Crystal structure of a thwarted mismatch glycosylase DNA repair complex

Tracey E.Barrett[1,2], Orlando D.Schärer[3,4], Renos Savva[1,5], Tom Brown[6], Josef Jiricny[7], Gregory L.Verdine[3] and Laurence H.Pearl[1,2,8,9]

[1]Department of Biochemistry and Molecular Biology, and [8]Joint UCL/LICR Crystallography Laboratory, University College London, Gower Street, London WC1E 6BT, [6]Department of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ, UK, [3]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA and [7]Institute for Medical Radiobiology, August Forel-Strasse 7, 8029 Zürich, Switzerland

[2]Present address: Centre for Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, UK
[4]Present address: Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands
[5]Present address: Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, UK

[9]Corresponding author
e-mail: l.pearl@biochem.ucl.ac.uk

The bacterial mismatch-specific uracil-DNA glycosylase (MUG) and eukaryotic thymine-DNA glycosylase (TDG) enzymes form a homologous family of DNA glycosylases that initiate base-excision repair of G:U/T mismatches. Despite low sequence homology, the MUG/TDG enzymes are structurally related to the uracil-DNA glycosylase enzymes, but have a very different mechanism for substrate recognition. We have now determined the crystal structure of the *Escherichia coli* MUG enzyme complexed with an oligonucleotide containing a non-hydrolysable deoxyuridine analogue mismatched with guanine, providing the first structure of an intact substrate-nucleotide productively bound to a hydrolytic DNA glycosylase. The structure of this complex explains the preference for G:U over G:T mispairs, and reveals an essentially non-specific pyrimidine-binding pocket that allows MUG/TDG enzymes to excise the alkylated base, 3,N^4-ethenocytosine. Together with structures for the free enzyme and for an abasic-DNA product complex, the MUG–substrate analogue complex reveals the conformational changes accompanying the catalytic cycle of substrate binding, base excision and product release.
Keywords: DNA repair/mismatch DNA glycosylase/structure

Introduction

The repair of many altered or mismatched bases in DNA proceeds via a base-excision repair pathway initiated by cleavage of the N-glycosidic bond linking the damaged or inappropriate base to the deoxyribose sugar (Seeberg

et al., 1995). The enzymes that achieve this step are more or less lesion specific, but mechanistically fall into one of two classes (Dodson *et al.*, 1994). Bifunctional DNA glycosylases displace the aberrant base by nucleophilic attack of a protein amino group on the C1′ carbon of the nucleotide, forming a covalent Schiff's base enzyme–DNA intermediate. This undergoes β-elimination and hydrolysis resulting in cleavage of the 3′-phosphodiester and release of the deoxyribose as a ring-opened α-β-unsaturated aldehyde (Dodson *et al.*, 1994). Monofunctional DNA glycosylases hydrolyse the N-glycosidic bond directly, leaving an intact phosphodiester backbone and an abasic C1′-hydroxy-2′-deoxyribose sugar ring, which is subsequently excised by an AP-endonuclease (Dianov and Lindahl, 1994).

Mutagenesis of a key catalytic residue in T4 endo V pyrimidine photodimer DNA glycosylase (Doi *et al.*, 1992) permitted structure determination of a complex between T4 endo V and a pyrimidine photodimer DNA substrate (Vassylyev *et al.*, 1995), providing the first direct view of DNA lesion recognition by a repair enzyme. Similar mutagenic approaches to the uracil-DNA glycosylases (Mol *et al.*, 1995; Panayotou *et al.*, 1998) produced mutants capable of binding but not hydrolysing uracil-DNA substrates. However, these mutants retain sufficient residual activity at crystallographic concentrations that only structures of enzyme–product complexes have so far been obtained for this class of enzymes (Slupphaug *et al.*, 1996; Barrett *et al.*, 1998b; Parikh *et al.*, 1998).

Chemical modification of the DNA offers an alternative, but more technically demanding, strategy for obtaining trapped repair enzyme–DNA complexes. An abasic pyrrolidine nucleotide (Schärer *et al.*, 1995, 1998) incorporated into an oligonucleotide, has been used to obtain the crystal structure of a DNA base-excision-product analogue complexed to the human N3-methyladenine DNA glycosylase (Lau *et al.*, 1998). We have previously shown that substitution of fluorine for either or both of the 2′ hydrogens in deoxyuridine renders it resistant to base excision by the human mismatch-specific thymine DNA glycosylase (Schärer *et al.*, 1997). We have now taken advantage of this observation to prepare co-crystals of the homologous *Escherichia coli* G:U/T mismatch-specific DNA glycosylase (MUG) bound to an oligonucleotide duplex containing mispairs between deoxyguanidine and 2′-fluoro-deoxyuridine, in which the scissile nucleotide is productively bound, but the N-glycosidic bond remains uncleaved. The structure of this complex at 2.85 Å resolution provides the first direct view of an enzyme–substrate analogue complex for a hydrolytic DNA glycosylase, revealing the structural basis for differential activity to G:U and G:T mismatches, and explaining the ability of MUG/thymine-DNA glycosylase (TDG) enzymes to excise the alkylation product, 3,N^4-ethenocytosine (Hang
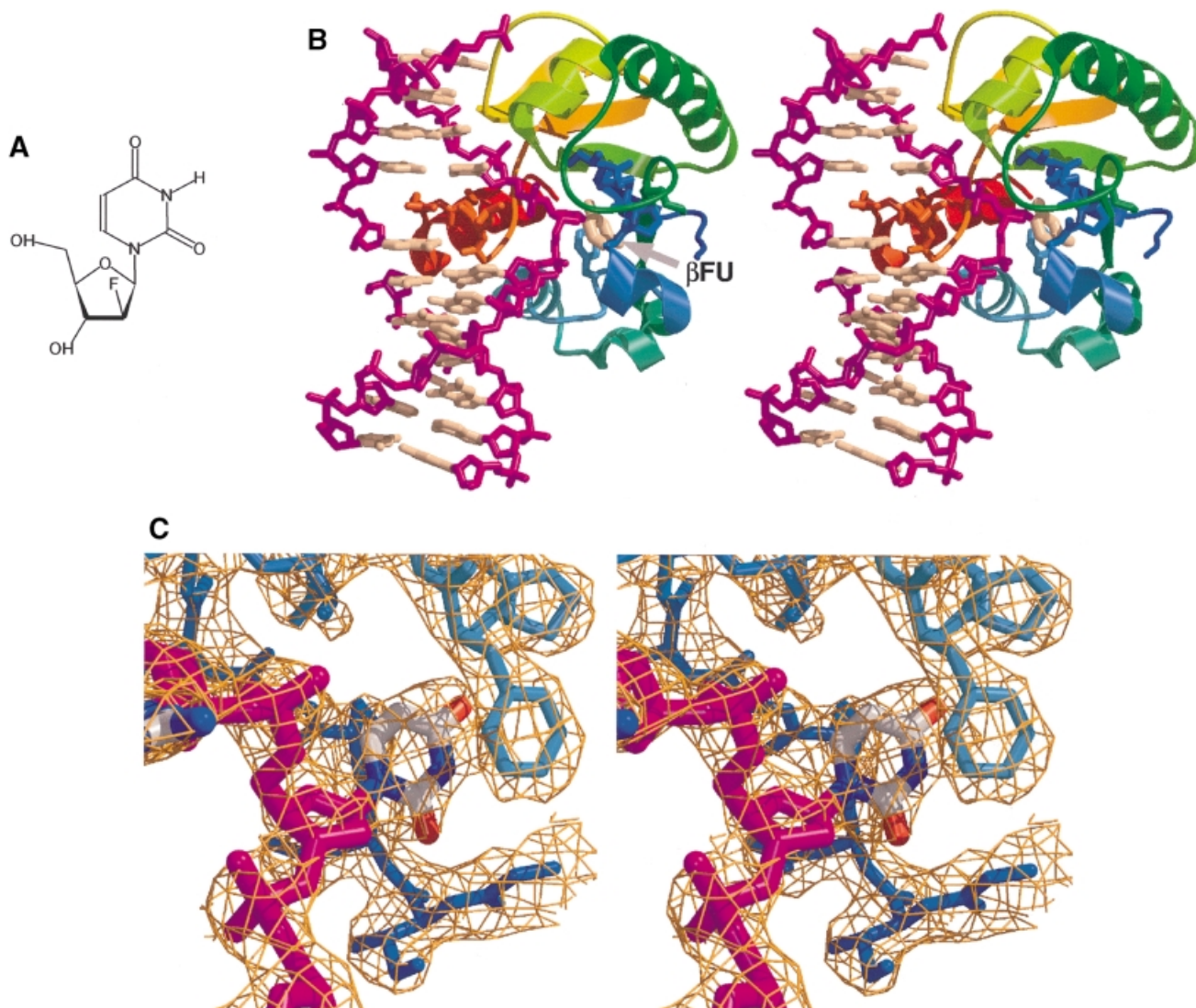
Fig. 1. Structure of MUG–βFU-DNA complex. (**A**) Chemical structure of the non-hydrolysable substrate analogue 1-(2′-deoxy-2′-fluoro-β-D-arabinofuranosyl)-uracil. (**B**) Stereo view of the MUG–βFU-DNA complex viewed towards the major groove of the DNA. The protein is shown as a secondary structure cartoon, coloured blue to red from the N- to C-terminus. (**C**) Electron density for the 'flipped-out' βFU nucleotide bound in the base-binding pocket. The electron density is from a $2F_o–F_c$ Fourier map, calculated with the omission of the βFU from the calculation, and contoured at 0.8 σ.

