

Optimization of an Unnatural Base Pair toward Natural-Like Replication

Young Jun Seo, Gil Tae Hwang, Phillip Ordoukhanian, and Floyd E. Romesberg*

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

Received October 7, 2008; E-mail: floyd@scripps.edu

Abstract: Predominantly hydrophobic unnatural nucleotides that selectively pair within duplex DNA as well as during polymerase-mediated replication have recently received much attention as the cornerstone of efforts to expand the genetic alphabet. We recently reported the results of a screen and subsequent lead hit optimization that led to identification of the unnatural base pair formed between the nucleotides **dMMO2** and **d5SICS**. This unnatural base pair is replicated by the Klenow fragment of *Escherichia coli* DNA polymerase I with better efficiency and fidelity than other candidates reported in the literature. However, its replication remains significantly less efficient than a natural base pair, and further optimization is necessary for its practical use. To better understand and optimize the slowest step of replication of the unnatural base pair, the insertion of **dMMO2** opposite **d5SICS**, we synthesized two **dMMO2** derivatives, **d5FM** and **dNaM**, which differ from the parent nucleobase in terms of shape, hydrophobicity, and polarizability. We find that both derivatives are inserted opposite **d5SICS** more efficiently than **dMMO2** and that overall the corresponding unnatural base pairs are generally replicated with higher efficiency and fidelity than the pair between **dMMO2** and **d5SICS**. In fact, in the case of the **dNaM:d5SICS** heteropair, the efficiency of each individual step of replication approaches that of a natural base pair, and the minimum overall fidelity ranges from 10^3 to 10^4 . In addition, the data allow us to propose a generalized model of unnatural base pair replication, which should aid in further optimization of the unnatural base pair and possibly in the design of additional unnatural base pairs that are replicated with truly natural-like efficiency and fidelity.

1. Introduction

Development of an unnatural base pair to supplement the natural base pairs, dA:dT and dG:dC, would increase the biotechnological utility of DNA and is the first, and likely most challenging, step toward expansion of the genetic code and creation of a semisynthetic organism. Such efforts were pioneered in the Benner laboratory through the design of nucleotides with unnatural hydrogen-bonding (H-bonding) topologies.^{1–5} More recent work from our laboratory^{6–13} and

the Hirao laboratory^{14–18} has built upon the remarkable observations from the Kool laboratory^{19–21} that H-bonds are not required for base pair replication and has focused on the pairing of nucleotides bearing predominantly hydrophobic nucleobases. These efforts have relied on rational design and structure–activity relationships^{22–27} and have yielded several unnatural base pair candidates that are replicated surprisingly

- (1) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322–8323.
- (2) Geyer, C. R.; Battersby, T. R.; Benner, S. A. *Structure (Cambridge, MA, U.S.)* **2003**, *11*, 1485–1498.
- (3) Moser, M. J.; Marshall, D. J.; Greiner, J. K.; Kieffer, C. D.; Killeen, A. A.; Ptacin, J. L.; Richmond, C. S.; Roesch, E. B.; Scherrer, C. W.; Sherrill, C. B.; Van Hout, C. Z.; Zanton, S. J.; Prudent, J. R. *Clin. Chem.* **2003**, *49*, 407–414.
- (4) Prudent, J. R. *Expert Rev. Mol. Diagn.* **2006**, *6*, 245–252.
- (5) Yang, Z.; Sismour, A. M.; Sheng, P.; Puskar, N. L.; Benner, S. A. *Nucleic Acids Res.* **2007**, *35*, 4238–4249.
- (6) Ogawa, A. K.; Wu, Y.; Berger, M.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 8803–8804.
- (7) Henry, A. A.; Yu, C.; Romesberg, F. E. *J. Am. Chem. Soc.* **2003**, *125*, 9638–9646.
- (8) Leconte, A. M.; Matsuda, S.; Hwang, G. T.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 4326–4329.
- (9) 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*, 10466–10473.
- (10) Leconte, A. M.; Hwang, G. T.; Matsuda, S.; Capek, P.; Hari, Y.; Romesberg, F. E. *J. Am. Chem. Soc.* **2008**, *130*, 2336–2343.
- (11) Matsuda, S.; Leconte, A. M.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 5551–5558.
- (12) Hwang, G. T.; Romesberg, F. E. *J. Am. Chem. Soc.* **2008**, *130*, 14872–14882.
- (13) Yu, C.; Henry, A. A.; Romesberg, F. E.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2002**, *41*, 3841–3844.
- (14) Mitsui, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S. *J. Am. Chem. Soc.* **2003**, *125*, 5298–5307.
- (15) Hirao, I. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622–627.
- (16) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nat. Methods* **2006**, *3*, 729–735.
- (17) Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. *J. Am. Chem. Soc.* **2007**, *129*, 15549–15555.
- (18) Hirao, I.; Harada, Y.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Yokoyama, S. *J. Am. Chem. Soc.* **2004**, *126*, 13298–13305.
- (19) Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191–6192.
- (20) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704–708.
- (21) Morales, J. C.; Kool, E. T. *Biochemistry* **2000**, *39*, 12979–12988.
- (22) McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585–11586.
- (23) Matsuda, S.; Henry, A. A.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2003**, *125*, 6134–6139.
- (24) Henry, A. A.; Olsen, A. G.; Matsuda, S.; Yu, C.; Geierstanger, B. H.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 6923–6931.

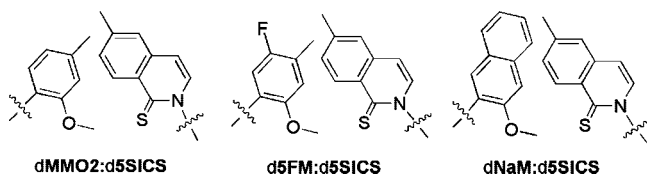


Figure 1. Unnatural base pairs used in this study. Only the nucleobase analogues are depicted, with the sugar and phosphates omitted for clarity.

well considering their unnatural structures.^{11,18,24} However, none of the candidates are replicated with the efficiency and fidelity of the natural base pairs. In addition, the effect of sequence on replication of the candidate unnatural base pairs remains largely unexplored, and while small variations are expected, any realistic unnatural base pair candidate must be shown to be replicated efficiently and accurately regardless of sequence context.

