Pure Appl. Chem., Vol. 76, Nos. 7–8, pp. 1563–1570, 2004.

Abasic site stabilization by aromatic DNA base surrogates: High-affinity binding to a base-flipping DNA-methyltransferase*

Ishwar Singh[1], Christine Beuck[2], Anupam Bhattacharya[1], Walburga Hecker[3], Virinder S. Parmar[1], Elmar Weinhold[2], and Oliver Seitz[4,‡]

[1]Bioorganic Laboratory, Department of Chemistry, University of Delhi, Delhi 110 007, India; [2]Institut für Organische Chemie, RWTH Aachen, Prof.-Pirlet-Strasse 1, 52 056 Aachen, Germany; [3]Max-Planck-Institut für Molekulare Physiologie, Otto-Hahn-Strasse 11, 44 227 Dortmund, Germany; [4]Institut für Chemie der Humboldt-Universität zu Berlin, Brook-Taylor-Str. 2, D-12489 Berlin, Germany

Abstract: DNA-methyltransferases catalyze the sequence-specific transfer of the methyl group of *S*-adenosylmethionine to target bases in genomic DNA. For gaining access to their target embedded within a double-helical structure, DNA-methyltransferases (DNA-MTases) rotate the target base out of the DNA helix. This base-flipping leads to the formation of an apparent abasic site. MTases such as cytosine-specific M·*Hha*I and M·*Hae*III and also the repair enzyme uracil DNA glycosylase (UDG) insert amino acid side chains into the opened space and/or rearrange base-pairing. The adenine-specific DNA MTase M·*Taq*I binds without amino acid insertion. This binding mode allows for a substitution of the orphaned thymine with larger DNA base surrogates without steric interference by inserted amino acid side chains. DNA containing pyrenyl, naphthyl, acenaphthyl, and biphenyl residues was tested in M·*Taq*I binding studies. The synthesis of DNA building blocks required the formation of a *C*-glycosidic bond, which was established by using protected 1-chloro-2-deoxyribose as glycosyl donor and organocuprates as glycosyl acceptors. It is shown that all of the base surrogates enhanced the binding affinity to M·*Taq*I. Incorporation of pyrene increased the binding affinity by a factor of 400. Interestingly, there is a correlation between the observed order of dissociation constants and the ability of a base surrogate to stabilize abasic sites in model duplexes.

Nucleic acids are like proteins and carbohydrates subject to a multitude of biological modification reactions. In genomic DNA, the exocyclic amino groups of adenine or cytosine or the C5 of cytosine often carry methyl groups [1,2]. DNA methylation serves diverse cellular functions [3–6]. For example, a specific methylation pattern endows many bacteria with the ability to distinguish self from non-self DNA. Some bacteria use N6-methylation of adenine as a means for cell cycle control. Methyl groups in DNA of humans and other mammals play an important role in the regulation of gene expression.

DNA-methylation is performed by methyltransferases which catalyze the sequence-specific transfer of the methyl group of *S*-adenosylmethionine to their respective target base [6,7]. For gaining access

*Lecture presented at the symposium "Chemistry of nucleic acids", as part of the 39th IUPAC Congress and 86th Conference of the Canadian Society for Chemistry: Chemistry at the Interfaces, Ottawa, Canada, 10–15 August 2003. Other Congress presentations are published in this issue, pp. 1295–1603.

‡Corresponding author: Tel.: 49 89 20937446; E-mail: oliver.seitz@chemie.hu-berlin.de

to their target embedded within a double-helical structure, DNA-methyltransferases (DNA-MTases) have evolved DNA-binding modes that lead to a local disruption of hydrogen-bonding and base-stacking interactions. Eventually, the target base is swung out of the interior of the helix, bound in an extrahelical conformation and placed in the vicinity of the catalytic machinery of the enzyme [8–10]. This so-called base-flipping has been first discovered in crystal structures of the MTase M·*Hha*1 complexed with DNA (Fig. 1A) and was shown to be characteristic for many other DNA-modifying enzymes, including the well-studied repair enzyme UDG-glycosylase [11].

