Chemistry of Glycosylases and Endonucleases Involved in Base-Excision Repair

Sheila S. David* and Scott D. Williams

Department of Chemistry, University of Utah, Salt Lake City, Utah 84112

Received January 6, 1998 (Revised Manuscript Received March 11, 1998)

Contents

I. Introduction 1221
II. Types of Base-Excision Repair Glycosylases 1223
A. Uracil-DNA Glycosylase (UDG) 1223
B. Mismatch-Specific Thymine-DNA Glycosylase (TDG) and Related Double-Strand Specific Uracil DNA Glycosylase (dsUDG) 1226
C. Alkylated Base Removal 1227
D. Endonuclease III and Related Enzymes 1229
E. Pyrimidine Dimer Glycosylases 1230
F. The FPG Protein (Fapy Glycosylase or MutM) 1231
G. Eukaryotic OG Glycosylases 1232
H. Mismatch-Specific Adenine Glycosylases: MutY and MYH 1233
III. Insights into Damage Recognition by BER Glycosylases from Structural and Biochemical Studies 1235
A. Uracil-DNA Glycosylase: Evidence for a Nucleotide-Flipping Mechanism 1235
B. DNA Recognition Motifs in a BER Superfamily 1237
1. Helix–Hairpin–Helix and GPD Motifs 1237
2. Fe–S Cluster Loop Domain of Endo III and MutY 1237
3. Substrate Recognition in the BER Superfamily: Endo III and AlkA 1239
C. Indirect Recognition of DNA Damage: T4-Endo V Bound to a Pyrimidine Dimer Containing DNA Duplex 1239
D. Mechanisms Facilitating Location of DNA Damage 1241
IV. Mechanistic Considerations 1242
A. Classification of Glycosylases vs Glycosylase/Lyase Enzymes 1242
B. Identification of Active-Site Amino Groups in T4-Endo V and FPG 1244
C. Identification of Imine Intermediates in the BER Superfamily 1245
1. Endonuclease III and Identification of a Mammalian Counterpart 1245
2. Yeast and Mammalian OG Glycosylase: Isolation, Sequencing, and Identification of a Critical Active-Site Lysine Residue 1245
D. Monofunctional vs Bifunctional Enzymes: The Controversy of MutY 1245
E. The Importance of an Aspartic or Glutamic Acid Residue in the Catalytic Activity of BER Glycosylases 1249
F. The Use of Synthetic DNA Analogues 1250
V. AP Endonucleases Associated with BER 1252
A. Classification of AP Endonucleases 1252
B. Substrate Specificity 1253
C. Mechanistic Properties of Class II AP Endonucleases 1255
D. Insight into Recognition of Abasic Sites by Exo III and HAP1 1256
VI. Prospectus for the Future 1257
VII. Abbreviations 1257
VIII. Acknowledgments 1257
IX. References 1258

I. Introduction

Cellular DNA continually suffers assault from exogenous and endogenous agents which lead to a wide variety of DNA modifications.[1] These modifications are often extremely detrimental to the cell leading to mutagenesis and carcinogenesis. An array of enzymes are given the daunting task of maintaining the integrity of DNA and many DNA repair mechanisms have been extensively characterized. Although different DNA repair pathways have been uncovered, these pathways are often remarkably conserved from bacteria to humans, which underscores their importance in maintaining the functional properties of DNA. There are an increasing number of examples illustrating the importance of DNA repair in the prevention of diseases such as cancer. For example, the relationship between aberrant nucleotide excision repair (NER) and the cancer-prone genetic disorder, xeroderma pigmentosum, has been well-documented.[2,3] More recently, a direct correlation between defective DNA mismatch repair and hereditary colon cancer has been established.[4] In the case of base-excision repair glycosylases, a possible link to cancer has recently been suggested by the observation that the human gene *hogg1*, which encodes for a repair enzyme involved in the repair of oxidatively damaged guanine residues, is located in a region of the chromosome often deleted in lung cancers.[5] Thus, these examples, as well as many others, establish that DNA repair enzymes play a critical role in the cell. This underscores the importance of a detailed understanding of their modes of damage recognition and repair mechanisms in order to gain insight into diseases such as cancer.