Efforts based on rational design have helped to elucidate the determinants of stability and efficient replication of DNA, but they are limited by our incomplete understanding of the forces involved. As an alternative to rational design, we recently screened a pool of 3600 predominantly hydrophobic base pair candidates to identify those that are most efficiently replicated by the exonuclease-deficient Klenow fragment of *Escherichia coli* DNA Pol I (Kf).¹⁰ Initial optimization of the most promising lead yielded the heteropair formed between dMMO2 and d5SICS (Figure 1), which is well recognized not only by Kf¹⁰ but also by a variety of other DNA polymerases.¹² In the sequence context examined, Kf synthesizes the d5SICS:dMMO2 heteropair (where dY:dX denotes the strand context with dY in the primer and dX in the template) with a second-order rate constant ($k_{\text{cat}}/K_{\text{M}}$, or efficiency) that is only 7-fold smaller than that of natural synthesis, but then continues primer extension by incorporation of the next correct dNTP with a rate that is 250-fold slower than extension of a natural base pair. Synthesis of the heteropair in the opposite strand context (i.e., dMMO2:d5SICS) is the rate-limiting step of replication, proceeding nearly 1000-fold slower than synthesis of a natural base pair, while extension is more efficient, occurring only 90-fold slower than extension of a natural base pair. Thus, although the unnatural base pair formed between d5SICS and dMMO2 is replicated more efficiently than those identified previously by rational design, neither the efficiency nor the fidelity of its replication approaches that of natural synthesis, and further optimization is still required.

Structure-activity relationship studies with both dMMO2 and d5SICS revealed that the substituents ortho to the glycosidic linkage—that is, the methoxy group of dMMO2 and the sulfur atom of d5SICS—are essential for polymerase recognition.^{10,11,13} These substituents likely act as H-bond acceptors in the developing minor groove, where they are required to engage polymerase-based H-bond donors.^{28–31} In contrast, substituents

positioned meta or para to the glycosidic linkage, which are expected to be disposed in the developing major groove, do not appear to directly engage the polymerase, and their effect on heteropair replication is more variable. Thus, the meta and para positions of dMMO2 and d5SICS appear to be best suited for derivatizations aimed at improving replication.

Here we report our efforts to optimize the heteropair by meta and para derivatization of dMMO2, whose incorporation into DNA is the slowest step of replication. We characterized the effects of a fluorine substituent at the meta position (d5FM, Figure 1) and increased aromatic surface area associated with ring fusion at the meta, para positions (dNaM, Figure 1). To more fully understand the generality of the unnatural base pairs, we fully characterized replication in two different sequence contexts. We find that both d5FMTP and dNaMTP are inserted opposite d5SICS with significantly greater efficiency than dMMO2TP, and moreover, we find that neither modification significantly interferes with any other step of replication. In fact, in the two sequence contexts examined, each individual step of replication of the heteropair formed between dNaM and d5SICS is within 8–140-fold and 6–490-fold, respectively, of a natural base pair, and overall fidelity ranges from 10^3 to 10^4 . Thus, the efficiency and fidelity of the heteropair's replication by Kf are beginning to approach those of natural synthesis and are likely sufficient for in vitro applications,^{32–40} and possibly for the initiation of in vivo efforts as well.⁴¹ In addition, the data allow us to propose a unified model of unnatural base pair replication, in which intrastrand intercalation plays a central role and which should prove invaluable for further optimization of these and other unnatural base pairs.

2. Results

The unnatural nucleotides d5FM and dNaM were designed to explore the effects of altered nucleobase polarity and surface area and were synthesized as described in the Experimental Section and Supporting Information. Both nucleotides were converted to the corresponding triphosphate or phosphoramidite, and each phosphoramidite was incorporated into primer and template oligonucleotides via automated DNA synthesis. Heteropair synthesis is gauged by determining the second-order rate constant ($k_{\text{cat}}/K_{\text{M}}$) for insertion of an unnatural triphosphate opposite an unnatural nucleotide in the template under steady-state conditions;⁴² extension is gauged by determining $k_{\text{cat}}/K_{\text{M}}$ for insertion of the next correct dNTP. Fidelities are calculated as the ratio of the second-order rate constant for synthesis or

- (25) Mitsui, T.; Kimoto, M.; Harada, Y.; Yokoyama, S.; Hirao, I. *J. Am. Chem. Soc.* **2005**, *127*, 8652–8658.
 (26) Zhang, X.; Lee, I.; Zhou, X.; Berdis, A. J. *J. Am. Chem. Soc.* **2006**, *128*, 143–149.
 (27) Kojima, N.; Inoue, K.; Nakajimi-Shibata, R.; Kawahara, S.; Ohtsuka, E. *Nucleic Acids Res.* **2003**, *31*, 7175–7188.
 (28) Spratt, T. E. *Biochemistry* **2001**, *40*, 2647–2652.
 (29) Leconte, A. M.; Chen, L.; Romesberg, F. E. *J. Am. Chem. Soc.* **2005**, *127*, 12470–12471.
 (30) Li, Y.; Waksman, G. *Protein Sci.* **2001**, *10*, 1225–1233.
 (31) Meyer, A. S.; Blandino, M.; Spratt, T. E. *J. Biol. Chem.* **2004**, *279*, 33043–33046.

- (32) Keefe, A. D.; Cload, S. T. *Curr. Opin. Chem. Biol.* **2008**, *12*, 448–456.
 (33) Shamah, S. M.; Healy, J. M.; Cload, S. T. *Acc. Chem. Res.* **2008**, *41*, 130–138.
 (34) Seeman, N. C. *Mol. Biotechnol.* **2007**, *37*, 246–257.
 (35) Seeman, N. C. *Trends Biochem. Sci.* **2005**, *30*, 119–235.
 (36) Ieven, M. *J. Clin. Virol.* **2007**, *40*, 259–276.
 (37) Svarovskaia, E. S.; Moser, M. J.; Bae, A. S.; Prudent, J. R.; Miller, M. D.; Borroto-Esoda, K. *J. Clin. Microbiol.* **2006**, *44*, 4237–4241.
 (38) Moser, M. J.; Christensen, D. R.; Norwood, D.; Prudent, J. R. *J. Mol. Diagn.* **2006**, *8*, 89–96.
 (39) Pietz, B. C.; Warden, M. B.; DuChateau, B. K.; Ellis, T. M. *Hum. Immunol.* **2005**, *66*, 1174–1182.
 (40) Johnson, S. C.; Marshall, D. J.; Harms, G.; Miller, C. M.; Sherrill, C. B.; Beaty, E. L.; Lederer, S. A.; Roesch, E. B.; Madsen, G.; Hoffman, G. L.; Laessig, R. H.; Kopish, G. J.; Baker, M. W.; Benner, S. A.; Farrell, P. M.; Prudent, J. R. *Clin. Chem.* **2004**, *50*, 2019–2027.
 (41) Wu, Y.; Fa, M.; Tae, E. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2002**, *124*, 14626–14630.
 (42) Creighton, S.; Bloom, L. B.; Goodman, M. F. *Methods Enzymol.* **1995**, *262*, 232–256.

