ORGANIC LETTERS

2000
Vol. 2, No. 9
1341−1344

Recognition of the Nonpolar Base 4-Methylindole in DNA by the DNA Repair Adenine Glycosylase MutY

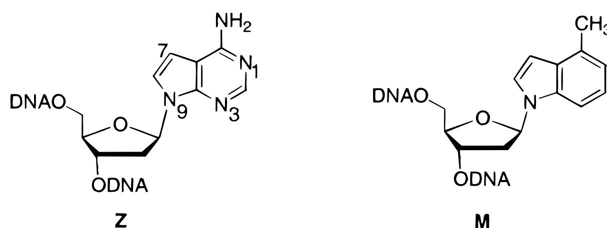
Cindy Lou Chepanoske, Charles R. Langelier, Nikolas H. Chmiel, and Sheila S. David*

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112

david@chemistry.chem.utah.edu

Received March 17, 2000

ABSTRACT



The DNA repair adenine glycosylase MutY efficiently recognizes 7-deaza-2′-deoxyadenosine (Z) and its nonpolar isostere 4-methylindole β-deoxynucleoside (M) opposite 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG) and G in DNA. Both wild-type and truncated MutY exhibit a 10- to 20-fold higher affinity for a duplex containing OG:M than OG:Z. More efficient recognition of M over Z by MutY may be to due the lack of hydrogen bonding with the OG that facilitates nucleotide flipping during the substrate recognition process.

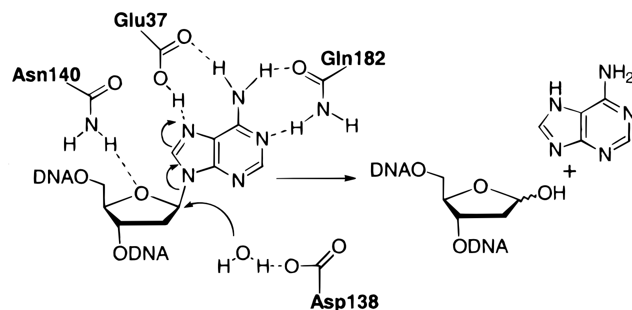
All organisms can efficiently counter DNA damage through elaborate pathways that recognize and repair a variety of potentially mutagenic and lethal base modifications.[1−3] A prevalent insult to the cell is oxidative damage to DNA, and one of the most stable products formed in vivo by reactive oxygen species is 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG).[4,5] Due to the propensity for insertion of A opposite OG by DNA polymerase during replication, the presence of OG can result in deleterious DNA mutations. In *Escherichia coli*, an enzymatic pathway has evolved dedicated to preventing mutations caused by OG, including the enzymes MutY, MutM, and MutT.[6] Homologues to enzymes in this pathway have also been found in other prokaryotes and eukaryotes.[7]

MutY is a base excision repair (BER) enzyme specific for OG mispaired with A, but also exhibits activity toward G:A and C:A mispairs in DNA.[6,8−10] MutY uses adenine glycosylase activity to remove misincorporated A residues, thereby preventing G to T transversion mutations.[11] The proposed mechanism of MutY based on the X-ray structure of the N-terminal domain with a bound adenine base[12] (Scheme 1) involves activation of a water molecule for nucleophilic attack of the glycosyl linkage, in a manner analogous to that proposed for other monofunctional DNA glycosylases.[7]