In the process of base-flipping, an apparent abasic site is formed and any base-flipping enzyme has to cope with energetic penalty introduced by the disruption of hydrogen-bonding and base-stacking interactions. It appears that based upon existing crystal structures, different mechanisms of abasic site stabilization have evolved. For example, the cytosine-specific DNA-MTase M·*Hha*I (Fig. 1A) inserts a glutamine side chain into the space opened after enzymatic base-flipping [11]. The glutamine residue forms hydrogen bonds with the orphaned guanine. The repair enzyme uracil DNA glycosylase (UDG) employs a similar tactic by using a leucine side chain as a mechanical wedge to "push" the uracil out

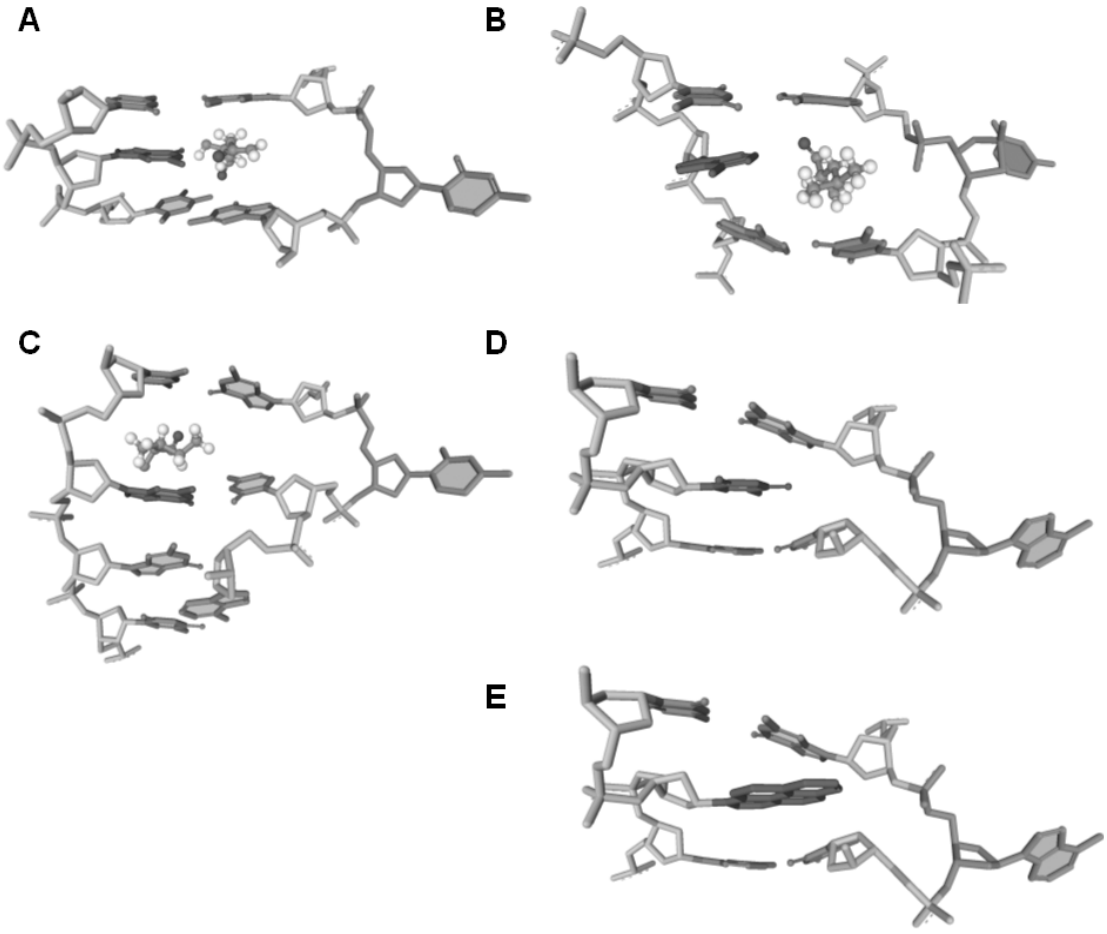


Fig. 1 DNA structures observed in cocrystals with (A) DNA MTase M·*Hha*I, (B) uracil DNA glycosylase (UDG), (C) DNA MTase M·*Hae*III, and (D) DNA MTase M·*Taq*I. The flipped bases are shown on the right. Note that in (B) the flipped base is already cleaved. Amino acids that reach into the duplex interior are represented as ball-and-stick models. (E) A modeled replacement, which illustrates that pyrene fits into the cavity left by the flipped adenine and potentially improves interstrand stacking.

of the duplex (Fig. 1B) [12]. In the structure of the cytosine DNA-Mtase M·*Hae*III, Ile 221 protrudes into the DNA helix, and the orphaned guanine pairs with the 3′-cytosine that flanks the flipped target cytosine (Fig. 1C) [13]. These binding modes are characterized by insertion of amino acid side chains into the DNA helix. This is in contrast to the structure observed in the binding of adenine-specific