Table 1. Second-Order Rate Constants (k_{cat}/K_M Values) for Heteropair and Mispair Synthesis and Extension with d**5SICS** in the Template (Context I)^a

5' -TAATACGACTCACTATAGGGAGC (Y)
 3' -ATTATGCTGAGTGATATCCCTCG (X) TCTAGGTTACGGCAGGATCGC

X	Y	synthesis ($M^{-1} \text{ min}^{-1}$)	extension ($M^{-1} \text{ min}^{-1}$)
T	A	$(3.2 \pm 0.6) \times 10^8$ ^b	$(1.7 \pm 0.2) \times 10^8$ ^b
5SICS	MMO2	$(3.6 \pm 0.7) \times 10^5$	$(1.9 \pm 0.2) \times 10^6$
	5FM	$(3.6 \pm 0.4) \times 10^6$	$(5.5 \pm 0.9) \times 10^6$
	NaM	$(5.0 \pm 0.1) \times 10^6$	$(1.2 \pm 0.4) \times 10^6$
	5SICS	$(2.7 \pm 0.9) \times 10^4$ ^c	$<1.0 \times 10^3$ ^c
	A	$(2.2 \pm 0.3) \times 10^4$ ^c	1.0×10^4 ^c
	G	$(1.3 \pm 0.8) \times 10^5$ ^c	$(4.9 \pm 0.5) \times 10^3$ ^c
	C	$<1.0 \times 10^3$ ^c	4.2×10^3 ^c
	T	$(1.3 \pm 0.4) \times 10^4$ ^c	$(4.0 \pm 0.2) \times 10^5$ ^c

^a See Experimental Section for details. ^b References 12 and 43. ^c Reference 10.

extension of the correct unnatural base pair to that for a mispair. Each heteropair and mispair is fully characterized in both possible strand contexts (i.e., dY:dX and dX:dY) and the minimum overall fidelity is determined as the product of the minimum single-step fidelities.

To evaluate the generality of the results, replication (synthesis and extension) of the heteropairs was characterized in two sequence contexts. Sequence context I is the same as that used previously to characterize a wide variety of unnatural base pairs,^{6–13,22–24,43–46} including the original analysis of the heteropair formed between d**MMO2** and d**5SICS**,¹⁰ and in this context the unnatural nucleotide in the template is flanked by a 3' dT and a 5' dG. In sequence context II, the natural nucleotides are inverted so that the unnatural nucleotide is flanked by a 3' dG and a 5' dT. Sequence contexts III and IV are the same as context I except that the 3' nucleotide that flanks the unnatural base in the template is a dA or a dC, respectively.

2.1. Unnatural Heteropair Replication in Sequence Context I. 2.1.1. Synthesis and Extension with d5SICS in the Template. We first examined the Kf-mediated insertion of d**5FMTP** and d**NaMTP** opposite d**5SICS** in sequence context I. For reference, the insertion of d**MMO2TP** opposite d**5SICS**, which is the step that most limits the replication of the parental heteropair in this sequence context, proceeds with a second-order rate constant of $3.6 \times 10^5 M^{-1} \text{ min}^{-1}$ (Table 1). The moderately efficient insertion of dGTP results in a minimum synthesis fidelity of only 3 for correct d**MMO2**:d**5SICS** heteropair synthesis in this sequence context. Characterization of the derivative base pairs in sequence context I revealed that Kf inserts d**5FMTP** and d**NaMTP** opposite d**5SICS** with efficiencies that are 10- and 14-fold higher than that for d**MMO2TP**, respectively (Table 1). Thus, the rate of the slowest step of heteropair synthesis is increased from 888-fold slower than natural synthesis in the case of d**MMO2** to only 88- or 64-fold slower for d**5FM** and d**NaM**, respectively. Correspondingly, the minimum synthesis fidelity is increased from only 3 for the synthesis of the d**MMO2**:d**5SICS** pair to 28 for d**5FM**:d**5SICS** and 38 for d**NaM**:d**5SICS**.

(43) Hwang, G. T.; Leconte, A. M.; Romesberg, F. E. *ChemBioChem* **2007**, *8*, 1606–1611.

(44) Hwang, G. T.; Romesberg, F. E. *Nucleic Acids Res.* **2006**, *34*, 2037–2045.

(45) Kim, Y.; Leconte, A. M.; Hari, Y.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7809–7812.

(46) Matsuda, S.; Henry, A. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 6369–6375.

Table 2. Second-Order Rate Constants (k_{cat}/K_M Values) for Heteropair and Mispair Synthesis and Extension with d**MMO2** or a Derivative in the Template (Context I)^a

