

# PROBING DNA STABILITY AND REPLICATION WITH HYDROPHOBIC BASES

Reported by Matthew Stone

October 1, 2001

## INTRODUCTION

Recent explorations employing artificial bases have allowed systematic examinations of base stacking in DNA duplexes and insight into the high rate of accuracy exhibited during DNA replication. Initial investigations of artificial hydrophobic bases by Kool and coworkers were pursued to evaluate the stability of duplex DNA in the absence of hydrogen bonding. Similar research on the contribution of hydrogen bonding had been performed in the past; however, the common strategy was to alter the bases by blocking or deleting hydrogen bonding functionalities.<sup>1</sup> These modifications often distorted the steric size and shape of the bases, which in turn influenced the stability of the DNA structure. In contrast to these studies, Kool and coworkers designed nucleotide bases that were steric mimics of natural bases, but lacked the ability to hydrogen bond. In the absence of hydrogen bonding they found evidence that the base stacking had a significant stabilizing effect on the structure of duplex DNA. Schultz and coworkers subsequently demonstrated that strongly stacking bases could be used to construct DNA duplexes with greater stabilities than duplexes composed exclusively of native bases.

Investigations have found new evidence that the high rate of accuracy at which DNA is replicated can be attributed to geometric constraints at the DNA polymerase active site. Bases that mimic a normal base pairing in a steric sense, but contain no complementary hydrogen bonds were found to be inserted at rates approaching those of natural bases. Unlike site directed mutagenesis, the use of artificial bases will not alter the structure of the active site. Systematic examination of bases by altering their minor groove functionality has suggested that hydrogen bond interactions may exist between residues in the polymerase active site and nucleotides. These interactions would assist in orientating bases in the active site and in detecting incorrect base pairings. Also some hydrophobic bases closely mimic a damaged DNA site. Resulting repair mechanisms seem to operate in much the same fashion as is observed for damaged natural bases. This could provide a model to study DNA repair mechanisms and also an opportunity explore how hydrophobic base might be used to encode for unnatural genetic data.

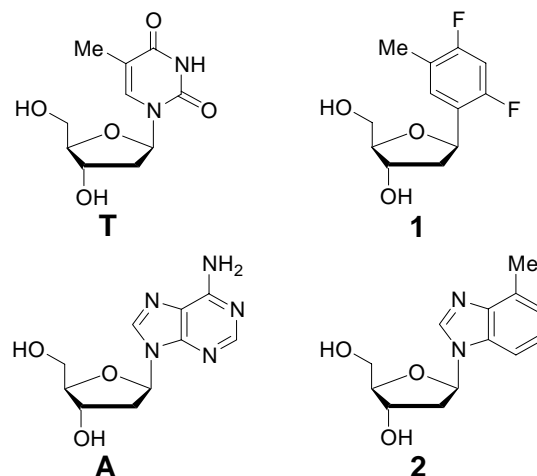
The substitution of hydrophobic bases into a DNA sequence to better understand the contributions that base stacking and hydrophobic interactions have on the stability of duplexes will be the focus of the first portion of this review. The use of hydrophobic bases to explore how polymerases

process nucleotides and their contribution to the fidelity of DNA replication will be discussed in the second portion.