DNA MTase M·*Taq*I to target DNA (Fig. 1D) [14]. M·*Taq*I binds without amino acid insertion and instead displaces the orphaned thymine toward the helix center in order to enhance interstrand stacking with the 5′ neighboring base of the target adenine.

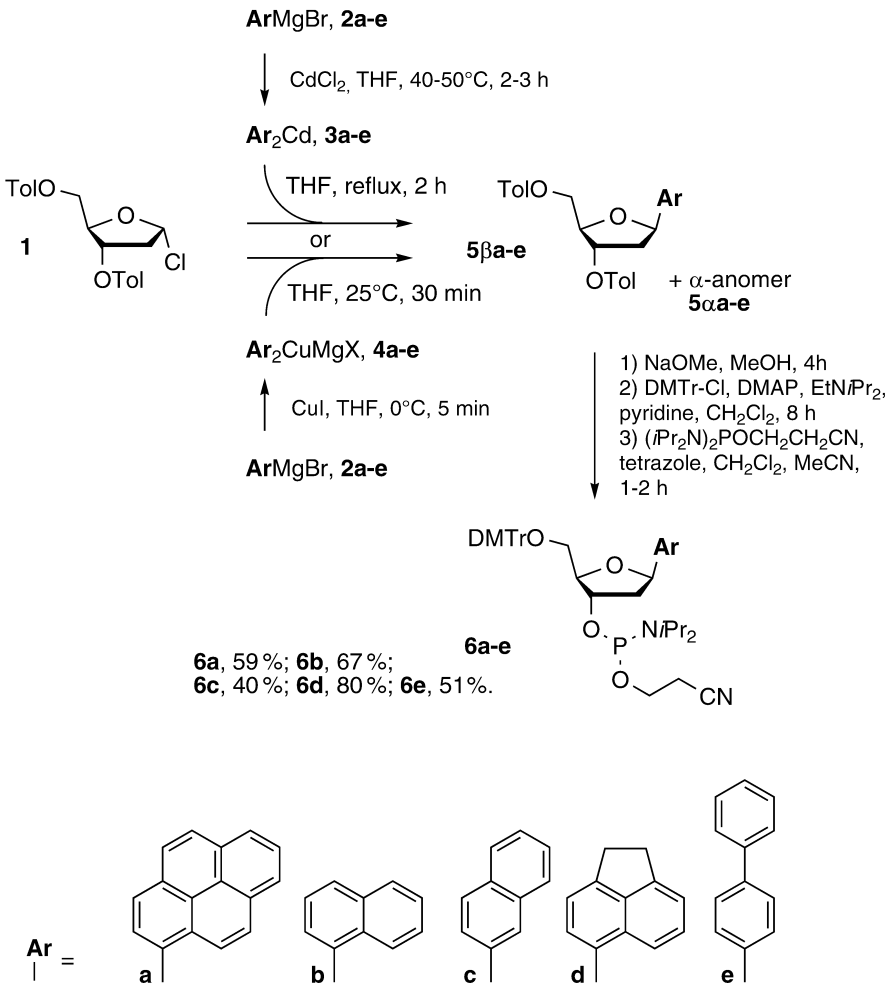
We reckoned that the mode of M·*Taq*I-DNA binding would allow for a substitution of the orphaned thymine with a larger aromatic DNA-base surrogate. Since M·*Taq*I binding occurs without concomitant amino acid side chain insertion, an aromatic base surrogate such as pyrene can fit well into the cavity left by the flipped target adenine (Fig. 1E) [15]. It was expected that more complete filling of the apparent abasic site would lead to enhanced interstrand stacking, which ultimately would result in a tightening of the MTase-DNA complex. The DNA-MTase could thus bind without paying the energetic cost for disrupting Watson–Crick hydrogen bonds and base-stacking interactions. Furthermore, it can be imagined that due to steric interference such binding enhancements would not occur with base-flipping enzymes that insert amino acid side chains into the opened space. If this were true, then the introduction of large base surrogates would allow a probing of the mechanism of base-flipping enzymes that have failed to cocrystallize with their DNA-substrates. In the following, we present our data on the effect of aromatic base surrogates on DNA structure [16] and DNAM·*Taq*I binding (no amino acid side-chain insertion) [15] and compare with work from Stivers [17,18] in which a pyrene nucleotide was explored in UDG (Leu191 insertion) binding studies.