5' -TAATACGACTCACTATAGGGAGC (Y)
 3' -ATTATGCTGAGTGATATCCCTCG (X) TCTAGGTTACGGCAGGATCGC

X	Y	synthesis ($M^{-1} \text{ min}^{-1}$)	extension ($M^{-1} \text{ min}^{-1}$)
MMO2	5SICS	$(4.7 \pm 0.4) \times 10^7$ ^b	$(6.7 \pm 1.1) \times 10^5$ ^b
	MMO2	$(1.2 \pm 0.1) \times 10^5$ ^c	$(5.3 \pm 0.9) \times 10^3$ ^c
	A	$(1.0 \pm 0.1) \times 10^5$ ^c	$(4.6 \pm 0.2) \times 10^4$ ^c
	G	$<1.0 \times 10^3$ ^c	$<1.0 \times 10^3$ ^c
5FM	5SICS	$(1.4 \pm 0.2) \times 10^7$	$(2.3 \pm 0.2) \times 10^6$
	5FM	$(8.9 \pm 0.7) \times 10^5$	$(2.6 \pm 0.8) \times 10^4$
	A	$(3.0 \pm 0.5) \times 10^5$	$(3.2 \pm 0.4) \times 10^5$
	G	$<1.0 \times 10^3$	$<1.0 \times 10^3$
NaM	5SICS	$(3.7 \pm 0.4) \times 10^7$	$(2.7 \pm 0.2) \times 10^6$
	NaM	$(3.4 \pm 0.8) \times 10^6$	$(1.1 \pm 0.3) \times 10^4$
	A	$(7.1 \pm 2.9) \times 10^5$	$(3.9 \pm 1.8) \times 10^4$
	G	$<1.0 \times 10^3$	$<1.0 \times 10^3$
	C	$<1.0 \times 10^3$	$(1.5 \pm 0.1) \times 10^6$
	T	$(5.6 \pm 0.5) \times 10^3$	$(5.7 \pm 0.3) \times 10^5$

^a See Experimental Section for details. ^b Reference 10. ^c Reference 9.

Perhaps the most remarkable attribute of the parental heteropair formed between d**MMO2** and d**5SICS**, relative to other unnatural base pairs that have been investigated, is its relatively efficient extension in both possible strand contexts. For d**MMO2**:d**5SICS** in sequence context I, the efficiency of extension is $1.9 \times 10^6 M^{-1} \text{ min}^{-1}$, which is only 90-fold lower than a natural base pair.¹⁰ The most efficiently extended mispair in this strand and sequence context is dT:d**5SICS**, which limits the minimum fidelity for correct heteropair extension to 5 (Table 1). However, since the synthesis of dT:d**5SICS** is not efficient, its impact on replication fidelity is minimized. As with d**MMO2**:d**5SICS**, extension of d**5FM**:d**5SICS** and d**NaM**:d**5SICS** is also very efficient in this sequence context (Table 1), resulting in minimum extension fidelities of 14 and 3, respectively. Thus, both of the derivatizations improve synthesis without interfering with extension, and when the effects on both steps are combined, these modifications increase the minimum overall replication fidelity from 130 for d**MMO2**:d**5SICS** to 3800 and 1260 for d**5FM**:d**5SICS** and d**NaM**:d**5SICS**, respectively.

2.1.2. Synthesis and Extension with dMMO2, d5FM, or dNaM in the Template. We next investigated whether the derivatization of d**MMO2** affects the efficiency or fidelity of replication in the opposite strand context (Table 2). For reference, with d**MMO2** in the template, Kf inserts d**5SICSTP** with an efficiency of $4.7 \times 10^7 M^{-1} \text{ min}^{-1}$, and while it does not insert dGTP, dCTP, or dTTP, it does insert dATP and d**MMO2TP**, albeit 470- and 390-fold slower than d**5SICSTP**, respectively. Kf inserts d**5SICSTP** slightly less efficiently with d**5FM** in the template than with d**MMO2** in the template; however, at $1.4 \times 10^7 M^{-1} \text{ min}^{-1}$, the efficiency remains high. Kf does not insert dGTP, dCTP, or dTTP opposite d**5FM**; however, it does insert d**5FMTP** with moderate efficiency, and dATP only slightly less efficiently, which results in the reduced minimum synthesis fidelity of 16 for d**5SICS**:d**5FM**. Problematically, the dA:d**5FM** mispair is also extended with moderate efficiency (see below), and this ultimately limits the overall fidelity of the d**5FM**-based heteropair in this strand and sequence context.

With d**NaM** in the template, the insertion of d**5SICSTP** is very efficient, in fact, slightly more efficient than with d**5FM**

in the template and only 8-fold less efficient than insertion of dATP opposite dT (Table 2). Neither dGTP nor dCTP is inserted opposite dNaM, but dTTP and dATP are inserted slightly more efficiently than they are opposite d5FM, and dNaMTP is inserted opposite itself with an efficiency of $3.4 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$, which limits the minimum synthesis fidelity of the d5SICS:dNaM heteropair in this sequence context to 11. However, the relatively inefficient extension of these mispairs minimizes their impact on overall fidelity (see below).

Relative to its synthesis, extension of the d5SICS:dMMO2 heteropair in sequence context I is somewhat less efficient ($k_{\text{cat}}/K_{\text{M}} = 6.7 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$, 254-fold lower than natural synthesis; Table 2), and this is the second most limiting step in replication of the parental heteropair. Characterization of mispair extension (Table 2) revealed that the dG:dMMO2 mispair is not efficiently extended, while the mispair with dA is extended more efficiently ($k_{\text{cat}}/K_{\text{M}} = 4.6 \times 10^4 \text{ M}^{-1} \text{ min}^{-1}$), and mispairs with dT and dC are extended as fast and actually 2-fold faster than the parental heteropair d5SICS:dMMO2. Extension of the dT and dC mispairs is efficient; however, because neither is synthesized at a detectable rate, they do not compromise replication. The fidelity of the parental heteropair replication in this strand and sequence context is ~ 7000 .

By comparison with d5SICS:dMMO2, extension of the d5SICS:d5FM heteropair is significantly more efficient ($k_{\text{cat}}/K_{\text{M}} = 4.1 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$); however, the mispairs with d5FM in the template are also generally extended more efficiently than those with dMMO2 in the template (Table 2). Extension of the dC:d5FM and dT:d5FM mispairs is virtually as efficient as that for the correct d5SICS:d5FM heteropair, but as with dMMO2 in the template, neither the dC:d5FM nor the dT:d5FM mispair is efficiently synthesized (see above). However, the dA:d5FM mispair is also extended more efficiently than the dA:dMMO2 mispair, and along with its moderately efficient synthesis (see above), the mispair with dA limits the overall fidelity of the d5SICS:d5FM heteropair in this sequence context to slightly greater than 300. Thus, fluorine substitution significantly reduces the fidelity of heteropair replication in this strand and sequence context.