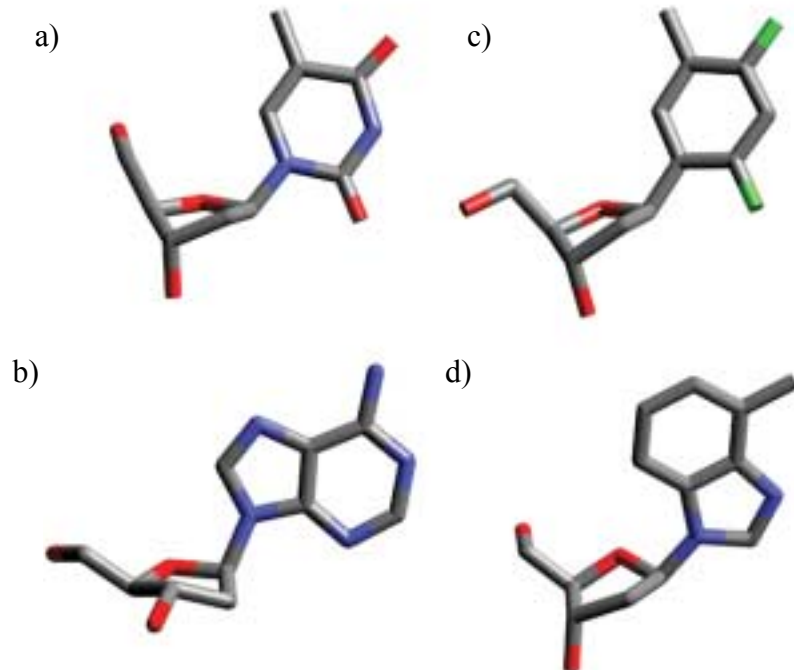
## EVALUATING DNA STABILITY IN THE ABSENCE OF HYDROGEN BONDING

### Internal Substitution of Non-Hydrogen Bonding Isosteres of Thymine and Adenine

There are many forces that contribute to the stability of DNA including hydrogen bonding, base stacking, steric effects, electrostatic repulsion and hydrophobic interactions. Isolating the individual contribution of these forces on DNA stability is difficult. Examination of artificial bases often has the problem of isolating the modifications from possible steric interactions. To eliminate possible steric contributions bases were designed that possessed nearly identical molecular footprints, but lacked the ability to hydrogen bond (Figure 1).<sup>2</sup> The structural mimics or isosteres were then used to examine base stacking in the absence of hydrogen bonding.



**Figure 1.** Deoxythymidine (T) and Deoxyadenosine (A) and their steric mimics (1 and 2).



**Figure 2.** X-ray crystal structures of deoxythymidine (a),<sup>5</sup> deoxyadenosine (b)<sup>4</sup> and their steric mimics 1 (c)<sup>3</sup> and 2 (d).<sup>6</sup>

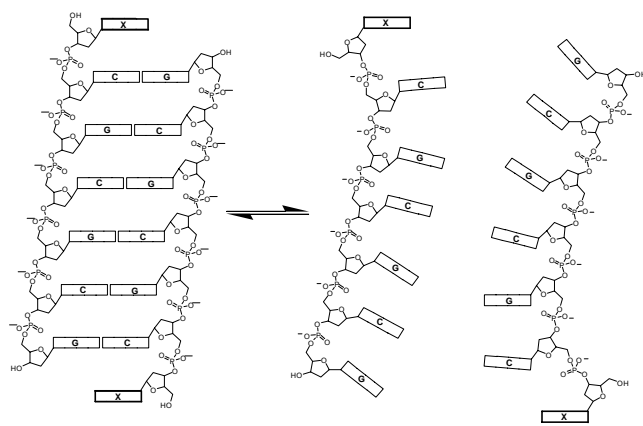
To detect conformational differences between nucleoside **1** and deoxyadenosine **T**, an X-ray crystal structure of **1** was obtained. It was found that **1** had a conformation nearly identical to a previously reported crystal structure of **T** (Figure 2).<sup>3</sup> The crystal structure of nucleoside **2** showed a syn base conformation base rather than the anti base conformation of **A** in the solid state. However subsequent NOE studies of **2** and **A** found that they had similar conformations in solution.<sup>4</sup> The only significant difference between **2** and **A** was the bond to the methyl group on **2**,

which was 0.18 Å longer than the corresponding bond to the amine on **3**. It has been established that **1** and **2** will function the same as **A** and **T** sterically but lack the ability to hydrogen bond.

To determine the effect on the stability of a DNA duplex, analogs **1** and **2** were internally substituted for natural bases. The stability of the resulting complex was determined monitoring the absorption of the solution at 260 nm or 280 nm while raising the temperature. The resulting absorbance curve was fitting to a two state approximation to determine the free energy of the duplex at room temperature.

Insertion of the bases **1** and **2** opposite each other destabilized the complex by 3.5 kcal/mol. This loss of energy was very close to a natural base mismatch within the same sequence. In this preliminary study it was not clear that the lower stability of the complex could be attributed solely the absence of hydrogen bonds. The most consistent finding of the study was that hydrophobic bases **1** and **2** were more stable when paired with themselves than with the four naturally occurring bases (Table 1). The preference for self-pairing was attributed to the cost of desolvating natural bases, which have the ability to hydrogen bond outside the duplex.