et al., 1998; Saparbaev and Laval, 1998). Together with structures for the free enzyme and for an enzyme–product complex (Barrett *et al.*, 1998b), this structure provides detailed insights into the subtle but important changes in the structure and conformation of the DNA and protein that accompany mismatch recognition and base excision by the MUG/TDG family of DNA glycosylases.

Results and discussion

Structure of MUG–DNA complex

The *E.coli* MUG was co-crystallized with an oligo-nucleotide: 5′-C-G-C-G-A-G-βFU-T-C-G-C-G-3′, where 'βFU' is 1-(2′-deoxy-2′-fluoro-β-D-arabinofuranosyl)-uracil (Figure 1A). The structure of the complex was determined by molecular replacement and refined at 2.85 Å resolution (see Materials and methods, and Table I). In principle, the oligonucleotide used in these co-crystals can form a blunt-ended self-complementary duplex with two

Table I. Crystallographic statistics

Data collection	Overall (outer shell)
R_{merge}	0.087 (0.278)
$I/\sigma (I)$	9.7 (2.9)
Completeness (%)	94.5 (96.7)
Multiplicity	2.3
No. unique reflections	5479

Structure refinement	
No. of atoms (protein + DNA)	1516
No. of atoms (solvent)	86
Resolution range (Å)	20.0–2.85
R_{cryst}	0.19
R_{free}	0.25

central G:βFU mismatched base pairs. However, as with the equivalent oligonucleotide containing deoxyuridine, formation of enzyme–DNA co-crystals is favoured by an alternative base pairing offset by six nucleotides (Barrett *et al.*, 1998b) (Figure 1B), giving a continuous nicked duplex running along the crystallographic ab 2-fold axis.

Previous studies of MUG (Barrett *et al.*, 1998b) identified a pocket in the enzyme, structurally similar to the uracil-binding pocket of 'classical' uracil-DNA glycosylases (UDGs) (Mol *et al.*, 1995; Savva *et al.*, 1995), and presumed to fulfil a similar role in providing a binding site for the base of a 'flipped-out' scissile nucleotide. Consistent with that, the abasic deoxyribose generated by base excision of a G:U mismatch lies close to the mouth of this pocket in a MUG–DNA excision product complex (Barrett *et al.*, 1998b). However, unlike DNA excision product complexes with UDG (Slupphaug *et al.*, 1996; Parikh *et al.*, 1998), where the cleaved base is retained, no electron density corresponding to the cleaved uracil was observed in the pocket in the MUG–DNA excision product complex (Barrett *et al.*, 1998b). In the co-crystals with the βFU oligonucleotide, difference maps calculated with phases from a model from which the βFU was omitted, show clear electron density for an intact nucleotide (Figure 1C). The βFU is bound with the base projecting into the predicted binding pocket, and packed between the aromatic ring of Phe30, and the side and main chains of Asn18 and Pro19. The electron density for the sugar ring is less well defined than that for the base, and no clear electron density is evident for the 2′ fluorine, suggesting a degree of pseudorotational disorder in the deoxyribose. The observed electron density and the orientation of the ring substituents are most consistent with the 2′-*endo* sugar pucker commonly observed in B-form DNA, which is also the preferred conformation of β, but not α, 2′-fluoro deoxypyrimidines (Lipnick and Fissekis, 1980).

Uracil/thymine discrimination

Unlike the UDGs, which are exquisitely specific for the excision of uracil, TDG will efficiently excise uracil or thymine but only from mismatched base pairs with guanine (Nedderman and Jiricny, 1994). The homologous *E.coli* MUG (Gallinari and Jiricny, 1996) preferentially excises uracil from G:U mismatches, and only excises thymine from G:T mismatches at very high enzyme concentrations (Barrett *et al.*, 1998b). In UDGs, rejection of thymine is conferred by a 'barrier' residue (Tyr90 in HSV-1 UDG), which packs against the 5′-position of bound uracil (Mol *et al.*, 1995; Savva *et al.*, 1995) and prevents binding of a 5′-methylated base. In the MUG/TDG family, the equivalent position is occupied by glycine, so that this major barrier to thymine is absent (Barrett *et al.*, 1998b). Nevertheless, the rate of G:T cleavage is slower than the rate of G:U cleavage for both enzymes, and in the case of MUG, significantly so, suggesting some other barrier to thymine binding, albeit of lower stringency than the tyrosine of UDGs. The position of the uracil base in the MUG–βFU DNA complex is such that the 5′-position is directed towards the side chain of Ser23, at the N-terminus of a short helix (20–27) present in the MUG (and by homology TDG) structure, but absent from UDGs (Mol *et al.*, 1995; Savva *et al.*, 1995). Modelling of the 5′-

methyl of thymine on to the observed uracil results in a steric clash with the side-chain hydroxyl of this residue, which would mitigate against thymine binding (Figure 2B). However, the side chain of Ser23 is free to rotate, and the helix itself shows a significant degree of plasticity (see below). Thus, unlike the tyrosine barrier in UDGs, which is rigidly supported by the core of the protein, crowding of the pyrimidine 5′-position by Ser23 would lower the efficiency of thymine excision, but not prevent it. Significantly, in mammalian TDGs, which excise uracil and thymine with more comparable efficiency (Waters and Swann, 1998), residue 23 is an alanine, whose smaller side chain presents much less steric interference with the 5′-methyl of thymine. The TDG homologue identifiable in *Schizosaccharomyces pombe* has serine rather than alanine at this position, and would be expected to display the greater preference for G:U over G:T of the bacterial MUG enzymes, although this remains to be tested.

The MUG pocket does not discriminate between cytosine and uracil

The Watson–Crick base-pairing edge of uracil in the βFU-DNA complex is directed towards the central β-sheet, which provides the bottom of the pocket. Both exocyclic carbonyls receive out-of-plane hydrogen bonds from peptide NH groups, directed to opposite faces of the ring (Figure 2C). The hydrogen bond from Ile17 to O2 provides no contribution to specificity, as cytosine also has an exocyclic O2, whereas in principle, the hydrogen bond from the NH of Phe30 to O4, would discriminate against the exocyclic N4-amino of a cytosine. However, the peptide NH of Phe30 is directed nearly perpendicularly to the plane of the base, so that it interacts favourably with the delocalized lone-pair of the amino-nitrogen rather than unfavourably with the amino-protons, and is thus unlikely to make a substantial contribution to rejection of cytosine. The exocyclic O4-carbonyl makes a further hydrogen bond to a water molecule in the bottom of the pocket, bound to the main-chain carbonyl of Thr67 and to a second water bound to the side chain of Lys68. Both these water molecules are also present in the free enzyme. These solvent interactions confer no specificity, however, and could equally well be made by the exocyclic N4-amino of cytosine. No interactions, direct or water-bridged, are made by the N3-imino of the ring. This relative lack of polar interactions with the bound base is in marked contrast to UDGs, which make up to five direct hydrogen bonds and one water-bridge.