S0009-2665(98)00321-5 CCC: $30.00
Published on Web 04/17/1998

Sheila S. David grew up in Ann Arbor, MI. She received her B.A. in Chemistry at Saint Olaf College in Northfield, MN in 1984. She did her graduate work at the University of Minnesota in bioinorganic chemistry in the laboratory of Professor Lawrence Que, Jr., and received her Ph.D. in 1989. Her interest in nucleic acid chemistry drew her to the laboratory of Professor Jacqueline K. Barton at the California Institute of Technology where she was an NIH postdoctoral fellow (1990–1992). She joined the faculty at the University of California, Santa Cruz, in the Department of Chemistry and Biochemistry as an Assistant Professor in 1992. During 1996, she moved her laboratory to Salt Lake City to join the faculty in the Department of Chemistry at the University of Utah. Her research interests are in the areas of DNA repair and metalloprotein–nucleic acid interactions. Presently, her laboratory is primarily focused on the fascinating DNA repair enzyme, MutY. For relaxation, Sheila enjoys participating in outdoor activities in the scenic mountains of Utah with friends (particularly, husband Pete Beal).

Scott Williams was born and raised in Boulder, CO. He earned his B.S. in Biochemistry and Cell Biology at the University of California, San Diego, in 1993. He then returned to Boulder to work with Dr. Ken Douglas as a research assistant in condensed matter physics at the University of Colorado. In 1995 he began his graduate studies at the University of California, Santa Cruz, under the direction of Dr. Sheila David. When Dr. David moved her laboratory to the University of Utah, Scott moved as well, and is continuing his thesis research in the field of DNA repair enzyme biochemistry. In addition to biological chemistry, Scott avidly enjoys mountain biking, Nordic skiing, and backpacking trips with his girlfriend. He also likes the occasional drive across the Rockies to visit his family in Colorado.

Intense interest in DNA damage recognition and DNA repair mechanisms is resulting in an increasingly more sophisticated understanding of the chemistry underlying DNA repair pathways. Notably, crystallographic studies on DNA repair enzymes have provided keen insight into the requirements for recognition of specific types of damaged DNA and into the chemistry of the repair processes. Two crystallographic studies of DNA repair enzymes bound to their substrate-containing DNA complexes are now available. These structures were obtained using the technique of site-directed mutagenesis to alter a specific amino acid(s) to inhibit enzymatic turnover while retaining specific substrate binding. Mutagenesis techniques have also provided insight into important amino acid residues needed for DNA damage recognition and for effecting DNA repair chemistry. The use of synthetic DNA substrates has allowed for a more detailed and quantitative analysis of the repair chemistry. Furthermore, advances in the methods for preparation of modified oligonucleotides have provided new avenues for exploring the DNA repair enzyme's substrate recognition by specific alterations of the substrate DNA.

DNA repair enzymes catalyze a variety of reactions upon DNA in the course of effecting repair.[3] However, most DNA repair enzymes require the participation of other enzymes to complete the repair process. The most common mechanism of DNA repair is excision repair, which involves removal of damaged or inappropriately base-paired DNA. Subsequently, separate DNA replication machinery is recruited to replace the excised region.[3] In the case of base-excision repair (BER), DNA repair glycosylases catalyze the removal of aberrant bases. Base-excision repair is distinguished from other types of excision repair, namely, nucleotide-excision repair (NER) and mismatch repair (MR). NER repairs a wide variety of DNA damage, such as pyrimidine dimers, by DNA phosphodiester hydrolysis to release a small oligonucleotide fragment (~12 nucleotides) containing the DNA damage. In contrast, MR corrects mismatched base pairs by removal of large DNA fragments (>1000 nucleotides). Recently, an alternative excision repair pathway for UV-damaged DNA has been identified in *S. pombe*, *N. crassa*, and *B. subtilis*.[6−11] This excision repair pathway cleaves the DNA phosphodiester backbone on the 5′-side of the UV-damaged site to initiate the repair process.