As with d5FM, the d5SICS:dNaM heteropair is extended more efficiently than the parental d5SICS:dMMO2 heteropair; however, in this case none of the mispairs are extended more efficiently than the corresponding mispairs with dMMO2 in the template (Table 2). While the minimum extension fidelity is only 2, d5SICS:dNaM is the first heteropair that is selectively extended in this strand context relative to all possible mispairs. Moreover, only the mispairs with dA and dNaM are synthesized and extended with any efficiency, and it is these mispairs that ultimately limit the fidelity of replication, but in this case, the minimum fidelity remains high at 3×10^3 . Thus, compared to the parental heteropair, the increased aromatic surface area of dNaM significantly improves the fidelity of dNaM:d5SICS replication without significantly reducing it in the opposite strand context.

2.2. Unnatural Heteropair Replication in Sequence Context II. **2.2.1. Synthesis and Extension with d5SICS in the Template.** To determine whether the improvements in heteropair replication are general, we next fully characterized the different heteropairs and mispairs in sequence context II. For reference, the synthesis and extension of a natural dA:dT base pair in sequence context II is virtually the same as in sequence context I (Table 3). The insertion of dMMO2 opposite d5SICS proceeds with an efficiency of $4.8 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$ (Table 3), and when

Table 3. Second-Order Rate Constants ($k_{\text{cat}}/K_{\text{M}}$ Values) for Heteropair and Mismatch Synthesis and Extension with d5SICS in the Template (Context II)^a

5' - TAATACGACTCACTATAGGGAGA (Ψ)
3' - ATTATGCTGAGTGATATCCCTCT (X) GCTAGGTTACGGCAGGATCGC

X	Y	synthesis ($\text{M}^{-1} \text{ min}^{-1}$)	extension ($\text{M}^{-1} \text{ min}^{-1}$)
T	A	$(3.3 \pm 0.1) \times 10^8$	$(3.1 \pm 0.4) \times 10^8$
5SICS	MMO2	$(4.8 \pm 1.8) \times 10^5$	$(3.2 \pm 0.1) \times 10^5$
	5FM	$(4.2 \pm 0.8) \times 10^6$	$(1.3 \pm 0.6) \times 10^6$
	NaM	$(4.1 \pm 0.1) \times 10^7$	$(6.8 \pm 2.2) \times 10^5$
	5SICS	$(2.4 \pm 0.2) \times 10^5$	$< 1.0 \times 10^3$
	A	$(5.9 \pm 0.8) \times 10^4$	$(3.6 \pm 1.6) \times 10^3$
	G	$(1.9 \pm 0.5) \times 10^5$	$< 1.0 \times 10^3$
	C	$< 1.0 \times 10^3$	$< 1.0 \times 10^3$
	T	$(8.5 \pm 1.1) \times 10^3$	$(3.3 \pm 0.2) \times 10^5$

^a See Experimental Section for details.

compared with the efficiencies of mispair synthesis, this results in a minimum fidelity for correct heteropair synthesis of 2.

As in sequence context I, insertion of d5FMTP and dNaMTP opposite d5SICS in sequence context II is significantly more efficient than insertion of dMMO2TP (Table 3). In fact, for dNaMTP, the relative increase in efficiency is significantly greater in sequence context II than it is in sequence context I. These increases in efficiency result in d5FM:d5SICS and dNaM:d5SICS heteropairs that are synthesized only 79-fold and, remarkably, only 8-fold less efficiently, respectively, than a natural base pair. When compared with the rates of mispair synthesis (Table 3), this results in a minimum synthesis fidelity of 18 and 170, respectively, for d5FM:d5SICS and dNaM:d5SICS heteropairs.

As observed in sequence context I, the extension of dNaM and d5FM paired opposite d5SICS in the template of sequence context II is slightly more efficient than with dMMO2 paired opposite d5SICS (Table 3). While the extension efficiencies of the heteropairs in this strand context are slightly lower in sequence context II than in sequence context I, this is also generally the case for the mispairs. The only mispair with d5SICS in sequence context II that is extended with any efficiency is that with dT ($k_{\text{cat}}/K_{\text{M}} = 3.3 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$). Extension of dT:d5SICS is as efficient as extension of dMMO2:d5SICS, but 2–4-fold less efficient than extension of d5FM:d5SICS or dNaM:d5SICS. Although the synthesis of dT:d5SICS is only marginally efficient (see above), extension of this mispair is sufficiently efficient that it limits the minimum overall replication fidelity of the heteropair. Nonetheless, the overall fidelity of this step is increased from 53 for dMMO2:d5SICS to 1800 for d5FM:d5SICS and, remarkably, to 10 000 for dNaM:d5SICS.

2.2.2. Synthesis and Extension with dMMO2, d5FM, or dNaM in the Template. Synthesis of the parental heteropair by insertion of d5SICSTP opposite dMMO2 proceeds with an efficiency of $6.6 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$ in sequence context II (Table 4). d5SICSTP is also inserted opposite d5FM and dNaM in sequence context II with remarkable efficiency ($8.0 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$ and $5.5 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$, respectively). The dA:dMMO2 mispair and the dMMO2:dMMO2 self-pair are synthesized nearly 20-fold more efficiently than in sequence context I, and these mispairs limit the minimum synthesis fidelity of the parental heteropair to 36 and of d5SICS:d5FM and d5SICS:dNaM heteropairs to 26 and 12, respectively. This slight reduction in synthesis fidelity does not compromise overall

Table 4. Second-Order Rate Constants (k_{cat}/K_M Values) for Heteropair and Mispair Synthesis and Extension with dMMO2 or a Derivative in the Template (Context II)^a