The primary determinant of specificity for uracil over cytosine in UDGs, is a conserved asparagine (147 in HSV-1 UDG), which lies at the bottom of the nucleotide-binding pocket, and together with the peptide NH of Gln87, provides three in-plane hydrogen bonds with the Watson–Crick base-pairing groups of bound uracil (Mol *et al.*, 1995; Savva *et al.*, 1995). Structural alignment of the common core of MUG and HSV-1 UDG identified residue 68 in *E.coli* MUG as the topological equivalent of Asn147 in UDG. Surprisingly, this residue is a lysine in *E.coli* MUG and an alanine in *Serratia marcescens* MUG, neither of which could provide the specific hydrogen-bonding pattern afforded by asparagine in the uracil-binding pocket of UDGs. We had speculated (Barrett *et al.*, 1998b) that some selectivity for O4 of uracil/
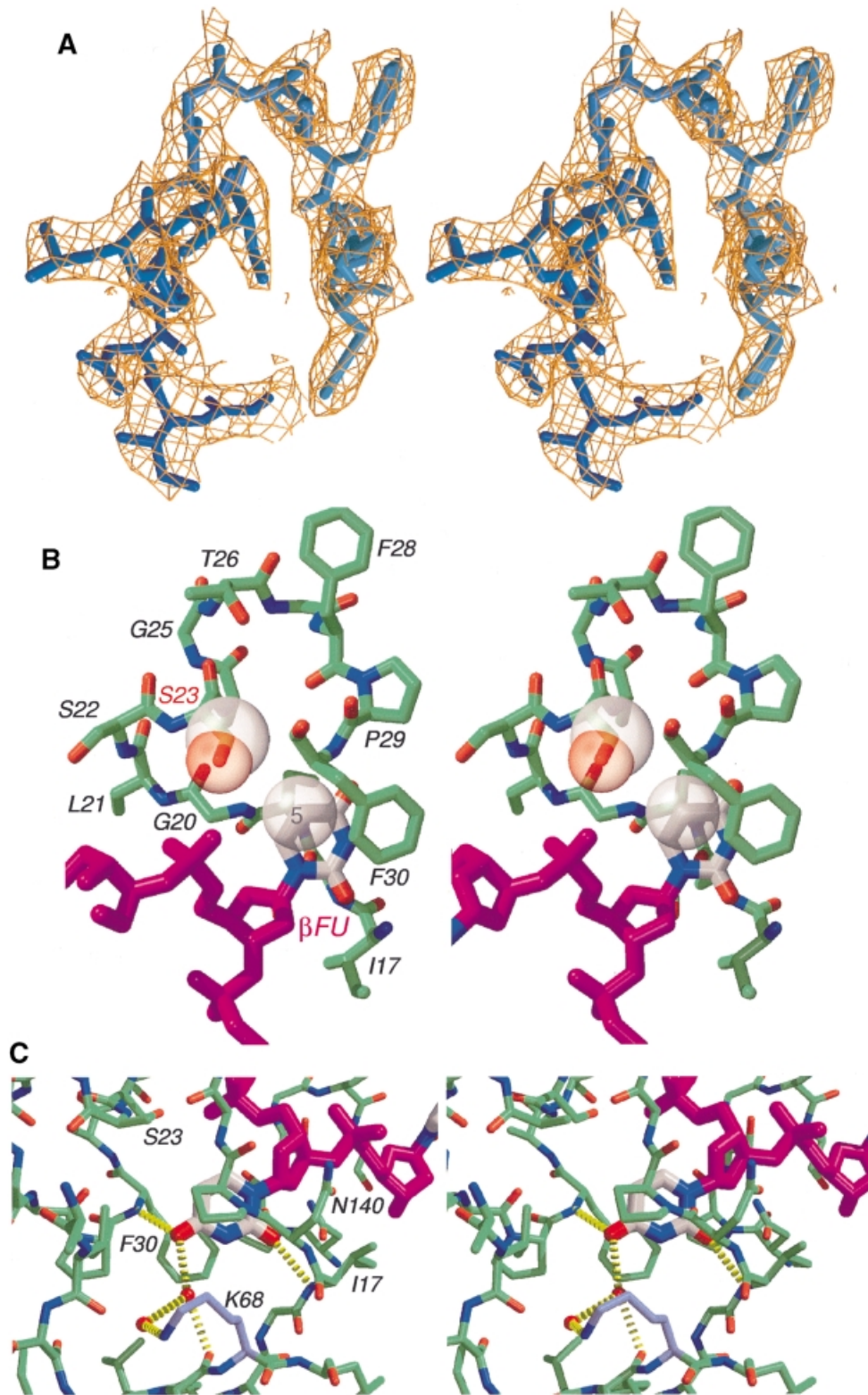
Fig. 2. Structure and interactions of the MUG base-binding pocket. (**A**) Electron density for the polypeptide chain from residues 17 to 30 which forms the upper part of the nucleotide-binding site. (**B**) Crowding of the C5-position of the bound uracil base by the side chain of Ser23 on the 20–27 helix, disfavours, but does not prevent, G:T excision. In TDGs, where G:T excision is efficient, the equivalent residue is alanine. (**C**) Hydrogen-bonding interactions made by uracil in the deep hydrophobic base-binding pocket of MUG. Lys68, previously suggested as a source of specificity, is highlighted.

thymine might be provided by rearrangement of the Lys68 side chain in a substrate complex, allowing the side-chain amino to hydrogen-bond to O4 of a bound uracil. In the substrate analogue complex presented here, it is clear that no such rearrangement of Lys68 occurs, and the side chain remains directed away from the bound nucleotide, involved
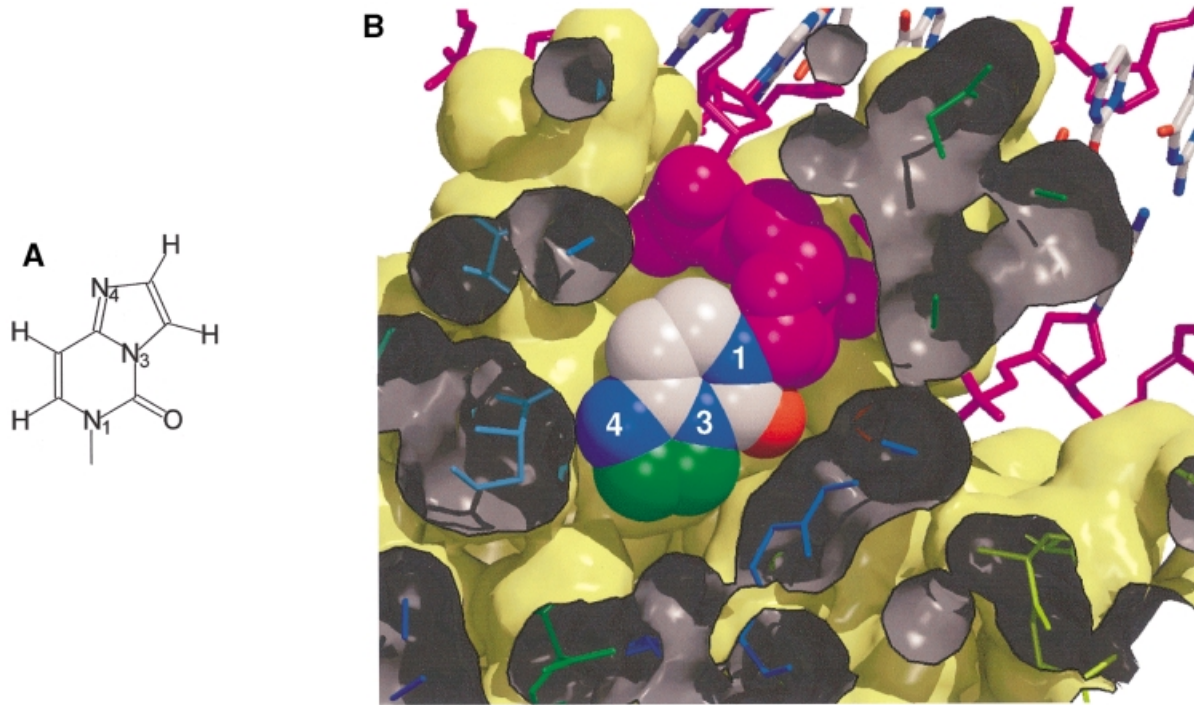
Fig. 3. Binding of 3,N^4-ethenocytosine (εC). (**A**) Chemical structure of εC. (**B**) Binding of εC in the pocket of MUG, modelled from the experimentally determined MUG–βFU-DNA structure. The etheno moiety bridging the imino N3 and exocyclic amino N4, is highlighted in green. Unlike uracil, 3,N^4-ethenocytosine entirely fills the available space, and is complementary to the pocket in both shape and chemistry.