For the incorporation of polycyclic, aromatic base surrogates into DNA a robust access to modified phosphoamidites **6** was required (Scheme 1). It was, hence, necessary to develop a methodology that establishes the *C*-glycosidic bond in building blocks **5**. We and others used the 2-deoxyribose donor **1** [19] and cadmium-organic reagents such as **3** as glycosyl acceptors [16,20,21]. Kool and coworkers have shown that the conversion of the aryl-Grignard reagents **2a–c** into less basic, but toxic cadmium-organic species **3a–c** was necessary in order to minimize the formation of elimination by-products such as glycals and furans [22]. In seeking a replacement for the toxic cadmium and with the aim of enhancing the sometimes low *C*-glycosylation yields we considered the use of alternative organometal reagents. Organocuprates are known as mild and powerful *C*-nucleophiles in coupling reactions with reactive alkyl halides. Notably, the reaction of glycopyranosyl halides with lithium dialkyl cuprates has been reported [23]. However, cuprate-mediated *C*-coupling of 1-halo-2-deoxysugars, which exhibit a pronounced sensitivity to base-induced elimination, have not been demonstrated. We supposed that Normant reagents would display a higher nucleophilicity than cadmium-organic reagents. Indeed, the use of Normant reagents **4a–e** allowed couplings at 25 °C as opposed to 66 °C required for the *C*-glycosylation of the cadmium reagents **3a–e.** In conclusion, the cuprate-mediated *C*-glycosylation proved superior to the use of organocadmium reagents in terms of both synthetic practicability and yields (Table 1). For completion of the building block synthesis, the toloyl groups were removed followed by regioselective introduction of the trityl group and phosphitylation.

Table 1 Yields of the *C*-glycosylation reactions[a].

	5a	**5b**	**5c**	**5d**	**5e**
With $\mathbf{Ar}_2$Cd, **3a–e**	72 % (4:1)	67 % (4:1)	47 % (2:1)	47 % (2:1)	41 % (2:1)
With $\mathbf{Ar}_2$CuMgX, **4a–e**	74 % (3:1)	81 % (4:1)	74 % (3:1)		75 % (2:1)