In contrast to internal substitutions, hydrophobic base pairs at the end of the duplex were found to increase stability. A **1-1** base pairing at the ends of a duplex was found to be 2.0 kcal/mol more stable than a corresponding **A-T** base pairing.<sup>7</sup> The additional stability was attributed to the increased hydrophobic and base stacking ability of **1**. As opposed to bases located in the middle of the duplex, bases at the end are exposed to water. This results in relatively



**Figure 3. Thermal decomplexation of DNA strands.**

**Table 1.  $T_m$  of duplexes with substituted **2** opposite natural nucleotides and **1**.<sup>3</sup>**

duplex	$T_m$ (°C)	$\Delta G_{25}^\circ$ (kcal/mol)
5'-CTTTTC $\overline{\text{A}}$ TTCTT 3'-GAAAAG $\overline{\text{T}}$ AAGAA	39.8	-12.4
5'-CTTTTC $\overline{\text{2}}$ TTCTT 3'-GAAAAG $\overline{\text{1}}$ AAGAA	30.1	-8.9
5'-CTTTTC $\overline{\text{2}}$ TTCTT 3'-GAAAAG $\overline{\text{A}}$ AAGAA	25.3	-8.0
5'-CTTTTC $\overline{\text{2}}$ TTCTT 3'-GAAAAG $\overline{\text{C}}$ AAGAA	24.4	-7.9
5'-CTTTTC $\overline{\text{2}}$ TTCTT 3'-GAAAAG $\overline{\text{G}}$ AAGAA	23.8	-7.6
5'-CTTTTC $\overline{\text{2}}$ TTCTT 3'-GAAAAG $\overline{\text{T}}$ AAGAA	20.8	-7.5

weaker hydrogen bonding interactions and the stability of the complex is more dependent on their stacking ability.

### External Substitution of Hydrophobic Bases to Examine Base Stacking

The relationship between stronger base stacking at the end of the duplex and a more stable complex was explored further. Self-complementary 5'd(XCGCGCG) strands of DNA were produced and the X base was varied (Figure 3). A study was

undertaken to determine what properties contribute to efficient base stacking.<sup>8</sup> Upon duplexation of two strands, the X bases were left unpaired or “dangling” off the ends of the duplex. The melting point of the duplex was then used as an indication of the stacking ability of the X base. Correlations were found with the polarizability and surface area excluded of the X base with the  $\Delta\Delta G^\circ$  of stacking. In contrast, no clear correlations were found between the  $\Delta\Delta G^\circ$  of stacking and the dipole moment of the base. Knowing that stacking abilities of bases can impart stability to the duplex, stable internal substitutions should be possible if a strongly stacking base was used. The strong propensity for base stacking would compensate for the lack of hydrogen bonds.

### Internal Substitution of Non-Isosteric Hydrophobic Base Pairs

Stabilizing internal substitution of hydrophobic bases using efficient base stacking offers the ability to incorporate fluorescent probes, metal binding ligands, and structures useful for in vitro experiments.<sup>9</sup> There also have been investigations into hydrophobic bases as universal or non-specific sites at which any of the four natural bases could pair with equal stability; however, these bases were found to be much less stable than natural pairings.<sup>10,11</sup> The substitution of pyrene base **3** opposite an abasic (lacking a base) site **4** demonstrated that a stable internal hydrophobic base substitution could occur (Figure 4).<sup>12</sup> The stability of pyrene base **3** was only slightly less than a natural base pairing. This was attributed to its strong stacking ability and its relatively large aromatic surface. Base **3** was also of interest and was studied as a fluorescent probe in other experiments.<sup>13</sup>

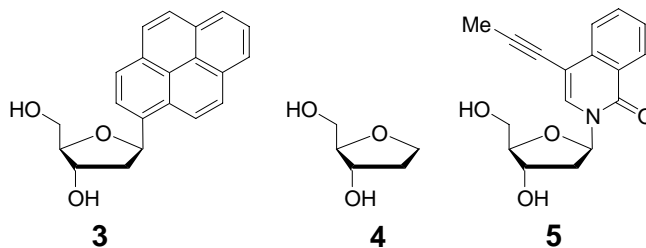


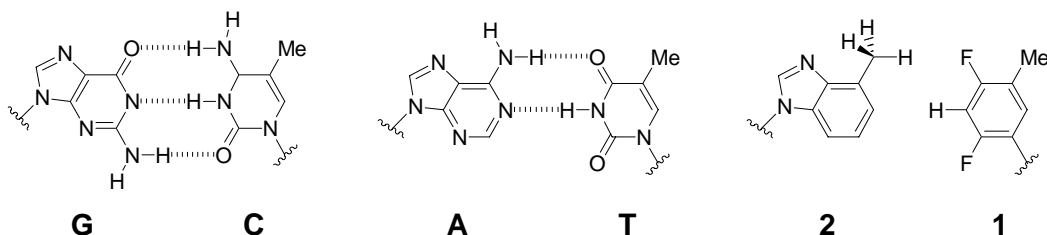
Figure 4. Hydrophobic nucleoside bases pairs that form stable duplexes even when they are internally substituted.