(1) Friedberg, E. C.; Walker, G. C.; Siede, W. *DNA Repair and Mutagenesis*; ASM Press: Washington, DC, 1995.
(2) Lindahl, T. *Nature* **1993**, *362*, 709−715.
(3) Lindahl, T.; Wood, R. D. *Science* **1999**, *286*, 1897−1905.
(4) Shinenaga, M. K.; Park, J.-W.; Cundy, K. C.; Gimeno, C. J.; Ames, B. N. *Methods Enzymol.* **1990**, *186*, 521−530.
(5) Dizdaroglu, M. *Biochemistry* **1985**, *24*, 4476−4481.
(6) Michaels, M. L.; Tchou, J.; Grollman, J.; Miller, J. H. *Biochemistry* **1992**, *31*, 10964−10968.
(7) David, S. S.; Williams, S. D. *Chem. Rev.* **1998**, *98*, 1221−1261.
(8) Radicella, J. P.; Clark, E. A.; Fox, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9674−9678.
(9) Tsai-Wu, J.-J.; Liu, H.-F.; Lu, A.-L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8779−8783.
(10) Porello, S. L.; Leyes, A. E.; David, S. S. *Biochemistry* **1998**, *37*, 14756−14764.
(11) Shibutani, S.; Takeshita, M.; Grollman, A. P. *Nature* **1991**, *349*, 431−434.
(12) Guan, Y.; Manuel, R. C.; Arvai, A. S.; Parikh, S. S.; Mol, C. D.; Miller, J. H.; Lloyd, R. S.; Tainer, J. A. *Nature Struct. Biol.* **1998**, *5*, 1058−1064.

10.1021/ol005831o CCC: $19.00 © 2000 American Chemical Society
Published on Web 04/14/2000

Scheme 1. Proposed Catalytic Mechanism of the Glycosylase Reaction Catalyzed by MutY, Adapted from Ref 12



MutY is a 39 kDa $[4Fe-4S]^{2+}$cluster-containing protein that can be separated into two distinct domains via limited proteolysis. The N-terminal domain (Met 1−Lys 225) is catalytically active and exhibits high sequence and structural homology to enzymes of the BER superfamily,[13−15] particularly those of the endonuclease III-like subfamily that contain a $[4Fe-4S]^{2+}$ center.[12,13] Though no structural information on the unique C-terminal domain (Gln 226−Val 350) of MutY is available, it was recently shown that this domain exhibits sequence homology to the d(OG)TPase MutT, suggesting that this domain may be involved in recognition of OG.[16]

Damage recognition by BER glycosylases has been proposed to utilize a "nucleotide flipping" mechanism[17] where the base being excised is extruded from the DNA helix and placed into a base-specific catalytic pocket. Evidence for this type of mechanism has been provided by X-ray crystallography of BER glycosylases bound to substrate- and substrate mimic-containing DNA.[18−22]

Our laboratory has been exploring the role of structural motifs and specific catalytic residues of MutY in the recognition and repair of OG:A and G:A mispairs, as well as determining the important features of the DNA substrate that facilitate recognition and repair. Previous results from our laboratory have shown that substrate recognition is facilitated by interactions of multiple regions of MutY, including the iron−sulfur cluster loop (FCL) domain and the C-terminal domain,[23,24] with the DNA substrate. In addition, we have observed that the presence of both A *and* OG have a dramatic influence on the recognition and repair properties of MutY.[10,24,25]

An approach to elucidating factors that influence damage recognition in BER enzymes is the use of substrate analogues incorporated into DNA that mimic structural features of the substrate but are resistant to enzymatic turnover.[7] We have used a variety of 2′-deoxyadenosine analogues to probe features of MutY that are important for recognition including 2′-deoxyformycin A (F), 7-deaza-2′-deoxyadenosine (Z),[25] and 2′-deoxy-2′-fluoroadenosine (FA).[26] In this work, we have incorporated the nonpolar isostere of A and Z, 4-methylindole β-deoxynucleoside (M), into a DNA duplex opposite G and OG and investigated the affinity of MutY for these duplexes compared to FA and Z (Figure 1). In addition, we have overexpressed a truncated form of MutY (Met 1−Lys 225, referred to henceforth as Stop 225) and investigated the consequence of removal of the C-terminal domain on the recognition of the M, Z, and FA analogues.