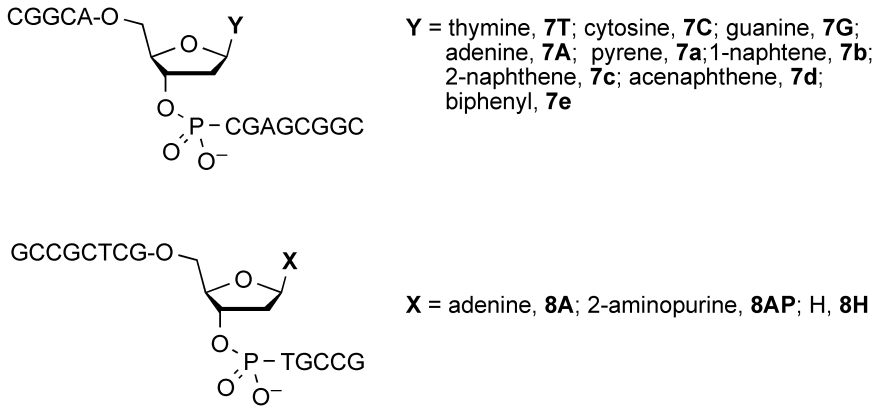
[a]The α:β ratio is given in brackets.



Scheme 1 Synthesis of phosphoramidites containing bulky base surrogates. DMTr = 4,4′-dimethoxytrityl, Tol = toluoyl.

The phosphoramidite building blocks **6a–e** were used in the automated synthesis of oligodeoxynucleotides **7a–e** (Scheme 2). It was first explored whether the aromatic base surrogates had any influence on DNA duplex stability. It was found that the substitution of thymine in **7T·8A** by the aromatic base surrogates in **7a–e** led to a decrease in the melting temperature T_M by 4.9–8.7 °C (Table 2, second row). The destabilization was lowest with the pyrene residue in **7a·8A**, which confirms results from Kool and coworkers who were the first to study pyrene base-pairing [24,25]. Interestingly, a comparison with the T_M = 55.7 °C and T_M = 56.1 °C measured for C-A and A-A pairs in **7C·8A** and **7A·8A**, respectively, revealed that base pairs comprised of adenine and pyrenyl, naphthyl, acenaphthyl, or biphenyl groups can confer higher stabilities than mismatched base pairs.

The T_M data confirmed the integrity of the duplex structures, which was deemed essential in order to perform the MTase binding studies. The binding affinity of duplexes **7·8A** to the M·*Taq*I MTase was investigated in solution (Table 2, third row). M·*Taq*I bound its natural substrate **7T·8A** with a dissociation constant K_D = 20 nM. Complexes with duplexes **7C·8A, 7G·8A**, and **7A·8A** containing target adenine opposite to C, G, and A displayed almost equal stability. This suggests that interference with hydrogen bonds to the target adenine has little effect on the M·*Taq*I binding affinity. This finding dif-



Scheme 2 Oligucleotides synthesized for hybridization and binding studies.

Table 2 Thermal stability T_M of duplexes, dissociation constants K_D of complexes with the adenine-specific DNA MTase M·*Taq*I and relative fluorescence of 2-aminopurine (2AP) containing duplexes[a].

```
5'-GCCGC T C G dRibX TGCCG-3'   8A; X=A, 8AP; X=2AP, 8H; X=H
3'-CGGCG A G C dRibY ACGGC-5'   7T, 7C, 7G, 7A, 7a-e
```

Y	X = adenine, **8A**		X = 2AP, **8AP**		X = H, **8H**	
	T_M[b]/°C	K_D[c]/nM^{-1}	T_M[b]/°C	Relative fluorescence[d]	T_M[b]/°C	ΔT_M[e]/°C
thymine, **7T**	65.9	20	63.2	1.0	49.8	0
cytosine, **7C**	55.7	30	nd	nd	50.6	+0.8
guanine, **7G**	61.0	30	56.1	4.3	48.7	–1.1
adenine, **7A**	56.1	25	nd	nd	52.5	+2.7
1-pyrenyl, **7a**	61.0	0.05	63.5	1.2	63.6	+13.8
4-biphenyl, **7e**	57.5	0.30	59.1	5.8	59.0	+9.2
acenaphthyl, **7d**	59.0	0.50	58.5	1.4	57.9	+8.1
1-naphthyl, **7b**	57.2	0.50	58.0	1.6	56.0	+6.2
2-naphthyl, **7c**	58.0	1.0	58.2	2.0	57.8	+8.0

[a]The double-stranded recognition sequence of M·*Taq*I is boxed.

