The ratio of fluorescence in the reaction containing both dNTPs and dXTP to the fluorescence of the reaction only containing dNTPs was used as a measure of both efficiency and selectivity of full length synthesis.

General Steady-State Kinetic Assay Protocol. Experiments were performed as in *Extension Screen – Individual Primer:Template* with the following exceptions. To ensure steady-state conditions,³⁵ enzyme concentrations ranged from 0.06 to 1.2 nM Kf and incubation times ranged from 2 to 12 min. Additionally, each experiment used nine triphosphate concentrations to more accurately calculate K_M . Data presented are the result of triplicate experiments.

Acknowledgment. Funding was provided by the National Institutes of Health (GM60005 to F.E.R.) and the Uehara Memorial Foundation (Y.H.).

Supporting Information Available: Details of synthesis and characterization of nucleotides; additional kinetic data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA078223D