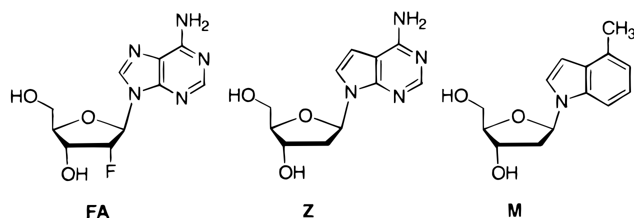


Figure 1. Substrate analogues and a nonpolar isostere of 2′-deoxyadenosine: 2′-deoxy-2′-fluoroadenosine (FA), 7-deaza-2′deoxyadenosine (Z), and 4-methylindole β-deoxynucleoside (M).

Under single-turnover conditions, neither MutY nor Stop 225 displays adenine glycosylase activity toward duplexes containing G:M or OG:M mispairs (data not shown). To determine whether the hydrophobic isostere M mimics features of the substrate, equilibrium dissociation constants (K_d) were determined for MutY and Stop 225 by a gel-retardation method[27] as described previously[26] for 30-mer duplexes containing the OG:X or G:X base pairs where X = M, Z, FA, or C (Table 1).

Table 1. Equilibrium Dissociation Constants for WT and Truncated MutY with the 30 Base Pair Duplex: (5′-CGATCATGGAGCCAC**Y**AGCTCCCGTTACAG-3′)• (3′-GCTAGTACCTCGGTG**X**TCGAGGGCAATGTC-5′), Where Y = OG or G and X = M, Z, FA, or C

central base pair	WT MutY (nM)	Stop 225 (nM)
OG:M	0.17 ± 0.05	1.4 ± 0.3
OG:Z	3.4 ± 1.7^{a}	15 ± 4
OG:FA	0.12 ± 0.05^{b}	90 ± 10
G:M	40 ± 6	20 ± 10
G:Z	35 ± 17^{a}	21 ± 6
G:FA	5.8 ± 0.6^{b}	80 ± 10
G:C	150 ± 50^{b}	130 ± 30

[a,b] K_d values listed for WT are from refs 25 and 26, respectively. All values have been corrected for activity for direct comparison.

(13) Michaels, M. L.; Pham, L.; Nghiem, Y.; Cruz, C.; Miller, J. H. *Nucleic Acids Res.* **1990**, *18*, 3841−3845.
(14) Manuel, R. C.; Czerwinski, E. W.; Lloyd, R. S. *J. Biol. Chem.* **1996**, *271*, 16218−16226.
(15) Gogos, A.; Cillo, J.; Clarke, N. D.; Lu, A.-L. *Biochemistry* **1996**, *35*, 16665−16671.
(16) Noll, D. M.; Gogos, A.; Granek, J. A.; Clarke, N. D. *Biochemistry* **1999**, *38*, 6374−6379.
(17) Roberts, R. J.; Cheng, X. *Ann. Rev. Biochem.* **1998**, *67*, 181−198.
(18) Mol, C. D.; Arvai, A. S.; Slupphaug, G.; Kavli, B.; Alseth, I.; Krokan, H. E.; Tainer, J. A. *Cell* **1995**, *80*, 1−20.
(19) Savva, R.; McAuley-Hecht, K.; Brown, T.; Pearl, L. *Nature* **1995**, *373*, 487−493.
(20) Vassylyev, D. G.; Kashiwagi, T.; Mikami, Y.; Ariyoshi, M.; Iwai, S.; Ohtsuka, E.; Morikawa, K. *Cell* **1995**, *83*, 773−782.
(21) Lau, A. Y.; Schärer, O. D.; Samson, L.; Verdine, G. L.; Ellenberger, T. *Cell* **1998**, *95*, 249−258.
(22) Hollis, T.; Ichikawa, Y.; Ellenberger, T. *EMBO J.* **2000**, *19*, 758−765.