in ion-pair interactions with the side chains of Asp4 and Asp71, at the surface of the protein. The equivalent residue to Lys68 of *E.coli* MUG, in the mammalian and yeast TDGs, is an asparagine, as in the UDGs. We had previously suggested that the TDGs, unlike the MUGs, might therefore be able to provide the specific hydrogen bonds with uracil observed in UDG complexes (Mol *et al.*, 1995; Savva *et al.*, 1995; Slupphaug *et al.*, 1996). However, the strand carrying the residue at this position in MUG is lower down in the pocket than, and adopts a different conformation from, its topological equivalent in UDGs. This places Lys68 in the wall of the pocket, rather than directly below the bound uracil, as is the asparagine in UDGs. Thus, an asparagine replacing Lys68, as in TDGs, has its amide head-group directed at an angle to the bound uracil and beyond hydrogen-bonding distance from any of its atoms. It seems likely that, contrary to our previous speculation, TDG, like MUG, does not provide specific hydrogen bonding able to distinguish uracil and thymine from cytosine. Instead, the environment presented to the bound base is overwhelmingly hydrophobic and essentially non-specific, in all of the MUG/TDG family. The ability of MUG/TDG enzymes to recognize and hydrolyse the pyrimidines in G:U and G:T mispairs, and their lack of activity against G:C, A:U or A:T base pairs, results from a combination of specificity for guanine on the complementary strand, and a weak 'push' component that can only 'flip-out' inherently unstable base pairs (Barrett *et al.*, 1998b).

Structural basis for excision of 3,N^4-ethenocytosine

3,N^4-ethenocytosine (εC) arises in DNA by the action of metabolites of environmental carcinogens such as vinyl

chloride and ethyl carbamate (Leithauser *et al.*, 1990; Nair *et al.*, 1995), and by endogenous products of lipid peroxidation (Chung *et al.*, 1996). In εC, the N4-amino and N3-imino groups of cytosine become cyclized into a stable imidazole ring, which completely disrupts the G:C base pair in which it arises (Figure 3A). Unlike comparable adducts of purines, εC is not a substrate for the mammalian (Hang *et al.*, 1996, 1997) or bacterial (Saparbaev *et al.*, 1995) alkylpurine DNA glycosylases, and the mechanism for its repair has been unclear. Recently, TDG and MUG have been identified as activities initiating base-excision repair of 3,N^4-ethenocytosine (εC) in mammalian cells (Hang *et al.*, 1998; Saparbaev and Laval, 1998) and *E.coli* (Saparbaev and Laval, 1998), respectively.

With the structure of the MUG–βFU complex, the ability of MUG/TDG enzymes to excise εC can be readily understood. When modelled from the bound uracil in the MUG–βFU substrate complex, the etheno moiety of εC is comfortably accommodated in the hydrophobic space at the bottom of the binding pocket, which is not occupied by bound uracil (Figure 3B). The etheno carbon attached to N3, is in van der Waals contact with the Cα of Gly16, which is conserved throughout the MUG/TDG family. None of the small differences in sequence between the MUG and TDG enzymes in this region would prevent the binding of εC in the position suggested by the MUG–βFU substrate complex.

Although MUG was characterized by its ability to excise uracil from G:U mismatches (Gallinari and Jiricny, 1996), it is some 50 times more active in the excision of εC from G:εC mispairs (Saparbaev and Laval, 1998). This increased efficiency reflects the relative ease with which the weakly hydrogen bonded and poorly stacked εC (Cullinan *et al.*, 1997) can be 'flipped-out' of a G:εC

mispair, and the favourable hydrophobic environment at the bottom of the MUG base-binding pocket in which it is then bound. The ability of MUG to select for G:U mispairs over G:C or A:U base pairs is a subtle balance between the energetically unfavourable disruption of the base pair, and the compensating interactions of the intercalation loop with the DNA, and the 'flipped-out' base with the binding pocket. With εC, which is incapable of making strong hydrogen bonds with any partner-base, or of stacking efficiently within the DNA duplex (Cullinan *et al.*, 1997; Korobka *et al.*, 1996), the energetic penalty to 'flipping-out' is greatly reduced whatever 'widowed' base is left. Burial of the flipped-out εC in the MUG base-binding pocket will be more favourable than burial of uracil, whose hydrogen-bonding requirements are not met by the pocket. The combination of a decreased penalty for base-pair disruption, and an enhanced binding energy for the flipped-out base effectively removes the mismatch specificity of MUG in εC excision, so that A:εC is excised at half the rate of G:εC, whereas excision of A:U is several orders of magnitude slower than excision of G:U. Nonetheless the requirement for second-strand interaction remains, and no activity is observed against εC in ssDNA (Saparbaev and Laval, 1998).

Conformational changes in the MUG reaction cycle

In previous studies (Barrett *et al.*, 1998b), we determined structures of the free MUG enzyme, and of a MUG complex with a base-excision reaction product. With the structure of the MUG–substrate analogue complex presented here, we are now able to follow the changes in the conformation of the enzyme and DNA as they progress through the reaction cycle of substrate binding, base excision and substrate release.

The most dramatic conformational change on formation of the enzyme–substrate-analogue complex occurs in the DNA, due to the flipping-out of the βFU nucleotide, leaving a widowed guanine on the distal strand. Nucleotide flipping has now been observed directly in three different classes of hydrolytic DNA glycosylases (Slupphaug *et al.*, 1996; Barrett *et al.*, 1998b; Lau *et al.*, 1998). In all cases, the space vacated in the base-stack by the flipped-out nucleotide is occupied by groups from the enzyme that make compensatory interactions with the 'proximal' strand containing the flipped-out nucleotide and thereby stabilize its extrahelical conformation. In MUG, the loop from 141–146 penetrates the body of the double helix, and the space left by the flipped-out nucleotide on the proximal strand becomes occupied by the main and side chains of Ser142, Gly143, Leu144 and Ser145 (Barrett *et al.*, 1998b). However, this process is qualitatively different in MUG, in that enzyme insertion occurs in both strands of the DNA. One consequence of this is that the bending of the DNA away from the enzyme observed in other systems (Slupphaug *et al.*, 1996; Lau *et al.*, 1998; Parikh *et al.*, 1998), is much less pronounced in MUG–DNA complexes. Instead, the phosphate backbone between the widowed guanine and the nucleotide on its 5′ side on the 'distal' strand of the DNA, adopts an extended conformation, separating the parallel planes of the sequential bases by >7 Å, and allowing intercalation of the guanidinium head-group of Arg146. Penetration of the 141–146 loop into the DNA allows formation of hydrogen bonds between

the carbonyl oxygen of Gly143 and the N1-imino and N2 exocyclic amino of the 'widowed' guanine, which we have previously suggested as a significant component of specificity for guanine in G:U/T mismatches (Barrett *et al.*, 1998b).

Significant movements occur in the short helix formed by residues 20–27 in MUG, on binding the βFU-DNA. In the free enzyme the side chain of Ser22 is directed into solvent, while the side chain of Ser23 is hydrogen bonded across the top of the pocket, to the peptide NH of Ala31. On binding the DNA substrate, the N-terminal turn of this helix rotates out from the body of the protein, pivoted at glycines 20 and 25, moving serines 22 and 23 by 2.5–3.0 Å. The C-terminal turn of this helix remodels from an α to a 3_{10} helical conformation, with the peptide carbonyl of Ser23 shifting its hydrogen bond from the peptide NH of Gly27 to the NH of Thr26 (Figure 4A and B). The α-helical hydrogen bonds from the peptide carbonyls of Leu21 and Ser22 to the peptide NHs of Gly25 and Thr26 in the free enzyme, are disrupted in the remodelled 3_{10} helix in the DNA substrate complex, making no compensating interactions, so that the average temperature factor for this turn, but not the surrounding chain, nearly doubles from 21 to 41 Å^2. In its new position, the side chain of Ser22 provides a water-bridged interaction with the 5′ phosphate of the nucleotide on the 5′ side of the 'flipped-out' βFU. The side chain of Ser23 loses its hydrogen bond to the NH of Ala31 and packs against C5 of the uracil ring in the complex. Apart from the remodelled helix, binding of the βFU-DNA causes only small movements and an overall decrease in temperature factors in the polypeptide encompassing the intercalation loop 141–150, and in the loop from 73 to 80, which is involved in interactions with the phosphate backbone. Overall, the r.m.s. deviation between free MUG and MUG complexed with βFU-DNA is 0.519 Å.