Motivated by the idea that a third stable base pair in DNA could be used to store additional genetic information, Shultz and coworkers subsequently demonstrated that base **5** could be used to construct duplexes possessing greater stabilities than duplexes containing native bases alone.<sup>9</sup> Similar to previously studied hydrophobic bases, **5** displayed a strong preference for self-pairing relative to the four natural base pairs. However, base **5** was able to form duplexes that were more stable than an A-T or C-G base pair would be at the same position. Many other structures have since been investigated, and many also had similar stabilizing capabilities.<sup>14</sup> The ability to form stable DNA duplexes with a third set of base pairs would allow genetic information to be stored. However, in order for the information to be used, these same stable nucleotides must be able to be processed by polymerases to replicate and translate the sequence. Besides the pursuit to expand the genetic alphabet there has been a great deal learned about how polymerases influence DNA sequence fidelity during replication.

## REPLICATION OF HYDROPHOBIC BASES

### Examining Geometric Constrains as a Mechanism for Polymerase Selectivity

In the replication of a bacterial DNA sequence there will be only one mismatch in the replication of  $10^9$  nucleotides.<sup>15</sup> As there are relatively small differences in energy between incorrectly hydrogen bonded base pairings and correct Watson and Crick base pairings, the level of accuracy displayed in replication is much greater than would be expected by comparing differences in energy (Figure 5). The accuracy of replication must, therefore, rely greatly on the ability of polymerases to regulate the insertion of the correct bases. The study of hydrophobic bases in DNA has provided strong evidence that supports geometric constrains at the polymerase active site greatly increase the level of accuracy in replication.<sup>16</sup> These constraints strongly favor Watson-Crick base pairings and disfavor other potential pairings. Isosteric bases fill the same space within the polymerase active site, and therefore even without hydrogen bonding should mimic a correct Watson-Crick base pairing. Processing of isosteric bases by the Klenow Fragment of *E. coli* polymerase I (KF) has been found to approach rates observed for natural bases.<sup>17</sup>

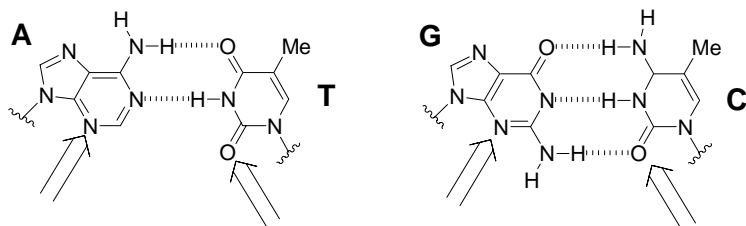


**Figure 5. Watson and Crick base pairing of guanine (G) with cytosine (C) and adenine (A) with thymine (T) along with steric mimics 1 and 2.**

Determining the rate of polymerase extension of a DNA strand was determined by monitoring the reaction at different times and separating the differing lengths of DNA using gel electrophoresis. In this way, non-hydrogen bonding base **1** was shown to template for adenine at selectivities approaching those of a natural base pairing.<sup>18</sup> The space filled by base analogs **1** and **2** have been shown to be nearly identical to those of adenine and thymine. When inserted into a strand of DNA, base analog **2** also exhibited selectively in templating dTTP (**1** triphosphate) 10-1000 fold over the four natural nucleoside triphosphates.<sup>19</sup> In addition, **1** templated dTTP by a 130-1900 fold preference over the four natural nucleoside triphosphates. The selectivity that hydrophobic bases show against natural bases can be attributed to the desolvation of the natural bases which loose hydrogen bonding when forced into polymerase active site. These results strongly support geometric constraints in a polymerase binding site as a contributing factor to the fidelity of DNA replication.