There are examples of DNA repair enzymes which catalyze direct chemical reversal of DNA damage.[3] An extensively characterized member of this class is DNA photolyase, which repairs cyclobutane pyrimidine dimers in DNA by utilizing the energy of visible light to break the cyclobutane ring of the dimer.[12] Other examples of direct reversal include alkyltransferase enzymes which remove aberrant alkylation of the phosphate backbone, and at the O-6 and O-4 positions of guanine and thymine, respectively.[13]

Since *Science* magazine heralded DNA repair enzymes as "Molecules of the Year" in December of 1994, a number of excellent review articles have been published which focus on various aspects of DNA repair. In addition, a textbook entitled "DNA Repair and Mutagenesis" was published in 1995.[3] The excitement centered around DNA repair has led to an explosion of research in this area during the past five years. In our review, we have focused on chemical aspects of recognition and DNA cleavage reactions catalyzed by base-excision repair glycosylases and apurinic-apyrimidinic (AP) endonucleases. We have not included many details associated with the biological aspects of this important process due

to our determination that these aspects have received sufficient coverage in other review articles. This is not meant to underestimate the importance of seminal biological experiments which have led to the identification of many DNA repair enzymes and determination of their biological roles in the cell. In our discussions of a particular DNA repair enzyme, previous review articles on that enzyme are duly noted for the reader to obtain additional information.

II. Types of Base-Excision Repair Glycosylases

Damage to individual DNA bases is repaired by the base-excision repair (BER) pathway.[3,14] The marquee players in the BER pathway are DNA glycosylase enzymes, which recognize a variety of modified or mismatched bases and catalyze cleavage of the N-glycosidic bond to release the inappropriate base from the deoxyribose ring. Many glycosylase enzymes also catalyze a β-elimination (or lyase) reaction to effect strand scission after base removal. Subsequent action of apurinic-apyrimidinic (AP) endonucleases and 3′- and 5′-phosphodiesterases remove the remaining sugar fragments to produce a single nucleotide gap with the proper 3′-hydroxyl (3′-OH) and 5′-phosphate ends to be acted upon by DNA polymerase, which adds the correct nucleotide. Finally, DNA ligase finishes the BER process by sealing the gap in the backbone.[15] The BER pathway is schematically represented in Figure 1A.

The BER pathway has been reconstituted in vitro with cell-free extracts or purified protein components and these experiments have established the minimal requirements for restoration of damaged DNA.[16−18] After base removal, a single-nucleotide gap is produced by the action of a 5′-AP endonuclease and deoxyribophosphodiesterase (dRPase) or by the successive β- and δ-elimination reactions of bifunctional glycosylases (see section V.A). This gap is then filled with the aid of various enzymes depending on the organism. For example, repair of uracil in DNA was achieved by the use of five *E. coli* proteins: uracil-DNA glycosylase, AP endonuclease IV, RecJ protein, DNA polymerase I, and DNA ligase.[16] This suggests that in vivo these five protein components are sufficient to effect the repair. Importantly, in these reconstitution experiments, the RecJ protein was required for the single-nucleotide replacement and in its absence a longer repair patch was produced by the 5′-nuclease function of DNA polymerase.

In mammalian cells, after base removal, two pathways have been observed which involve short-patch or long-patch removal and replacement (Figure 1B). The short-patch pathway involves single nucleotide gap replacement and requires the action of DNA polymerase β.[19,20] Recent reports have also shown that specific protein−protein contacts between DNA polymerase β and the major human AP endonuclease, HAP1 (section V) coordinate the sequential reactions of these enzymes on abasic site-containing DNA.[21] Interestingly, the replacement synthesis is also mediated by the interaction of XRCC1 (X-ray cross complementation protein 1) with DNA polymerase β and DNA ligase III.[22,23] This short-patch pathway can be reconstituted in vitro using human uracil-DNA glycosylase, HAP1, DNA polymerase β, XRCC1, and DNA ligase I or III.[23] A second pathway for completion of BER in eukaryotes involving polymerase β or δ results in the removal and replacement of longer patches of 2−10 nucleotides.[19,24] This pathway involves the action of flap-endonuclease (FEN-1) and depends on the presence of PCNA (proliferating cell nuclear antigen).[25,26] In mammalian cells, the long-patch repair pathway occurs at oxidized or reduced AP sites, while normal AP sites are processed by the short-patch pathway.[26]