Base excision and release

Hydrolysis of the N-glycosidic bond will convert a bound substrate into the reaction products, free uracil base and a DNA molecule with an abasic deoxyribose. In crystallographic studies of human UDG complexes with uracil-DNA oligonucleotides, where base excision occurs in the crystallization experiments (Slupphaug *et al.*, 1996; Parikh *et al.*, 1998), the excised base and the abasic-DNA are both retained in the complex. In the complex obtained by co-crystallization of MUG with an oligonucleotide containing a G:U mismatch (Barrett *et al.*, 1998b) the abasic-DNA product remains bound to the enzyme, but no uracil was visible. We cannot dismiss the possibility that the MUG–product complex is formed by rebinding to a product released after base excision by a different MUG molecule. Indeed such an explanation has been presented for an abasic-product DNA complex obtained with a human UDG mutant (Parikh *et al.*, 1998). In UDGs, the β-strands providing the groups forming the bottom of the uracil-binding pocket are embedded within the hydrophobic core of the protein, so that the uracil-binding site is completely closed at the bottom. Thus, the only exit for an excised uracil is through the open mouth of the pocket, which is completely occluded by the 'flipped-out' segment of the proximal strand of the DNA duplex, in UDG–product complexes (Slupphaug *et al.*, 1996;
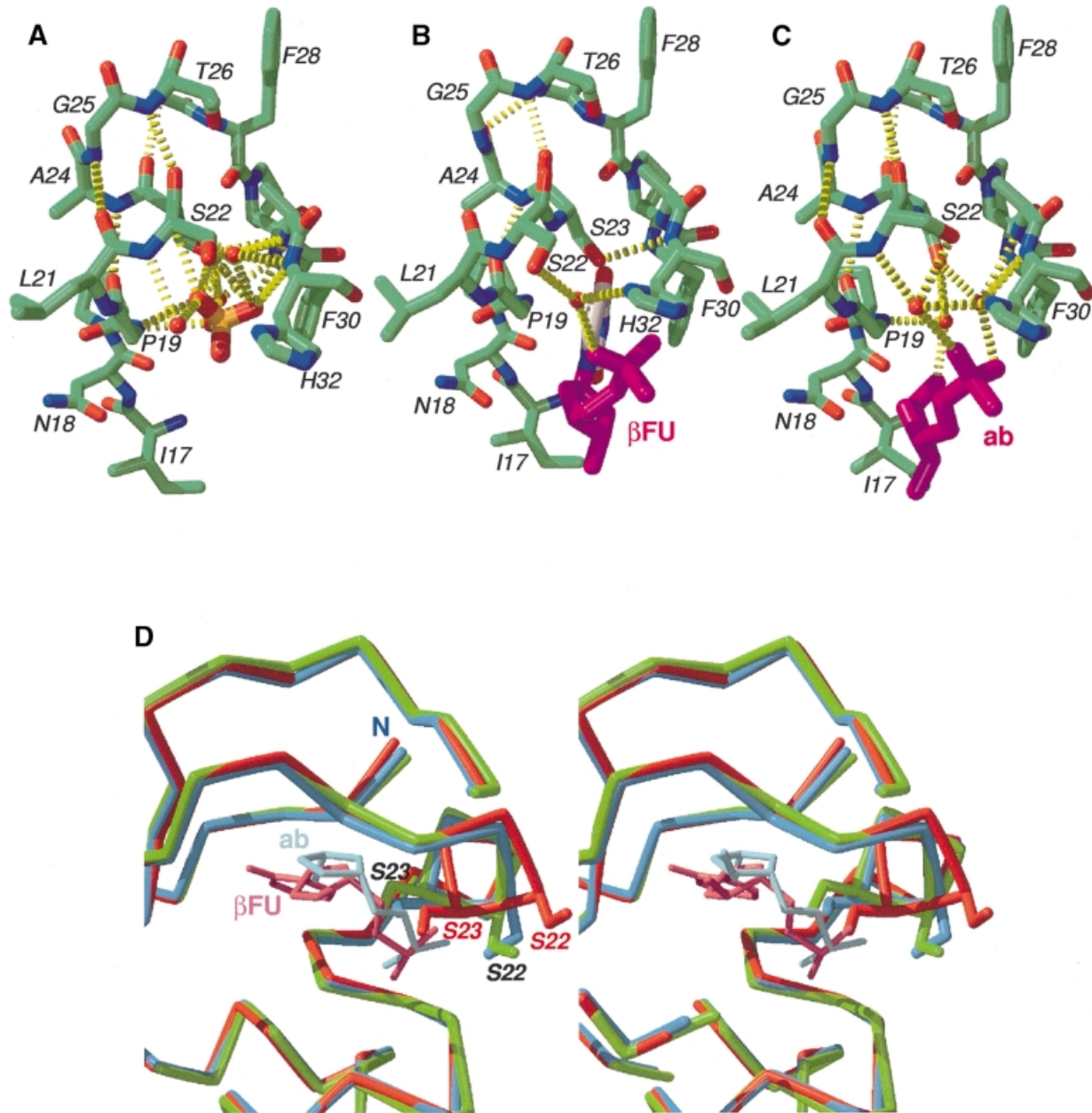
Fig. 4. Remodelling of the 20–27 helix on nucleotide binding. (**A**) Conformation and hydrogen-bonding pattern of the 'plastic' helix formed by residues 20–27, in the free MUG enzyme. A sulfate ion is bound at the N-terminus of the helix; (**B**) as (A) but in the MUG–βFU-DNA complex; (**C**) as (B) but in the MUG–abasic-DNA complex. (**D**) Stereo-pair of superposition of free enzyme (yellow), MUG–βFU-DNA complex (red) and MUG–abasic-DNA complex (blue), showing the motion of the 20–27 helix during the reaction cycle.

Parikh *et al.*, 1998). Release of excised uracil base from UDG therefore requires prior dissociation of the abasic-DNA product.

Although the mouth of the binding pocket in MUG is similar to that in UDG, and is also occluded by the proximal DNA strand in both substrate analogue and product complexes, the pocket itself is much deeper and offers the tantalizing possibility of an alternative exit for the release of an excised base without the requirement for dissociation of the abasic-DNA. The bottom of the pocket in MUG is provided by the three N-terminal residues of the protein, which lead into the edge strand of the β-α-β fold, and are fully exposed at the surface of the protein. This segment does not entirely close off the bottom of the pocket, so that a small hole connects to the surface