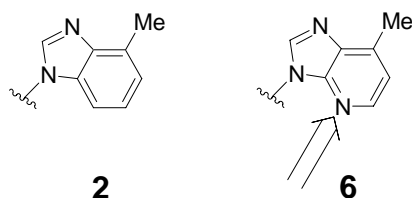
## Examining Minor Groove Hydrogen Bonding That Reinforces Polymerase Selectivity

The formation of hydrogen bonds within the active site of polymerases has also been proposed as an additional method to reinforce steric constraints. A hydrogen bonding interaction in the minor groove would restrict the base from forming non-Watson-Crick base hydrogen bond patterns. X-ray crystal structures have revealed hydrogen bond donating side chains within the active sites of DNA polymerases.<sup>20</sup> These side chain donors have been proposed to interact with the minor groove hydrogen bond acceptors on DNA bases (Figure 6). As a method to detect minor groove hydrogen bonding interactions, a structural modification was made to base analog **2**, giving it a minor groove hydrogen bond acceptor (Figure 7).<sup>21</sup> Bases **2** and **6** were then processed by Klenow Fragment (KF<sup>-</sup>) of *E. coli* polymerase I (lacking 3' exo nuclease activity).



**Figure 6.** Minor groove hydrogen bond acceptors indicated for guanine (G), cytosine (C), adenine (A) and thymine (T).

a)



b)

5'-TCGCAGCCGTCCAX  
3'-AGCGTCGGCAGGT 1CGTAAA

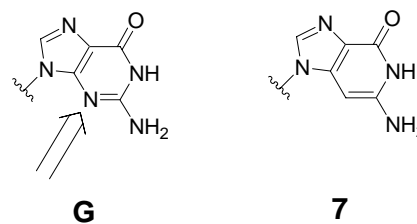
**Figure 7.** a) Hydrophobic adenine nucleosides analogs **2** and **6** with and without minor groove acceptor respectively. b) 5' Primer and 3' template strands where X is **2** or **6**.

If there were hydrogen bonding interactions between the side chain in active site and the minor groove of the nucleotide, the loss in reactivity upon the deletion of the acceptor or donor functionalities should be the same as the deletion of both them at the same time. This experiment employed guanine and a minor groove deleted base analog 3-deazaguanine **7** in position Y (Figure 8). Two amino acids were separately mutated, and a similar loss of reactivity for

6

KF<sup>-</sup> showed a significant difference in its ability to extend the primer strand, depending on the presence of base **2** or **6** in position X (Figure 7). When **6** was placed at position X, the primer strand was extended 300-fold more efficiently than when **2** was at the same position. This would suggest that there was a hydrogen bonding interaction between the minor groove of that base and the active site of KF<sup>-</sup>. Because other explanations for this result were possible, a similar study was conducted by Spratt, which examined mutated potential hydrogen donating side chains in the active site of KF<sup>-</sup>.<sup>22</sup>

a)



b)

5'-GCACCGCAGCGCAY  
3'-CGTGGCGTCGCGTCGTCAGCGTC

**Figure 8.** a) Guanine (G) and minor groove deleted 3-deazaguanine (**7**). b) 5' Primer and 3' template strands where Y is G or **7**.

both guanine and 3-deazaguanine **7** were observed. A similar reduction of reactivity was observed for the unmutated  $\text{KF}^-$  when 3-deazaguanine was at position Y. Although not conclusive, these results strongly suggested that the presence of an important hydrogen bonding interaction between the active site of  $\text{KF}^-$  and the minor groove of bases. This interaction may help to eliminate base pairing other than the Watson-Crick type pairing, thus increasing selectivity.

### Examining DNA Repair Mechanism

There are rate-limiting steps in the process of base insertion by polymerases. A mispaired base can increase the amount of time that is required to move forward to the next step. During this time, polymerases have the opportunity to dissociate, thus terminating the elongation of the strand. A similar phenomenon to base mismatching was observed after the insertion of pyrene nucleoside triphosphate **8**, which was selectively inserted across from an abasic site **4**.<sup>23</sup> Although the polymerase was able to insert **8**, it was not capable of extending the strand past that nucleotide (Figure 9), most likely because its non-Watson and Crick shape was detected as a base pair mismatch. This behavior is not isolated to **8**, but was exhibited by a variety of non-isoteric hydrophobic bases, including self-pairing base **9**.<sup>9,10</sup>

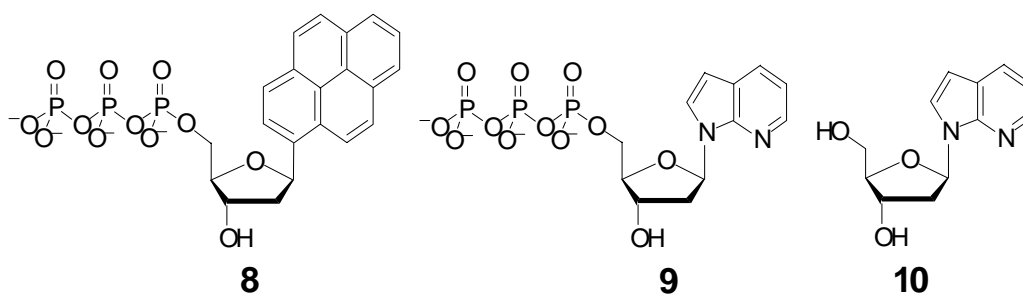


Figure 9. Pyrene nucleoside triphosphate (**8**), 7-aza-indole nucleoside triphosphate (**9**) and 7-aza-indole nucleoside (**10**).