MutY has a high affinity for duplexes containing OG:M ($K_d = 0.17 \pm 0.05$ nM), which rivals the affinity observed with other 2′-deoxyadenosine analogues. Indeed, a comparable K_d value was obtained using the FA analogue ($K_d = 0.12 \pm 0.05$ nM), which has minimal alteration of structure and recognition elements relative to that of A. This result is surprising since in the X-ray structure of the N-terminal domain of MutY with a bound adenine there are a large number of specific hydrogen-bonding contacts with the adenine base (Scheme 1). This is further underscored by the 20-fold higher affinity of MutY for the OG:M-containing duplex relative to the OG:Z duplex, in which only the hydrogen-bond contact with N-7 has been disrupted. An explanation for the high affinity of MutY for the OG:M duplex may be the lack of hydrogen bonding between the OG and M, which may facilitate an important recognition event. Indeed, the absence of an energetic cost associated with disrupting the base pair may considerably ease a nucleotide flipping process and therefore compensate for the absence of favorable hydrogen-bonding interactions between MutY and M.

These results also illustrate that high-affinity recognition of M requires the presence of OG. The binding affinity of MutY to the duplexes with the M and Z analogues opposite G are similar, and MutY binds G:Z and G:M duplexes with a 10- to 200-fold lower affinity, respectively, than the corresponding OG-containing duplexes. We and others have observed a similar trend with other analogues previously,[25,26,28,29] highlighting the importance of OG in the base pair for recognition. In the recognition of the A analogues opposite G, the nature of the A analogue appears to be more significant based on the higher affinity of MutY for the G:FA duplex relative to G:M and G:Z. However, based on the considerable differences in their structures, it is surprising that the M and Z analogues behave similarly. Furthermore, MutY binds with higher affinity to both M and Z opposite G than the corresponding G:C-containing (nonspecific) duplex.

The dissociation constants of Stop 225 for M, Z, and FA in duplexes opposite OG and G are also listed in Table 1. Interestingly, Stop 225 does not bind the FA-containing duplex with high affinity nor does this truncated form retain the ability to discriminate between FA opposite OG and G ($K_d = 90 \pm 10$ nM and 80 ± 10 nM, respectively). This result is in agreement with the proposed role of the C-terminal domain in OG recognition, and kinetic experiments that show that the C-terminal domain confers specificity for OG:A mispairs over G:A mispairs.[16,24]

(23) Chepanoske, C. L.; Golinelli, M.-P.; Williams, S. D.; David, S. S. *Arch. Biochem. Biophys.*, in press.
(24) David, S. S.; Chmiel, N. H. Unpublished results.
(25) Porello, S. L.; Williams, S. D.; Kuhn, H.; Michaels, M. L.; David, S. S. *J. Am. Chem. Soc.* **1996**, *118*, 10684−10692.
(26) Chepanoske, C. L.; Porello, S. L.; Fujiwara, T.; Sugiyama, H.; David, S. S. *Nucleic Acids Res.* **1999**, *27*, 3197−3204.
(27) Carey, J. *Methods Enzymol.* **1991**, *208*, 103−117.
(28) Bulychev, N. V.; Varaprasad, C. V.; Dorman, G.; Miller, J. H.; Eisenberg, M.; Grollman, A. P.; Johnson, F. *Biochemistry* **1996**, *35*, 13147−13156.
(29) Deng, L.; Schärer, O. D.; Verdine, G. L. *J. Am. Chem. Soc.* **1997**, *119*, 7865−7866.

Surprisingly, Stop 225 exhibits a 60-fold higher affinity for the OG:M duplex (1.4 ± 0.3 nM) compared to that of the duplex containing an OG:FA mispair ($K_d = 90 \pm 10$ nM). Thus, removal of the C-terminal domain has resulted in an 8-fold decrease in affinity for OG:M in contrast to the greater than 700-fold decrease in affinity for the OG:FA duplex. Furthermore, Stop 225 exhibits affinities similar to those of the WT enzyme for M and Z opposite G. In contrast to the results with Z and FA, Stop 225 retains preference for OG over G opposite M. Taken together, these results are consistent with the role proposed for the C-terminal domain in the base flipping process based on the decrease in the pre-steady-state glycosylase rate observed for Stop 225 with OG:A mispairs.[16,26] The presence of the C-terminal domain may not be as critical in the recognition of base pairs that are more readily extruded from the duplex, such as the M analogue. Furthermore, this suggests a novel recognition mechanism for MutY where both bases in the damaged and mismatched base pair are recognized, possibly by expulsion of both bases from the DNA duplex.