The BER pathway appears to be critical and is highly conserved from bacteria to humans. Indeed, DNA glycosylase and glycosylase/lyase enzymes have been isolated from numerous sources including bacteria, yeast, and mammals. Much of the pioneering work was performed with bacteria and has proved critical for identification of similar enzymes in eukaryotes. DNA glycosylases attend to a wide variety of lesions arising from both endogenous and exogenous factors. The types of damage include alkylation, oxidation, and hydrolysis. Examples of these reactions in DNA are illustrated in Figure 2. Base modifications resulting from alkylation, oxidation, or deamination are potentially mutagenic by virtue of alterations in duplex base-pairing motifs or lethal to the cell by inhibiting DNA synthesis.

Several DNA glycosylases with different substrate targets have evolved to assist in the restoration of damaged DNA. Examples of well-characterized DNA glycosylases are listed in Table 1. Most BER enzymes are named after the substrate initially determined for these enzymes. As will be illustrated in the discussions below, subsequent work often indicated that the first substrate identified was not necessarily the best or the most biologically relevant substrate. Another complicating feature of many BER glycosylases is that they recognize a variety of substrates. The relative ranking of substrates for BER glycosylases in causing detrimental effects in vivo is unclear and remains an important and challenging issue to be resolved.

A. Uracil-DNA Glycosylase (UDG)

The major source of uracil in DNA is from cytosine residues which have been hydrolytically deaminated.[1,27] The presence of uracil in DNA may also arise as the result of misincorporation of 2′-deoxyuracil during DNA replication under the unusual circumstance when dUTP has eluded processing by cellular dUTPase.[28] The detection of cytosine deamination within DNA prompted the search for enzymes involved in the repair of uracil in DNA. Using a [^{3}H]-dU substrate oligonucleotide, an activity was purified from *E. coli* that released [^{3}H]uracil.[29] The enzyme was dubbed "uracil glycosidase" and has the distinction of being the first enzyme that removes aberrant bases to be discovered.[29] This famous enzyme, renamed uracil-DNA glycosylase (UDG), has been extensively studied and previously reviewed.[28,30] Cytosine deamination events result in the formation of G:U base pairs which, if left unrepaired, will result in G:C to A:T transversion mutations. In *E. coli* strains lacking a functional UDG, an increase in G:C