(Figure 5A). While the dimensions of this hole in the observed complexes are insufficient to allow passage of an excised base, the four N-terminal residues of this strand are not hydrogen bonded into the central β-sheet, and are held in place only by the packing of Val2 against Ala24 and Gly27 in the mobile helix formed by residues 20–27 (see above), and by hydrogen bonds from the carbonyl of Val2 and the side chain of Asp4 to the side chains of lysines 68 and 95, respectively. These interactions would allow some flexibility in the position of this strand, conceivably coupled to the observed motions in the 20–27 helix, so that the hole could be sufficiently enlarged, at least transiently, to allow the exit of an excised base through the bottom of the pocket. In the *S.marcescens* MUG sequence this N-terminal strand begins two residues
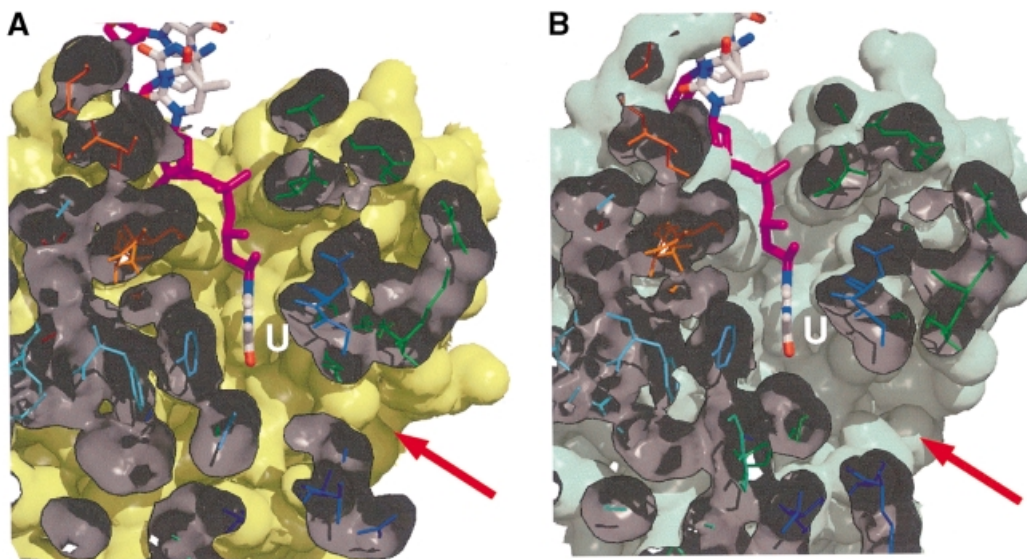
Fig. 5. MUG base-binding pocket is open at both ends. (**A**) Surface contour of the base-binding pocket in *E.coli* MUG, transected to show the continuous tunnel connecting the opposite faces of the protein. (**B**) As (A), but calculated from the 63% identical *S.marcescens* MUG sequence modelled on to the *E.coli* MUG structure using SWISSMODEL (Guex and Peitsch, 1997). The putative exit-hole in *S.marcescens* MUG is substantially enlarged due to the absence of two N-terminal residues relative to *E.coli* MUG.
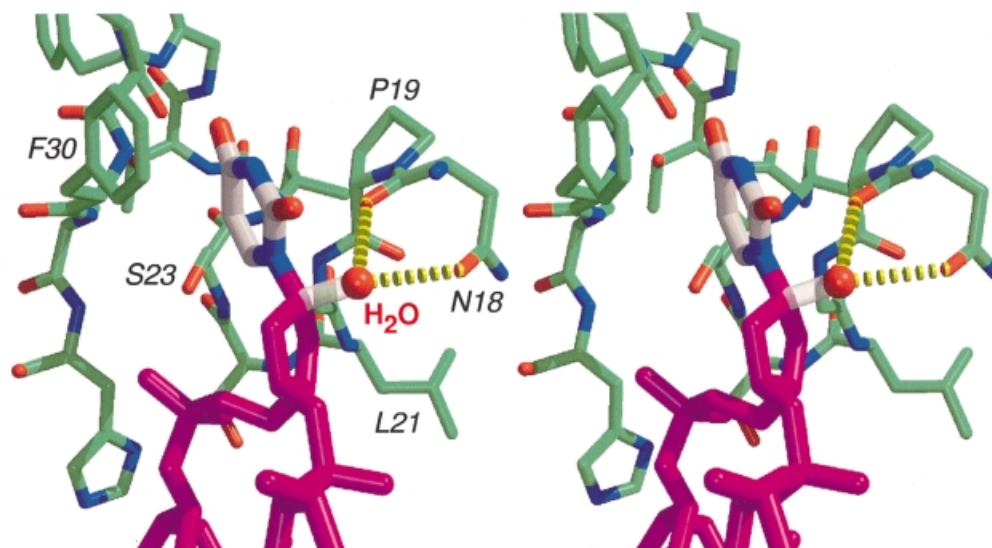

Fig. 6. Nucleophilic attack by bound water. The putative nucleophilic water molecule observed in the structure of free MUG bound to the side chain and main chain of Asn18, was superimposed in the same position relative to the protein in the structure of MUG with the bound βFU substrate analogue to model the stage of the reaction immediately prior to formation of the transition state. The water thus positioned is ~1.7 Å from C1′ of the nucleotide, and optimally directed for an in-line nucleophilic attack on C1′.

later, so that the hole would already be of sufficient size to allow release of a base, without need for conformational change (Figure 5B). Thus, in principle the MUG enzymes appear to provide a 'back-door' escape route for excised bases, independent of the release of the abasic-DNA. Further work will be required to test this hypothesis.

The transition from the βFU-DNA substrate analogue complex to the abasic product complex, which corresponds to excision and possible release of uracil, produces no significant changes in the conformation of the DNA overall. However, there are small changes in the position of the 'flipped-out' deoxyribose, where removal of the base and inversion to the β-anomer causes a shift of ~1 Å out of the pocket, and a rotation towards the catalytic residue Asn18. Concomitant with the movement of the

scissile nucleotide, the $\alpha/3_{10}$ helix at residues 20–27 'clicks' back into the fully hydrogen-bonded α-helical conformation adopted by this helix in the free enzyme (Figure 4C). Restitution of the α-helix moves Ser22 ~2 Å back into the body of the protein, breaking its water-bridged interaction with the 5′ phosphate of the nucleotide on the 5′ side of the βFU in the substrate analogue complex, and forming a new water-bridge with the 5′ phosphate of the abasic product nucleotide itself. Ser23 moves ~1.7 Å down into the space previously occupied by the uracil in the βFU-DNA complex, providing a second water-bridge to the 5′ phosphate of the abasic nucleotide. The increased set of interactions made by this phosphate in the product complex produces a substantial decrease in its temperature factor relative to the βFU-

substrate complex, whereas temperature factors of the adjacent phosphates on the same strand, whose environments are essentially unchanged by loss of the base, show much smaller decreases.

The space 'vacated' by the uracil on excision, is re-occupied by solvent molecules in the product complex. One water lies close to the position previously occupied by O2 of uracil, and bridges the O1′ hydroxyl to the peptide NH of Ile17 and to the side chain of Asn140. A water molecule is also bound at this position in the free enzyme. A second water, unique to the product complex, lies close to the position occupied by C5 of the uracil ring in the substrate complex, and bridges the O1′ hydroxyl of the abasic deoxyribose to the side chain of Ser23 and the peptide NH of Gly20.

Abasic product release

The final stage of the reaction cycle involves the release of the abasic-DNA product. Kinetic analysis (Waters and Swann, 1998) and binding studies (Barrett et al., 1998b) of TDG and MUG have independently demonstrated high affinity for abasic-DNA reaction products, so that product release is the rate-limiting step in substrate turnover by these enzymes. With UDGs the situation is less clear, as the catalytic domain of HSV-1 UDG displays no significant affinity for abasic sites in surface plasmon resonance (SPR) experiments (Panayotou et al., 1998), whereas human UDG displays nanomolar affinity for abasic sites when measured by SPR (Parikh et al., 1998). Whether these results represent real differences in mechanism between UDGs is uncertain, as the rate measured for product release by human UDG (0.2 s^{-1}) is difficult to reconcile with rates exceeding 40 s^{-1} for substrate turnover by the same enzyme (Kavli et al., 1996). The situation is further complicated by the observation that different forms of human UDG varying in regions outside the catalytic core, do not show high affinity for abasic sites (Bharita et al., 1998).

Release of the abasic-DNA product causes few structural changes in the enzyme, apart from a small relaxation of those segments including the intercalation loop 141–150, and the loop from 73 to 80, which are involved in direct DNA interactions in the product complex, which experience a small increase in temperature factors. Although it makes no significant movements on abasic-product release, the 'plastic' helix at residues 20–27 becomes markedly better ordered with the departure of the DNA, and interacts with a substantial network of solvent molecules that fill the empty DNA-binding site and pocket. During the reaction cycle this helix behaves like a helical spring, becoming tensed into the relatively unstable 3_{10} conformation when a base is bound in the pocket, and relaxing back to the stable well-ordered α-conformation on base excision and product release (Figure 4D).

Catalytic mechanism of MUG/TDG DNA glycosylases

On the basis of crystal structures for human (Mol et al., 1995) and herpes simplex virus (Savva et al., 1995) two very different catalytic mechanisms for the catalysed hydrolysis of the N-glycosidic bond were proposed. The mechanism of Mol et al. (1995) proposed an in-line displacement of the uracil, by nucleophilic attack of the imidazole of a conserved histidine residue on the C1′ of the deoxyribose, yielding a covalent enzyme–DNA intermediate. Subsequent attack by water or hydroxide would displace the histidine, releasing the abasic 1′-hydroxy-2′-deoxyribose product. This mechanism, which combines aspects of DNA lyase (Dodson et al., 1994) and β-glycohydrolase (McCarter and Withers, 1994) mechanisms, is attractive in that the consecutive in-line displacement reactions produce an abasic sugar with the β-anomeric configuration observed for base-excision products, once released from the enzyme active site (Barrett et al., 1998a). However, doubt is cast upon this mechanism by subsequent structures of human UDG complexes with base-excision repair products (Slupphaug et al., 1996; Parikh et al., 1998), which clearly show the bound abasic sugar to be in the α-anomeric configuration, and the proposed catalytic histidine to be on the wrong face of the sugar ring to achieve an in-line displacement of the base.