5' - TAATACGACTCACTATAGGGAGA (Y)
3' - ATTATGCTGAGTGATATCCCTCT (X) GCTAGGTTACGGCAGGATCGC

X	Y	synthesis ($\text{M}^{-1} \text{min}^{-1}$)	extension ($\text{M}^{-1} \text{min}^{-1}$)
MMO2	5SICS	$(6.6 \pm 0.5) \times 10^7$	$(1.7 \pm 0.8) \times 10^6$
	MMO2	$(1.8 \pm 1.2) \times 10^6$	$<1.0 \times 10^3$
	A	$(1.7 \pm 0.3) \times 10^6$	$(1.1 \pm 0.3) \times 10^4$
	G	$(7.9 \pm 2.1) \times 10^3$	$<1.0 \times 10^3$
	C	$(3.0 \pm 0.8) \times 10^3$	$(4.4 \pm 1.0) \times 10^5$
NaM	T	$(5.2 \pm 2.7) \times 10^3$	$(2.0 \pm 0.6) \times 10^6$
	5SICS	$(5.5 \pm 0.7) \times 10^7$	$(4.1 \pm 0.1) \times 10^6$
	NaM	$(4.8 \pm 1.2) \times 10^6$	$<1.0 \times 10^3$
	A	$(1.4 \pm 0.4) \times 10^6$	$(1.4 \pm 0.1) \times 10^4$
	G	$(3.7 \pm 0.4) \times 10^3$	$<1.0 \times 10^3$
5FM	C	$(6.4 \pm 2.6) \times 10^3$	$(3.0 \pm 0.7) \times 10^5$
	T	$(8.3 \pm 1.4) \times 10^3$	$(8.3 \pm 0.6) \times 10^5$
	5SICS	$(8.0 \pm 0.8) \times 10^7$	$(8.2 \pm 0.3) \times 10^6$
	5FM	$(2.7 \pm 0.6) \times 10^6$	$<1.0 \times 10^3$
	A	$(3.1 \pm 0.1) \times 10^6$	$(4.4 \pm 0.1) \times 10^4$
	G	$(1.2 \pm 0.4) \times 10^3$	$<1.0 \times 10^3$
	C	$(7.5 \pm 0.6) \times 10^3$	$(1.5 \pm 0.4) \times 10^6$
T	$(1.2 \pm 0.6) \times 10^4$	$(4.1 \pm 1.0) \times 10^6$	

^a See Experimental Section for details.

fidelity, as in each case the most efficiently synthesized mispairs are extended much less efficiently than the correct heteropairs (see below).

Once synthesized, d5SICS:dMMO2, d5SICS:d5FM, and d5SICS:dNaM heteropairs are all extended with high efficiency (Table 4), but as with each of the mispairs, these efficiencies are somewhat lower than those observed in sequence context I. The mispairs with dT are consistently the most efficiently extended, but because they are not efficiently synthesized, they do not compromise the minimum overall fidelity. For each heteropair in this strand and sequence context, the minimum overall fidelity is limited by mispairing with dA. The minimum overall fidelity of the dMMO2- and d5FM-based heteropairs is 6100 and 4700, respectively, while the dNaM-based heteropair has a remarkable minimum overall fidelity of 12100.

2.3. Synthesis of dNaM:d5SICS in Sequence Contexts III and IV. Virtually without exception, both unnatural and natural triphosphates are inserted opposite unnatural nucleobases in the template more efficiently in sequence context II than in context I. The difference is most pronounced for insertion of dNaMTP, which is almost 10-fold faster in sequence context II than in context I. To further probe this sequence dependence, we examined two additional sequence contexts, III and IV, which differ from sequence context I only at the position immediately 3' to the unnatural nucleotide in the template. In the four sequence contexts I–IV, the natural base pair in the primer–template preceding the unnatural nucleotide is dA:dT, dC:dG, dT:dA, and dG:dC, respectively. Insertion of dNaMTP opposite d5SICS proceeds with an efficiency of $5.9 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$ in sequence context III and $6.8 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$ in sequence context IV. Thus, the rate of unnatural base pair synthesis is nearly identical in contexts I and III, as well as in contexts II and IV, suggesting that the identity, but not the strand context, of the flanking natural base pair is important for unnatural triphosphate insertion.

3. Discussion

The unnatural base pair formed between d5SICS and dMMO2 was identified by optimizing the most promising lead

from a screen of 3600 candidates. While the heteropair formed between dMMO2 and d5SICS is replicated better than the unnatural base pairs we have examined previously, its replication still requires further optimization to approach that of a natural base pair. Toward this goal we synthesized and characterized the dMMO2 analogues d5FM and dNaM. These derivatives differ from dMMO2 in nucleobase dipole moment, shape, aromatic surface area, and hydrophobicity, which have all been shown to significantly impact unnatural base pair replication.^{7,13,24,41,45} Modifications are made at the positions meta and para to the glycosidic linkage, which have been shown to be tolerant of modification.¹⁰

Unnatural Heteropair Replication and Its Sequence Dependence. Insertion of dMMO2 opposite d5SICS limits the synthesis of the parental heteropair in both sequence contexts characterized. We find that the efficiency of this step of replication is significantly increased both by the fluorine substituent of d5FM (by ~10-fold) and by the increased aromatic surface area of dNaM (by 14–85-fold, depending on sequence context). In the opposite strand context, dMMO2, d5FM, and dNaM all template the insertion of d5SICSTP with similarly high efficiency. Both the fluorine substituent and increased aromatic surface area also generally increase the rate of natural dNTP insertion, with the exception of dGTP in context II, where the rate of mispair synthesis is decreased 6- and 2-fold, respectively. Both unnatural pair and mispair synthesis were consistently more efficient in sequence context II than context I. In addition, the efficiency of dNaMTP insertion opposite d5SICS is the same in sequence contexts I and III as well as in contexts II and IV. These data suggest that the observed sequence dependencies reflect the nature of the natural base pair that precedes the unnatural base pair being synthesized, with a dG:dC or a dC:dG resulting in similar and relatively more efficient insertion than a dA:dT or dT:dA, which also resulted in similar efficiencies. Thus, we conclude that the identity of the flanking base pair is important, while its strand context is not.

d5FM:d5SICS and d5SICS:d5FM heteropairs are extended 3–6-fold more efficiently than the parental heteropair between dMMO2 and d5SICS, while the dNaM:d5SICS heteropair is extended with an efficiency similar to the parental heteropair in the same sequence context, and the d5SICS:dNaM heteropair is extended 2–4-fold more efficiently. Relative to dMMO2, the mispairs with a natural nucleotide in the primer paired opposite d5FM are also generally extended slightly faster, while those with dNaM are not. As with heteropair and mispair synthesis, we also observe a sequence dependence on the rate of mispair extension, which is generally more efficient in sequence context I than in context II. The only exceptions are with dMMO2, d5FM, or dNaM in the template, where the pairs with d5SICS or dT are extended slightly more efficiently in context I. While extension in the two sequence contexts involves the insertion of a different natural triphosphate, dCTP in context I and dATP in context II, this difference is unlikely to be the origin of the observed differences in extension efficiency, as the rate of natural base pair extension is identical in the two sequence contexts.