To determine if the unexpectedly high affinity of MutY for the OG:M-containing duplex correlates with a decreased stability of the duplex, melting temperatures (T_m) were measured and compared with duplexes containing a central OG:A or OG:Z mispair (Table 2).

Table 2. UV Melting Temperatures (T_m) of 11-Mer Duplexes of the Sequence d(5′-GAGCT**OG**GTGGC-3)•d(3′-CTCGA**X**CACCG-5′) (3−10 μM Duplex DNA Was Used in a Buffer Containing 10 mM $KHPO_4$, 1 M NaCl, and 1 mM EDTA)

X	T_m (°C)
A	55.9 ± 0.5
Z	49.6 ± 0.8
M	43.4 ± 0.4

Indeed, the stability of the duplexes containing both the M and Z analogues has decreased relative to the OG:A mispair. The destabilization of the duplex provided by the introduction of M may facilitate a conformational change of the DNA necessary for efficient binding before catalysis. In addition, Stop 225, which lacks the C-terminal domain, can still bind OG/G:M duplexes with a greater affinity than other adenine analogues which do not significantly affect base pairing stabilization of the duplex. This trend is also reflected in the data of the Z analogue. In the case of the WT enzyme, this analogue opposite OG and G is one of the least efficiently recognized; however, the affinity of Stop 225 for Z opposite OG and G is considerably better than that of the FA analogue. Thus, duplex destabilization may be a factor leading to the improved recognition of this analogue relative to FA with Stop 225. These results are consistent with a mechanism that utilizes the C-terminal domain for recognition and involves disruption of the duplex near the base pair.

Designed hydrophobic DNA bases and base analogues incapable of hydrogen bonding have been used to study the

molecular details of a variety of biological processes,[30−32] in addition to adding diversity to coding properties of DNA.[33] For example, Kool and co-workers have investigated hydrogen bonding, steric requirements, and minor groove interactions of DNA polymerase using hydrophobic base analogues.[34−36] These studies have shown that non-hydrogen-bonding hydrophobic isosteres of dATP and dTTP are effectively inserted into DNA with high fidelity by DNA polymerases, suggesting that shape and steric considerations may be more important than hydrogen-bond formation for polymerase fidelity. The work presented herein describes a novel use for hydrophobic bases, as illustrated with 4-methylindole, to delineate factors influencing DNA damage recognition and repair. Our results show that high-affinity binding of MutY can be obtained with M which is incapable of hydrogen bonding to OG. This suggests that MutY may not require specific contacts with the adenine base for efficient recognition of the substrate; however, such contacts may be necessary for catalysis of adenine removal. In addition, these results show that the presence of OG may be more important for the initial substrate recognition event preceding catalysis. Furthermore, this work provides additional evidence consistent with the use of a nucleotide flipping mechanism for substrate recognition by MutY and the participation of the C-terminal domain in this process.

Acknowledgment. This work was supported by NIH Grant CA 67985 and the Chemistry Department of the University of Utah. S.S.D. is an A. P. Sloan Research Fellow (1998−2000). We thank Eric Kool for providing the inspiration for this work during his visit to the University of Utah.

Supporting Information Available: Protein overexpression, purification, and experimental details of the dissociation constant (K_d) determinations and UV melting temperature (T_m) determinations. This material is available free of charge via the Internet at http://www.pubs.acs.org.

OL005831O

(30) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473−1487.
(31) Matray, T. J. and Kool.; E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191−6192.
(32) Guckian, K. M.; Morales, J. C.; Kool, E. T. *J. Org. Chem.* **1998**, *63*, 9652−9656.
(33) McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585−11586.
(34) Morales, J. C.; Kool, E. T. *Nature Struct. Biol.* **1998**, *5*, 950−954.
(35) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704−708.
(36) Morales, J. C.; Kool, E. T. *J. Am. Chem. Soc.* **1999**, *121*, 2323.