The mechanism of Savva et al. (1995) proposes a nucleophilic attack on the C1′ of the deoxyribose by a water or hydroxide ion bound to and activated by a conserved aspartate. The same conserved histidine invoked in the covalent mechanism (Mol et al., 1995) is proposed to act as a general acid, protonating the O2 of uracil and thereby destabilizing the N-glycosidic bond and facilitating in-line displacement of the base by the nucleophile. This mechanism generates an initial abasic deoxyribose product in the α-anomer, as is observed experimentally, and the proposed role for the conserved histidine is confirmed by observation of a hydrogen bond from the imidazole ring to the O2 of the cleaved, but retained uracil base (Slupphaug et al., 1996; Parikh et al., 1998). The substantial loss of hydrolytic activity that results from mutation of the catalytic aspartate to asparagine (Slupphaug et al., 1996; Panayotou et al., 1998) lends further support for this mechanism, which has become generally accepted (Krokan et al., 1997).

Structural homology between MUG/TDGs and UDGs (Barrett et al., 1998b), particularly around the mouth of the base-binding pocket, suggests that the catalytic mechanisms will be at least related. However, the conserved catalytic aspartate and histidine of UDGs become asparagines (positions 18 and 140), so the concerted general acid/base mechanism of UDGs (Savva et al., 1995) cannot simply apply. Asn140 in MUG is clearly unable to fulfil the role of the conserved histidine in UDGs, which protonates (or at least hydrogen bonds to) the O2 of bound uracil, improving the leaving-group properties of the base (Savva et al., 1995). Unlike its UDG counterpart, Asn140 does not move at any stage of the base-excision cycle, and is >4 Å from the O2 of the uracil in the βFU-DNA complex. Nor is Asn140 conserved in the MUG/TDG family, being methionine in mammalian TDGs and glycine in S.pombe, making its involvement in catalysis very unlikely. Although unable to act as a general base like its UDG aspartate counterpart, Asn18 in MUG, which is conserved in the known MUG/TDG sequences, nonetheless binds and presents a water molecule in a similar position and orientation to the catalytic aspartate of UDGs (Barrett et al., 1998b). Mutants of human UDG in which the catalytic aspartate is mutated to asparagine,

are still capable of performing base excision under crystallization conditions (Slupphaug *et al.*, 1996; Parikh *et al.*, 1998). This active-site asparagine is conserved in all the known MUG/TDG sequences, and its mutation to alanine (which cannot bind a water molecule) in TDG abolishes its hydrolytic activity (J.Jiricny, unpublished results). The recently described SMUG (Haushalter *et al.*, 1999) and TMUDG (Sandigursky and Franklin, 1999) uracil-DNA glycosylase families appear to represent intermediates between UDG and MUG/TDG in terms of their catalytic apparatus. Thus, SMUGs and TMUDGs conserve the general-acid histidine of UDGs which is absent in MUG/TDGs, but replace the general-base aspartate of UDGs with the asparagine of MUG/TDGs in SMUGs, and alanine or glycine in TMUDGs.

When the structures of the MUG–βFU-DNA complex and the free enzyme are superimposed, the water molecule bound between the side-chain amide and peptide carbonyl of Asn18 in free MUG (Barrett *et al.*, 1998b) is ~1.7 Å from the C1′ of the βFU and perfectly aligned for nucleophilic attack on the deoxyribose, and in-line displacement of the base from the opposite side of ring (Figure 6). The conformation of the nucleotide facilitates such an attack, with the glycosidic bond oriented antiperiplanar with respect to one of the lone pairs on the deoxyribose ring oxygen. This orientation allows for a favourable orbital overlap between the lone pair electrons and the nearly vacant *p*-orbital developed on C1′ in the transition state of the base-excision reaction. The presence of the 2′-fluoro group in βFU destabilize this transition state by withdrawing further electron density from the already electron-poor C1′ and renders the glycosidic bond resistant to hydrolysis. Thus, although MUG/TDG enzymes lack the powerful concerted general acid/base mechanism that make UDGs the fastest DNA glycosylases yet characterized, they retain the essential enzymic characteristics of presenting the reactants in high local concentration and favourable orientation.

Conclusion

The MUG–βFU complex presented here, has provided a clear understanding of the unusual properties of the MUG base-binding pocket, revealing the lack of specific interactions afforded to 'natural' pyrimidines, and providing an unambiguous structural explanation for the ability of MUG and TDG enzymes to excise 3,N^4-ethenocytosine. The MUG/TDG enzymes incorporate a remarkable multifunctionality, providing mismatch and alkylation damage glycosylase activity within the same enzyme. Which of these activities represents the response to the selective pressure under which MUG/TDG enzymes evolved, is uncertain, as neither activity would be a negative factor in an evolution driven by a requirement for the other. At least in *E.coli*, which has an efficient UDG enzyme for uracil excision in general, repair of εC may well be the most important role of MUG, in the absence of any other εC-DNA glycosylase activity. However, MUG/TDGs are by no means ubiquitous, and no MUG/TDG homologues have been identified so far in genomes of bacteria other than *E.coli* and *S.marcescens*, in any archaeal genomes or in the genomes of *Saccharomyces cerevisiae* and *Caenorhabditis elegans*. Assuming these organisms are equally susceptible to this type of alkylation damage,

some other more ubiquitous εC-DNA glycosylase activity remains to be discovered.

Direct visualization of enzyme–substrate complexes is a perennial 'holy grail' in structural biology, but is by definition extraordinarily difficult to achieve. With DNA-repair enzymes, chemical modification of substrate DNA provides the most powerful means for obtaining detailed structural and mechanistic insights into the recognition and repair of DNA lesions, without the inevitable structural artefacts generated by mutation of catalytic residues. The virtually isosteric replacement of a single atom (H→F) in the βFU-nucleotide employed in this study, represents the limit of what can be achieved in minimizing structural perturbation of the enzyme–substrate complex, while still preventing turnover.

Materials and methods

DNA synthesis

1-(2′-deoxy-2′-fluoro-β-D-arabonofuranosyl)-uracil (βFU) nucleoside was prepared according to published procedures (Howell *et al.*, 1988). The nucleoside was converted to the β-cyanoethyl-*N,N*′-diisopropyl-phosphoramidite derivative after protection of the 5′-hydroxyl as a dimethoxytrityl ether, as previously described (Schärer *et al.*, 1997).

Oligonucleotides were synthesized on an Applied Biosystems 394 DNA synthesizer using cyanoethyl phosphoramidite chemistry, deprotected in concentrated aqueous ammonia for 8 h at 55°C, and purified by reverse-phase HPLC. After purification, the major product was evaporated to dryness and desalted using a Pharmacia NAP 10 column (Sephadex G25), according to the manufacturer's instructions.

Crystallization, data collection and refinement

Escherichia coli MUG was expressed and purified, combined with the pre-annealed βFU oligonucleotide in a molar ratio of 1:2, and crystallized as previously described for other MUG–DNA complexes (Barrett *et al.*, 1998b). The MUG–βFU DNA complex crystallizes in space group P4$_2$2$_1$2 with unit cell dimensions a = 102.0 Å, c = 45.1 Å. Data to 2.85 Å were collected at 100 K from a single crystal of minimum dimension 30 μm on beam line 7.2 (λ = 1.488 Å) at the Synchrotron Radiation Source CLRC Daresbury Laboratory, and recorded on a MAR 345 Image Plate Detector. Diffraction images were integrated using DENZO and merged and reduced using SCALEPACK and other programs of the CCP4 program suite (CCP4, 1994). The structure was solved by molecular replacement with the structure of free MUG (Protein Databank code: 1MUG) and refined using REFMAC (CCP4, 1994) and X-PLOR (Brunger, 1992), with manual intervention using 'O' (Jones *et al.*, 1991). Statistics for the data collection and refinement are given in Table I. Coordinates will be deposited in the Protein Databank. Molecular graphics images were generated using Robert Esnouf's adaptation of MOLSCRIPT (Kraulis, 1991) and RASTER3D (Merrit and Murphy, 1994), with surfaces generated using SURFNET (Laskowski, 1995).

Acknowledgements

References

Barrett,T.E., Savva,R., Barlow,T., Brown,T., Jiricny,J. and Pearl,L.H. (1998a) Structure of a DNA base-excision product resembling a cisplatin inter-strand adduct. *Nature Struct. Biol.*, **5**, 697–701.

Barrett,T.E., Savva,R., Panayotou,G., Barlow,T., Brown,T., Jiricny,J. and Pearl,L.H. (1998b) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell*, **92**, 117–129.

Bharita,S., Krokan,H.E., Kristiansen,L., Otterlei,M. and Slupphaug,G. (1998) Human mitochondrial uracil-DNA glycosylase preform (UNG1) is processed to two forms one of which is resistant to inhibition by AP sites. *Nucleic Acids Res.*, **26**, 4953–4959.