Pursuing the next step towards expanding the genetic alphabet, Schultz and coworkers sought to efficiently replicate an entire strand of DNA containing their non-isoteric, but stable hydrophobic bases. They investigated calf thymus DNA polymerase  $\beta$  ( $\text{Pol } \beta$ ), which had been found to be capable of extending DNA past a damaged site that stalled other polymerases.<sup>24</sup> It was found that  $\text{Pol } \beta$  was able to continue replication of the DNA sequence past 7-aza-indole nucleoside **10**, but only in the presence of  $\text{KF } \text{exo}^+$ .<sup>25</sup> In the absence of  $\text{KF } \text{exo}^+$ , no insertions were observed. Their results were consistent with  $\text{KF } \text{exo}^+$  selectively using **9** to template for **10** after this the rest of the sequence was extended by  $\text{Pol } \beta$ . It remains to be seen if this strategy is universal and would allow efficient replication of other hydrophobic nucleotide bases. If this is the case, then the possibility cannot be excluded at this stage that hydrophobic base pairs in DNA sequences may be capable of processing unnatural amino acids.

## CONCLUSION

The substitution of hydrophobic nucleotides into natural nucleotide sequences has offered insight into the stability of DNA duplexes and shown further evidence that steric, solvophobic, and aromatic stacking are important to duplex stability. Steric and hydrophobic interactions were also found to be crucial in the accurate replication of a DNA sequence. Hydrophobic bases have been shown to offer a valuable tool to systematically investigate DNA stability and replication.

## REFERENCES

- (1) Stawzewski, P.; Tamm, C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 36-57.
- (2) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew., Chem. Int. Edit. Engl.* **2000**, *39*, 991-1009.
- (3) Guckian, K. M.; Morales, J. C.; Kool, E. T. *J. Org. Chem.* **1998**, *63*, 9652-9656.
- (4) Young, D. W.; Tollin, P.; Wilson, H. R. *Acta. Crystallogr. Sect. B* **1969**, *25*, 1423.
- (5) Sato, T. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1984**, *40*, 880.
- (6) Guckian, K. M.; Kool, E. T. *Angew. Chem., Int. Edit. Engl.* **1997**, *36*, 2825-2828.
- (7) Schweitzer, B. A.; Kool, E. T. *J. Am. Chem. Soc.* **1995**, *117*, 1863-1872.
- (8) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 2213-2222.
- (9) Ogawa, A. K.; Wu, Y. Q.; McMinn, D. L.; Liu, J. Q.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 3274-3287.
- (10) Nichols, R.; Andrews, P. C.; Zhang, P.; Bergstrom, D. E. *Nature* **1994**, *369*, 492-493.
- (11) Loakes, D.; Hill, F.; Brown, D. M.; Sailisbury, S. A. *J. Mol. Biol.* **1997**, *270*, 426-435.
- (12) Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191-6192.
- (13) Wu, Y. Q.; Ogawa, A. K.; Berger, M.; McMinn, D. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 7621-7632.
- (14) Echols, H. *Annu. Rev. Biochem.* **1991**, *60*, 477-511.
- (15) Kool, E. T. *Biopolymers* **1998**, *48*, 3-17.
- (16) Moran, S.; Ren, R. X. F.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10506-10511.
- (17) Barsky, D.; Kool, E. T.; Colvin, M. E. *J. Biomol. Struct. Dyn.* **1999**, *16*, 1119-1134.
- (18) Morales, J. C.; Kool, E. T. *Biochemistry* **2000**, *39*, 12979-12988.
- (19) Spratt, T. E. *Biochemistry* **2001**, *40*, 2647-2652.
- (20) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704-708.
- (21) Hoffmann, J. S.; Pillaire, M. J.; Maga, G.; Podust, V.; Hübeshner, U.; Villani, G. *Proc. Natl. Acad. Sci.* **1995**, *92*, 5356-5360.
- (22) Tae, L. T.; Wu, Y.; Xia, G.; Schultz, P. G.; Romesburg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 7439-7440.