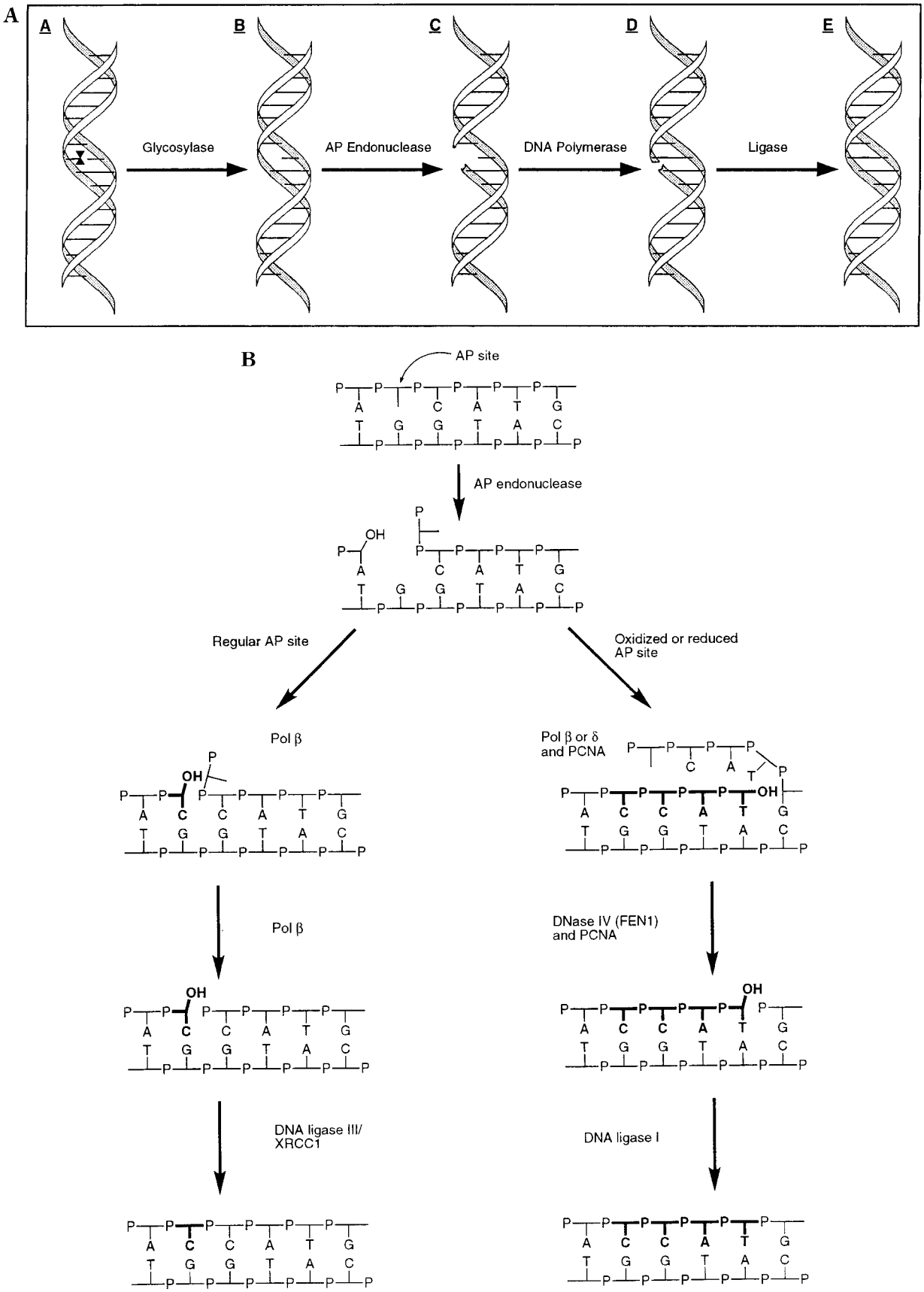

Figure 1. (A) Schematic representation of the base-excision repair (BER) pathway. Base modification (or mismatched base) in the DNA duplex A is recognized and removed by a BER glycosylase to form an AP site B. Either lyase activity intrinsic to the BER glycosylase or separate AP endonuclease activity at the resulting AP site causes DNA strand cleavage. 5′- and 3′-phosphodiesterases may also be recruited to remove remaining sugar fragments to provide a one nucleotide gap with 3′-hydroxyl and 5′-phosphate ends C. DNA polymerase activity replaces the nucleotide gap with the correct nucleotide D, and the phosphodiester backbone is sealed by ligase E. (B) Different repair pathways that occur after removal of the base at the AP site in human cells. (Adapted from ref 26.)