Brunger,A. (1992) *X-PLOR Version 3.1. A system for X-ray Crystallography and NMR.* Yale University Press, New Haven, CT.

CCP4 (1994) Programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.

Chung,F.L., Chen,H.J.C. and Nath,R.G. (1996) Lipid peroxidation as a potential endogenous source for the formation of exocyclic DNA adducts. *Carcinogenesis*, **17**, 2105–2111.

Cullinan,D., Johnson,F., Grollman,A.P., Eisenberg,M. and de los Santos,C. (1997) Solution structure of a DNA duplex containing the exocyclic lesion $3,N^4$-etheno-2′-deoxycytidine opposite 2′-deoxyguanosine. *Biochemistry*, **36**, 11933–11943.

Dianov,G. and Lindahl,T. (1994) Reconstitution of the DNA-base excision-repair pathway. *Curr. Biol.*, **4**, 1069–1076.

Dodson,M.L., Michaels,M.L. and Lloyd,R.S. (1994) Unified catalytic mechanism for DNA glycosylases. *J. Biol. Chem.*, **269**, 32709–32712.

Doi,T., Recktenwald,A., Karaki,Y., Kikuchi,M., Morikawa,K., Ikehara,M., Inaoka,T., Hori,N. and Ohtsuka,E. (1992) Role of the basic amino acid cluster and Glu-23 in pyrimidine dimer glycosylase activity of T4 endonuclease V. *Proc. Natl Acad. Sci. USA*, **89**, 9420–9424.

Gallinari,P. and Jiricny,J. (1996) A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature*, **383**, 735–738.

Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.

Hang,B., Chenna,A., Rao,S. and Singer,B. (1996) $1,N^6$-ethenoadenine and $2,N^4$-ethenocytosine are excised by seperate human DNA glycosylases. *Carcinogenesis*, **17**, 155–157.

Hang,B., Singer,B., Margison,G.P. and Elder,R.H. (1997) Targetted disruption of alkylpurine-DNA-*N*-glycosylase in mice eliminates repair of $1,N^6$-ethenoadenine and hypoxanthine but not of 3,N4-ethenocytosine or 8-oxoguanine. *Proc. Natl Acad. Sci. USA*, **94**, 12869–12874.

Hang,B., Medina,M., Fraenkel-Conrat,H. and Singer,B. (1998) A 55-kDa protein isolated from human cells shows DNA glycosylase activity toward $3,N^4$-ethenocytosine and the G/T mismatch. *Proc. Natl Acad. Sci. USA*, **95**, 13561–13566.

Haushalter,K.A., Stukenberg,P.T., Kirschner,M.W. and Verdine,G.L. (1999) Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. *Curr. Biol.*, **9**, 174–185.

Howell,H.G., Brodfuehrer,P.R., Brundidge,S.P., Benigni,D.A. and Sapino,C.J. (1988) Antiviral nucleosides. A stereospecific total synthesis of 2′-fluoro-2′-deoxy-β-D-arabinofuranosyl nucleosides. *J. Org. Chem.*, **53**, 85–88.

Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Kavli,B., Slupphaug,G., Mol,C.D., Arvai,A.S., Petersen,S.B., Tainer,J.A. and Krokan,H.E. (1996) Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO J.*, **15**, 3442–3447.

Korobka,A., Cullinan,D., Cosman,M., Grollman,A.P., Patel,D.J., Eisenberg,M. and de los Santos,C. (1996) Solution structure on an oligonucleotide duplex containing the exocyclic lesion $3,N^4$-etheno-2′-deoxycytidine opposite 2′-deoxyadenosine, determined by NMR spectroscopy and restrained molecular dynamics. *Biochemistry*, **35**, 13310–13318.

Kraulis,P.J. (1991) MOLSCRIPT—A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Krokan,H.E., Standal,R. and Slupphaug,G. (1997) DNA glycosylases in the base excision repair of DNA. *Biochem. J.*, **325**, 1–16.

Laskowski,R.A. (1995) Surfnet—a program for visualizing molecular-surfaces, cavities and intermolecular interactions. *J. Mol. Graph.*, **13**, 323.

Lau,A.Y., Scharer,O.D., Samson,L., Verdine,G.L. and Ellenberger,T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: Mechanisms for nucleotide flipping and base excision. *Cell*, **95**, 249–258.

Leithauser,M.T., Liem,A., Stewart,B.C., Miller,E.C. and Miller,J.A. (1990) $1,N^6$-ethenoadenosine formation, mutagenicity and murine tumor-induction as indicators of the generation of an electrophilic epoxide metabolite of the closely related carcinogens ethyl carbamate (urethane) and vinyl carbamate. *Carcinogenesis*, **11**, 463–473.

Lipnick,R.L. and Fissekis,J.D. (1980) A comparative conformational study of certain 2′-deoxy-2′-fluoro-arabinofuranosylcytosine nucleosides. *Biochim. Biophys. Acta*, **608**, 96–102.

McCarter,J.D. and Withers,S.G. (1994) Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.*, **4**, 885–892.

Merrit,E.A. and Murphy,M.E.P. (1994) Raster3D version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr.*, **50**, 869–873.

Mol,C.D., Arvai,A.S., Slupphaug,G., Kavli,B., Alseth,I., Krokan,H.E. and Tainer,J.A. (1995) Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, **80**, 869–878.

Nair,J., Barbin,A., Guichard,Y. and Bartsch,H. (1995) $1,N^6$-ethenodeoxyadenosine and $3,N^4$-ethenodeoxycytidine in liver DNA from humans and untreated rodents detected by immunoaffinity ^{32}P postlabeling. *Carcinogenesis*, **16**, 613–617.

Nedderman,P. and Jiricny,J. (1994) Efficient removal of uracil from G-U mispairs by the mismatch-specific thymine DNA glycosylase from HeLa-cells. *Proc. Natl Acad. Sci. USA*, **91**, 1642–1646.

Panayotou,G., Brown,T., Barlow,T., Pearl,L.H. and Savva,R. (1998) Direct measurement of the substrate preference of uracil-DNA glycosylase. *J. Biol. Chem.*, **273**, 45–50.

Parikh,S.S., Mol,C.D., Slupphaug,G., Bharati,S., Krokan,H.E. and Tainer,J.A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J.*, **17**, 5214–5226.

Sandigursky,M. and Franklin,W.A. (1999) Thermostable uracil-DNA glycosylase from *Thermotoga maritima*, a member of a novel class of DNA repair enzymes. *Curr. Biol.*, **9**, 531–534.

Saparbaev,M. and Laval,J. (1998) $3,N^4$-ethenocytosine, a highly mutagenic adduct, is a primary substrate for *Escherichia coli* double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase. *Proc. Natl Acad. Sci. USA*, **95**, 8508–8513.

Saparbaev,M., Kleibl,K. and Laval,J. (1995) *Escherichia coli*, *Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair $1,N^6$-ethenoadenine when present in DNA. *Nucleic Acids Res.*, **23**, 3750–3755.

Savva,R., McAuley-Hecht,K., Brown,T. and Pearl,L.H. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487–493.

Schärer,O.D., Ortholand,J.-Y., Ganesan,A., Ezaz-Nikpay,K. and Verdine,G.L. (1995) Specific binding of the DNA-repair enzyme AlkA to a pyrrolidine-based inhibitor. *J. Am. Chem. Soc.*, **117**, 6623–6624.

Schärer,O.D., Kawate,T., Gallinari,P., Jiricny,J. and Verdine,G.L. (1997) Investigation of the mechanisms of DNA binding of the human G/T glycosylase using designed inhibitors. *Proc. Natl Acad. Sci. USA*, **94**, 4878–4883.

Schärer,O.D., Nash,H.M., Jiricny,J., Laval,J. and Verdine,G.L. (1998) Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *J. Biol. Chem.*, **273**, 8592–8597.

Seeberg,E., Eide,L. and Bjørås,M. (1995) The base excision repair pathway. *Trends Biochem. Sci.*, **20**, 391–397.

Slupphaug,G., Mol,C.D., Kavli,B., Arvai,A.S., Krokan,H.E. and Tainer,J.A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87–91.

Vassylyev,D.G., Kashiwagi,T., Mikami,Y., Ariyoshi,M., Iwai,S., Ohtsuka,E. and Morikawa,K. (1995) Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: Structural basis for damaged DNA recognition. *Cell*, **83**, 773–782.

Waters,T.R. and Swann,P.F. (1998) Kinetics of the action of thymine DNA glycosylase. *J. Biol. Chem.*, **273**, 20007–20014.