[b]Measured as denaturation curves at 1 μM DNA concentration in a buffered solution (100 mM NaCl, 10 mM NaH_2PO_4, pH 7.0) by using a temperature gradient of 0.2 °C/min.

[c]K_D values were calculated by curve fitting of fluorescence titration curves acquired by titrating M·*Taq*I to a mixture of duplexes **7·8A** and a 36mer competitor duplex which contained 2-aminopurine at the target position. The fluorescence signal generated was followed after each enzyme addition until no further fluorescence change was observed to assure that final fluorescence intensities reflect equilibrium and true binding affinities are obtained.

[d] Increase of fluorescence of duplex **7·8AP** relative to the duplex **7T·8AP**.

[e]ΔT_M values are based on the T_M value of **8H·7T**. nd, not determined.

fers from results obtained with cytosine-specific M·*Hha*I DNA-Mtase, which exhibited up to 20-fold enhanced affinity upon binding to mismatched target cytosine [26].

While the natural nucleobases in matched duplex **7T·8A** and mismatched duplexes **7C·8A**, **7G·8A**, and **7A·8A** conferred almost equal stabilities of DNA-M·TaqI complexes, all base surrogates in duplexes **7a–e·8A** were found to stabilize the DNA-enzyme complex. The K_D values were in the subnanomolar range, which corresponds to 20–60-fold enhancements of the binding affinity. The pyrenyl residue in duplex **7a·8A** showed a particularly pronounced tightening of the DNA-M·*Taq*I complex.

The binding affinity was enhanced by a factor of 400, which corresponded to a dissociation constant K_D = 50 pM. The binding studies, hence, revealed that the affinity of binding to M·*Taq*I was enhanced upon introduction of aromatic base surrogates, but remained virtually unchanged upon introduction of mismatched natural nucleobases. In seeking an explanation, we considered two major effects. A base surrogate that replaces the orphaned thymine can possibly enhance binding affinity by increasing the proclivity of the target adenine to accommodate an extrahelical conformation. In this case, the DNA-M·*Taq*I binding would profit from an unstacked target base. Alternatively, the binding affinity can be enhanced by filling the opened space after enzymatic base-flipping had occurred. Here, reinsertion of the target base into the interior of the duplex would be prevented and a better filling of the apparent abasic site would compensate the energetic penalty that arises from the enzyme-induced abasic site formation.

The first possibility was investigated in fluorescence studies based on the use of 2-aminopurine (2AP) [27] as fluorescent base analog of adenine [16]. 2AP is virtually nonfluorescent when stacked within a DNA duplex, but fluoresces upon disruption of base-stacking [28–30]. In duplexes **7·8AP**, 2AP replaced the target adenine and served as a probe of base-stacking interactions. Table 2 lists the relative fluorescence of 2AP-containing duplexes. It became apparent that pyrenyl, naphthyl, and acenaphthyl residues in duplexes **7·8AP** increased the 2AP fluorescence by a factor of 1.2–2.0. The mismatched guanine in **7G·8AP**, and particularly the biphenyl base in **7e·8AP**, resulted in large 4.6–5.8-fold enhancements of 2AP fluorescence. This data suggests that guanine and the biphenyl base are more efficient in disrupting base-stacking than pyrene, naphthene, and acenaphthene. Nevertheless, the pyrene residue in **7a·8A** despite being a less potent base unstacker conferred tighter DNA-M·*Taq*I binding than guanine and biphenyl residues in **7G·8A** and **7e·8A**. It can be concluded that disruption of target base-stacking is not the decisive determinant.