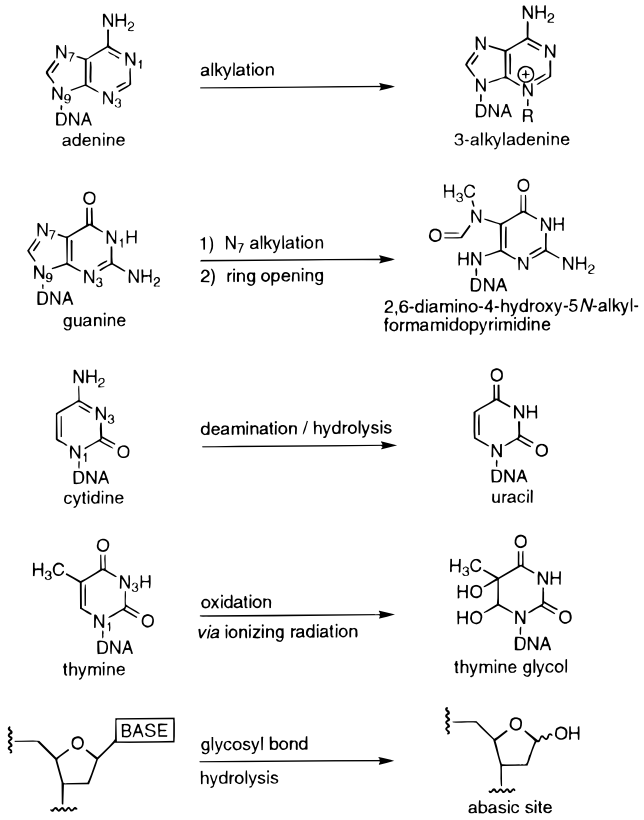

Figure 2. Examples of common modifications of the four DNA bases, all of which have potential cellular mutagenicity and toxicity. Glycosyl bond hydrolysis can occur regardless of base type (although more frequent for purines) and therefore "BASE" refers to any of the four bases.

to A:T mutations is observed, consistent with cytosine deamination as the dominant pathway for uracil formation in DNA.[31] UDG is an enzyme common to the pox and herpes virus families as well as prokaryotic and eukaryotic organisms.[30] Additionally, the genes from eukaryotic, prokaryotic, and viral organisms for UDG are highly conserved, indicating a similar active site and mechanism.[32]

UDG specifically cleaves uracil residues from both double- and single-stranded DNA, with slightly higher activity toward single-stranded DNA.[30] However, uracil in single-stranded loop regions is inefficiently removed by UDG[33] and the extent of its removal can be modulated by the presence of a single-stranded binding protein.[34] UDG is not active toward uracil-containing RNA, mononucleotides, or nucleosides.[35] Efficient UDG recognition of uracil-containing DNA appears to require phosphate backbone contacts since processing of small DNA substrates is observed only when phosphorylated at both the 5′ and 3′-ends.[36] Specific phosphate contacts have also been observed in the crystallographic studies of UDG bound to DNA (see section III.A). Consistent with these observations, the minimal substrate for UDGs is pd(UN)p.[36] UDG's ability to recognize uracil in both duplex and single-stranded DNA sets it apart from other known DNA glycosylases, which act upon their substrates in duplex DNA only.

Indeed, the substrate specificity of this BER enzyme is rather remarkable in light of its crucial ability to exclude thymine residues as substrates

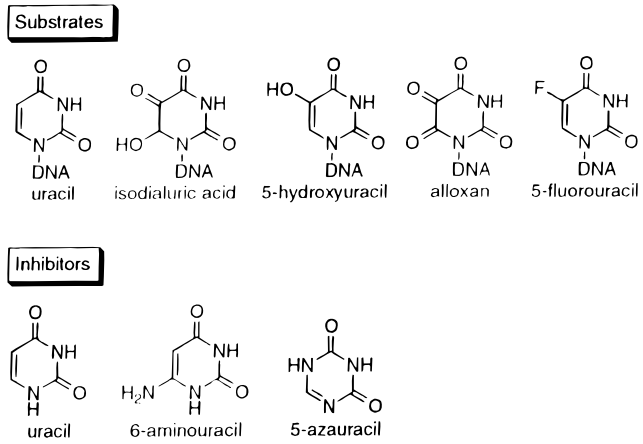

Figure 3. Uracil-DNA glycosylase (UDG) substrates and inhibitors. These substrates are listed independent of UDG source. There are small differences in specificity depending on source; although all UDG enzymes remove uracil. See text for more details.