Model for Unnatural Base Pair Replication. Previously, we proposed a model of unnatural base pair extension wherein pairs formed between analogues with large aromatic nucleobases are difficult to extend due to an intercalative mode of pairing that results in distortions of the primer terminus.⁹ We thus attribute the generally less efficient heteropair and mispair extension in

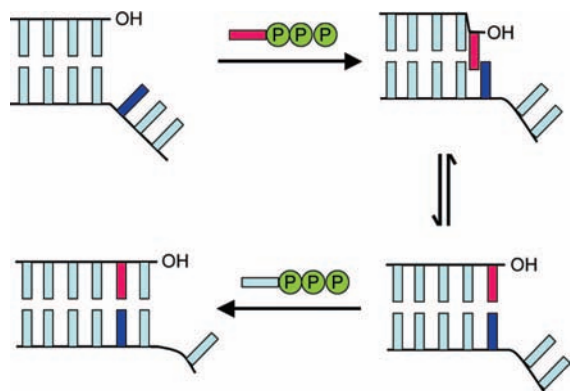


Figure 2. Model for unnatural base pair replication. The unnatural nucleobase in the template is shown in dark blue and the natural or unnatural nucleobase of the incoming dNTP is shown in red.

sequence context II to increased intercalation of the primer terminus nucleobase in the presence of a flanking dC:dG pair, which stacks more favorably in DNA than a dA:dT base pair.^{47–49} This effect is more pronounced with d5SICS in the template, due to its aromatic surface area which is suitably disposed to pack with the nucleobase of the incoming dNTP. The effect is less pronounced with dMMO2 and d5FM in the template, due to reduced surface area, with dNaM in the template, due to geometrical constraints that preclude packing, and with dT in the primer, possibly due to a reduced ability to intercalate. Thus, in general the primer terminus may be sufficiently flexible to allow some distortion, provided that templating base is incapable of mediating edge-on interactions and sufficiently stabilizing packing interactions are available via intercalation.

Unlike for extension, data collected previously have not led to the development of a model for unnatural base pair synthesis, but the systematic characterization of unnatural pair and mispair synthesis reported here provides more insight. The observed sequence dependencies must result from the identity of the natural nucleotide in the template 3' to the unnatural nucleotide, and possibly with its pairing nucleotide in the primer (the 5' nucleotide is isolated by a sharp turn in the template and packing interactions with the polymerase). Thus, we infer that the 3' dG in the template, and/or its pairing dC in the primer of sequence context II, favors natural and unnatural dNTP insertion opposite an unnatural nucleotide, relative to the dA:dT base pair of sequence context I. The rates of dNaM:d5SICS synthesis in sequence contexts III and IV support this conclusion and further suggest that the rates depend on the flanking *base pair* but not on its strand context (i.e., which nucleotide is in the primer strand and which is in the template strand).

The data prompt us to conclude that the intercalative model of base pair extension is also applicable to unnatural base pair

and mispair synthesis (Figure 2). Indeed, at least with Kf, an intercalative mode of interaction, both during and after synthesis, may dominate when the templating nucleobase is incapable of edge-on interactions but does allow for stable intercalation. This model suggests that efficient replication of predominantly hydrophobic unnatural base pairs requires nucleobase intercalation that is sufficient to facilitate synthesis but not so stabilizing as to inhibit deintercalation and continued extension. If further experiments support this model, it should prove invaluable for continued optimization of the unnatural base pair and perhaps for the design of additional unnatural base pair candidates as well.

Efforts to Expand the Genetic Alphabet. From a practical perspective, there are two important and related criteria for the evaluation of any unnatural base pair's replication: the overall efficiency and fidelity of its replication. The overall efficiency with which the parental heteropair formed between d5SICS and dMMO2 is replicated is limited by the relatively inefficient synthesis of dMMO2:d5SICS and, to a lesser degree, by the extension of d5SICS:dMMO2 (sequence context I) or by the synthesis and extension of dMMO2:d5SICS (sequence context II). The fidelity of heteropair synthesis is most limited by the self-pairs or mispairs with purines, while the fidelity of extension is limited by the mispairs with pyrimidines (Table 5). Despite their inefficient synthesis, the mispairs with dT most limit overall replication fidelity with d5SICS in the template, and despite their relatively inefficient extension, the mispairs with dA most limit the overall fidelity with dMMO2 in the template. Overall, these mispairs limit the fidelity of d5SICS:dMMO2 to between 6000 and 7000 and that of dMMO2:d5SICS to 53 and 130 in sequence contexts I and II, respectively (Table 5).

Derivatization of the dMMO2 scaffold with a meta fluorine atom or meta/para-linked aromatic surface area increases by 1–2 orders of magnitude the efficiency of the step that most limits replication: insertion of the triphosphate opposite d5SICS in the template, due to both an increased k_{cat} and a decreased K_{M} . Along with smaller effects on the other steps of replication, this results in overall fidelities that are, in general, significantly increased relative to the parental heteropair. With d5FM in the primer strand and d5SICS in the template strand, the fidelity increases to almost 4000-fold in context I and only slightly less in context II. However, while the overall fidelity remains high with d5SICS in the primer and d5FM in the template in sequence context II (5000-fold), it is lower in context I (300-fold). The moderate fidelity in context I results from the relatively efficient insertion and extension of the dA:d5FM mispair.