Enzymatic base-flipping leads to the formation of an apparent abasic site. In the case of DNA-M·*Taq*I binding, this cavity in the interior of DNA remains unfilled [14]. It could, hence, be possible that the aromatic base surrogates stabilize the abasic site and restore contiguous base-stacking interactions. The ability to stabilize abasic sites was evaluated by T_M studies with model duplexes **7·8H** in which a tetrahydrofurane nucleotide mimicked enzymatic base-flipping (Table 2, rows 6 and 7). A comparison of the T_M = 49.8 °C measured for the abasic site-containing duplex **7T·8H** with the T_M = 65.9 °C provided by the T-A pair in **7T·8A** demonstrates how efficient abasic sites are in decreasing duplex stability. This destabilization also occurred upon introduction of abasic site opposite to the other natural nucleobases as evidenced by T_M = 50.6, 48.7, and 52.5 °C for C, G, and A containing duplexes **7C·8H**, **7G·8H**, and **7A·8H**, respectively. In contrast to the natural nucleobases, each of the aromatic base surrogates in **7a–e·8H** was able to enhance the T_M of abasic site-containing duplexes. The pyrene residue conferred the highest abasic site stabilization. A comparison of the T_M = 63.6 °C obtained for duplex **7a·8H** containing pyrene opposite to the abasic site and the T_M = 65.9 °C measured for the non-modified duplex **7T·8A** suggests that the pyrene-abasic site pair provided a "pairing strength" that almost reached that of an T-A base pair. The correlation between the ability to stabilize abasic sites and to enhance the binding affinities to M·*Taq*I appears striking. The pyrenyl residue provided the highest thermal stability to abasic site duplexes and was also most efficient in tightening the DNA-M·*Taq*I complex. In addition, the biphenyl base was superior to the acenaphthyl and naphthyl residues in terms of both abasic site stabilization and enhancement of binding affinities to M·*Taq*I. Therefore, it appears plausible to conclude that the pyrene residue confers the remarkable 400-fold enhancement of binding affinity mainly by compensating the energetic penalty from enzyme-induced abasic site formation.

It is instructive to compare the data on M·*Taq*I binding with studies of uracil DNA glycosylase (UDG) and variants with a duplex that contained a pyrenyl residue opposite to 2′-fluoro-2′-deoxyuridine (U^β) as uridine analog [17,18]. UDG fills the abasic site formed upon flipping of the target uracil by insertion of the side chain of leucine 191 [12]. Replacement of the A-U^β base pair by a pyrenyl-U^β base pair led to a 3.7-fold enhancement of binding affinity. In binding to UDG, pyrene might facilitate base-flipping since uracil is less tightly bound in the interior of the duplex. On the other

hand, steric interferences with the inserted leucine side chain would destabilize the DNA-UDG complex once base-flipping had occurred. As a result, only a minor change in the binding affinity to UDG is observed. This is in contrast to the 400-fold enhancement of binding affinity to M·*Taq*I that was observed when the T-A base pair in **7T·8A** was replaced by the pyrene-A pair in **7a·8A**. In M·*Taq*I, the pyrenyl residue fits well into the opened space formed after base-flipping since amino acids are not inserted into the DNA duplex. As a result, pyrene can restore the base stack possibly by employing both intra- and interstrand stacking interactions. This interpretation is in agreement with the results from binding studies to UDG mutants in which the inserted Leu 191 had been replaced by glycine [17,18]. Removal of the critical leucine side chain reduced the binding affinity by a factor of 43. The pyrene residue completely restored the damaging effect of the L191G mutation.

From the studies of UDG and M·*Taq*I binding to pyrene-containing duplexes, a picture emerges that suggests that significant binding enhancements only occur in cases where insertion of amino acid side chains into the duplex interior is avoided. M·*Taq*I binds by omitting amino acid insertion. A large hydrophobic base surrogate such as pyrene can be easily accommodated into the abasic site formed after enzymatic base-flipping. As a result, base surrogates that are able to stabilize the abasic site proved efficient in enhancing DNA-enzyme binding. Such binding enhancements were not observed in binding studies with UDG that inserts a leucine side chain into the space opened after enzymatic base-flipping. The comparison between M·*Taq*I and UDG illustrates that duplexes with aromatic base surrogates opposite to the target base might allow us to distinguish between different base-flipping mechanisms.

REFERENCES

1. I. Ahmad and D. N. Rao. *Crit. Rev. Biochem. Mol. Biol.* **31**, 361–380 (1996).
2. A. Jeltsch. *ChemBioChem* **3**, 275–293 (2002).
3. K. D. Robertson and P. A. Jones. *Carcinogenesis* **21**, 461–467 (2000).
4. S. A. Wajed, P. W. Laird, T. R. DeMeester. *Ann. Surg.* **234**, 10–20 (2001).
5. R. L. Momparler and V. Bovenzi. *J. Cell. Physiol.* **183**, 145–154 (2000).
6. X. D. Cheng and R. J. Roberts. *Nucleic Acids Res.* **29**, 3784–3795 (2001).
7. X. D. Cheng. *Ann. Rev. Biophys. Biomol. Struct.* **24**, 293–318 (1995).
8. R. J. Roberts and X. D. Cheng. *Ann. Rev. Biochem.* **67**, 181–198 (1998).
9. R. J. Roberts. *Cell* **82**, 9–12 (1995).
10. D. P. Hornby and G. C. Ford. *Curr. Opin. Biotechnol.* **9**, 354–358 (1998).
11. S. Klimasauskas, S. Kumar, R. J. Roberts, X. Cheng. *Cell* **76**, 357–369 (1994).
12. S. S. Parikh, G. Walcher, G. D. Jones, G. Slupphaug, H. E. Krokan, G. M. Blackburn, J. A. Tainer. *Proc. Natl. Acad. Sci. USA* **97**, 5083–5088 (2000).
13. K. M. Reinisch, L. Chen, G. L. Verdine, W. N. Lipscomb. *Cell* **82**, 143–153 (1995).
14. K. Goedecke, M. Pignot, R. S. Goody, A. J. Scheidig, E. Weinhold. *Nat. Struct. Biol.* **8**, 121–125 (2001).
15. C. Beuck, I. Singh, A. Bhattacharya, W. Hecker, V. S. Parmar, O. Seitz, E. Weinhold. *Angew. Chem., Int. Ed.* **42**, 3958–3960 (2003).
16. I. Singh, W. Hecker, A. K. Prasad, V. S. Parmar, O. Seitz. *Chem. Commun.* 500–501 (2002).
17. Y. L. Jiang, K. Kwon, J. T. Stivers. *J. Biol. Chem.* **276**, 42347–42354 (2001).
18. Y. L. Jiang and J. T. Stivers. *Biochemistry* **41**, 11248–11254 (2002).
19. M. Hoffer. *Chem. Ber.-Recl.* **93**, 2777–2781 (1960).
20. A. K. Ogawa, Y. Q. Wu, M. Berger, P. G. Schultz, F. E. Romesberg. *J. Am. Chem. Soc.* **122**, 8803–8804 (2000).
21. R. X. F. Ren, N. C. Chaudhuri, P. L. Paris, S. Rumney, E. T. Kool. *J. Am. Chem. Soc.* **118**, 7671–7678 (1996).
22. N. C. Chaudhuri and E. T. Kool. *Tetrahedron Lett.* **36**, 1795–1798 (1995).
23. R. Bihovsky, C. Selick, I. Giusti. *J. Org. Chem.* **53**, 4026–4031 (1988).

24. T. J. Matray and E. T. Kool. *J. Am. Chem. Soc.* **120**, 6191–6192 (1998).
25. T. J. Matray and E. T. Kool. *Nature* **399**, 704–708 (1999).
26. S. Klimasauskas and R. J. Roberts. *Nucleic Acids Res.* **23**, 1388–1395 (1995).
27. D. C. Ward, E. Reich, L. Stryer. *J. Biol. Chem.* **244**, 1228–1237 (1969).
28. C. R. Guest, R. A. Hochstrasser, L. C. Sowers, D. P. Millar. *Biochemistry* **30**, 3271–3279 (1991).
29. B. W. Allan and N. O. Reich. *Biochemistry* **35**, 14757–14762 (1996).
30. B. Holz, S. Klimasauskas, S. Serva, E. Weinhold. *Nucleic Acids Res.* **26**, 1076–1083 (1998).