Both the efficiency and overall fidelity of the heteropair formed between dNaM and d5SICS are remarkable in all strand and sequence contexts examined; the efficiency of every step of replication is within 8–140-fold and 6–490-fold of a natural

Table 5. Minimum Single Step and Overall Replication Fidelities^a

primer	template	minimum synthesis fidelity		minimum extension fidelity		minimum replication fidelity ^b	
		context I	context II	context I	context II	context I	context II
d5SICS	dMMO2	390 (dMMO2, dA)	36 (dMMO2, dA)	0.6 (dC)	0.9 (dT)	6730 (dA)	6100 (dA)
dMMO2	d5SICS	2.8 (dG)	2.0 (d5SICS, dG)	4.8 (dT)	0.9 (dT)	130 (dT)	53 (dT)
d5SICS	d5FM	16 (d5FM)	26 (dA, d5FM)	1.2 (dT)	2.0 (dT)	330 (dA)	4700 (dA)
d5FM	d5SICS	28 (dG)	18 (d5SICS)	14 (dT)	3.9 (dT)	3800 (dT)	1800 (dT)
d5SICS	dNaM	11 (dNaM)	12 (dNaM)	2.7 (dC)	4.9 (dT)	2670 (dNaM)	12 100 (dA)
dNaM	d5SICS	38 (dG)	170 (dNaM)	3.0 (dT)	2.1 (dT)	1260 (dT)	10 000 (dT)

^a See text for details. The mispairs that most limit fidelity are indicated in parentheses. ^b Minimum replication fidelity is the product of synthesis and extension fidelity.

base pair in the two sequence contexts examined, and the overall fidelities range from 10^3 to 10^4 . The fidelities are by far the highest reported for any unnatural base pair and result, at least in part, from the consistent orthogonality of the mispairs that are most efficiently synthesized and those that are most efficiently extended. The high efficiencies and fidelities associated with the heteropair formed between dNaM and d5SICS appear to be generally independent of sequence, are beginning to rival those of natural synthesis, and are likely to be sufficient for most in vitro applications of an expanded genetic alphabet. Given these remarkable properties, we also believe that the heteropair is sufficiently promising to initiate in vivo efforts as part of the long-term goal to expand the genetic code and create a semisynthetic organism. Such efforts are now underway.

4. Experimental Section

General Methods. All reactions were carried out in oven-dried glassware under inert atmosphere and all solvents were dried over 4 Å molecular sieves with the exceptions of dichloromethane, which was distilled from CaH_2 , and tetrahydrofuran, which was distilled from sodium and potassium metal. All other reagents were purchased from Aldrich. ^1H , ^{13}C , and ^{31}P NMR spectra were recorded on Varian Mercury 300, Varian Inova-400, or Bruker AMX-400 spectrometers. High-resolution mass spectroscopic data were obtained from the facilities at The Scripps Research Institute. Polynucleotide kinase was purchased from New England Biolabs, Kf from GE Healthcare, and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ from MP Biomedicals.

Nucleoside Synthesis. The synthesis of d5SICS was described previously¹⁰ and the synthesis of d5FM and dNaM are described in detail in the Supporting Information. Briefly, for dNaM, 2-methoxynaphthalene was lithiated and coupled to 2-deoxy-3,5-*O*-(tetraisopropylidisiloxane-1,3-diyl)-D-erythropentofranose. The resulting diol was cyclized under Mitsunobu conditions and deprotection yielded the free nucleoside. For d5FM, 4-fluoro-3-methylanisole was iodinated with AgNO_3 and I_2 and then coupled to *tert*-butyl-[[3-(*tert*-butyldimethylsilyloxy)-2,3-dihydrofuran-2-yl]methoxy]dimethylsilane. In each case anomeric mixtures of nucleosides were obtained and the α anomer was purified by column chromatography and confirmed by NOE and HOMO-COSY NMR spectroscopy (Supporting Information).

Oligonucleotide Synthesis. Oligonucleotides were prepared by the β -cyanoethyl phosphoramidite method on controlled pore glass supports (1 μmol) by use of an Applied Biosystems Inc. 392 DNA/RNA synthesizer. After automated synthesis, the oligonucleotides were cleaved from the support and deprotected by heating in aqueous ammonia solution at 55 °C for 12 h. The crude product was further purified by polyacrylamide gel electrophoresis, followed by electroelution. The resulting purified oligonucleotides were precipitated in 80% ethanol and dried overnight. Oligonucleotides were characterized by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (Applied Biosystems Voyager DE-PRO System 6008), and their concentrations were determined spectrophotometrically by use of standard extinction coefficients for the natural nucleotides and extinction coefficients at 260 nm of 1.9×10^{-2} , 1.5×10^{-3} , and $4.1 \times 10^{-3} \text{ M}^{-1} \text{ m}^{-1}$ at for d5SICS, d5FM, and dNaM, respectively.

Gel-Based Kinetic Assay. Primer oligonucleotides were 5'-radiolabeled with $[\text{}^{32}\text{P}]\text{ATP}$ and T4 polynucleotide kinase. Template and primers were annealed in Kf reaction buffer by heating to 90 °C followed by slow cooling to ambient temperature. Assay conditions included 40 nM primer/template, 0.1–1.2 nM Kf, 50 mM Tris-HCl, pH 7.5, 10 mM MgCl_2 , 1 mM dithiothreitol (DTT), and 50 $\mu\text{g}/\text{mL}$ acetylated bovine serum albumin (BSA). The reactions were carried out under steady-state conditions by combining the DNA–enzyme mixture with an equal volume (5 μL) of $2 \times$ dNTP stock solution, with varying concentrations of unnatural triphosphate (1–2000 μM), incubating at 25 °C for 1–10 min, and quenching by the addition of 20 μL of loading dye [95% formamide, 20 mM ethylenediaminetetraacetic acid (EDTA), and sufficient amounts of bromophenol blue and xylene cyanol]. The reaction mixture was then analyzed by 15% polyacrylamide and 8 M urea denaturing gel electrophoresis. Radioactivity was quantified by use of a Phosphorimager and the ImageQuant program (Molecular Dynamics) with overnight exposures. k_{obs} values were plotted against triphosphate concentration and used to fit the Michaelis–Menten equation (Kaleidagraph, Synergy Software). The data presented are averages and standard deviations of three independent determinations.

Acknowledgment. Funding was provided by the National Institutes of Health (GM060005 to F.E.R.) and Korea Research Foundation (KRF-2006-352-C00047 to Y.J.S.).

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

JA807853M

(47) Danilov, V. I. *Mol. Biol. Rep.* **1975**, *2*, 263–266.

(48) Johnson, T.; Zhu, J.; Wartell, R. M. *Biochemistry* **1998**, *37*, 12343–12350.

(49) Bommarito, S.; Peyret, N.; SantaLucia, J., Jr. *Nucleic Acids Res.* **2000**, *28*, 1929–1934.