while so efficiently cleaving uracil from DNA. UDG has also been found to excise 5-fluorouracil (5-FU) from DNA, albeit slowly.[37,38] 5-FU is a cancer chemotherapeutic agent and its pharmacological properties may be mitigated by the action of UDG.[37,38] Human UDG (hUDG) has also been shown to recognize uracil derivatives generated from oxidative damage to cytosine using γ-irradiation.[39] GC/MS analysis of supernatant fractions of aerated aqueous hUDG reactions with γ-irradiated DNA detected 5-hydroxyuracil, isodialuric acid, and alloxan in addition to uracil.[39] These three additional substrates for hUDG were the only ones identified out of 12 total lesions detected in the γ-irradiated DNA. *E. coli* UDG has been shown to excise isodialuric acid from oxidatively damaged DNA[40] and 5-hydroxyuracil from oligonucleotide substrates containing a single 5-hydroxyuracil base.[41] These oxidized UDG substrates appear to be removed slowly relative to uracil; however, a quantitative measurement of the relative substrate specificity within the same sequence context and under the same conditions has not been reported. These results suggest a possible involvement of UDG in the repair of oxidative DNA lesions.

The UDG enzymes can be inhibited by free uracil base and the modified uracil bases 6-aminouracil and 5-azauracil (Figure 3).[42] 5-Fluorouracil serves as a weak inhibitor for UDG, while 5-bromouracil and 5-methyluracil are ineffective. Indeed, a number of modified uracil bases are unable to inhibit UDG. These ineffective modified uracil bases encompass changes in positions 1−4 of the pyrimidine ring. The only effective inhibitors have modest changes at the 5 and 6 positions, as in 5-FU. The ability of UDG to remove isodialuric acid from DNA indicates that planarity is not required for recognition, and also indicates a higher tolerance of UDG for substitutions at the 5,6 positions of uracil.[39,40] These results led to the early suggestion that UDG contains a pocket which is specific for uracil and that this serves as the mechanism for exclusive substrate recognition. This idea was later confirmed by the X-ray structure of UDG bound to uracil-containing DNA (section III.A).

Table 1. Types of Base-Excision Repair Glycosylases[a]

enzyme	sources	representative substrate[b]	bifunctional	BER superfamily member
uracil DNA glycosylase	viral	uracil	no	no
	bacterial	uracil	no	no
	plant	uracil	no	no
	human	uracil	no	no
	S. cerevisiae	uracil	no	no
thymine DNA glycosylase	human	G:**T**	no	no
	M. thermoautotrophicum	G:**T**	no?	yes
double strand uracil DNA glycosylase	insects	G:**U**	no	no
AlkA[c]	*E. coli*	3-MeA	no	yes
Tag[c]	*E. coli*	3-MeA	no	no
MAG[c]	*S. cerevisiae*	3-MeA	no	no
MPG[c]	plant	3-MeA	no	no
	rat	3-MeA	no	no
	human	3-MeA	no	no
endonuclease III	*E. coli*	thymine glycol	yes	yes
NTG1/SCR1 and SCR2[d]	*S. cerevisiae*	thymine glycol	yes	yes
Nth-Spo[e]	*S. pombe*	thymine glycol	yes	yes
endonuclease III-like	bovine	thymine glycol	yes	yes
	human	thymine glycol	yes	yes
endonuclease VIII	*E. coli*	thymine glycol	yes	no
T4-endonuclease V	bacteriophage T4	*cis-syn* thymidyl(3′→5′)thymidine	yes	no
UV-endonuclease	*M. luteus*	*cis-syn* thymidyl(3′→5′) thymidine	yes	yes
CPD glycosylase	*S. cerevisiae*	*cis-syn* thymidyl(3′→5′) thymidine	yes	?
FPG[f]	bacterial	OG	yes	no
yOGG1[g]	*S. cerevisiae*	OG	yes	yes
mOGG1[g]	mouse	OG	yes	yes
hOGG1[g]	human	OG	yes	yes
MutY	*E. coli*	OG:**A**	no[i]	yes
MYH[h]	bovine	OG:**A**	?	yes
	human	OG:**A**	?	yes

[a] Appropriate references can be found in the text [b] OG: 7,8-dihydro-8-oxoguanine. 3-MeA: 3-methyladenine. [c] 3-Methyladenine glycosylase. [d] NTG1: endonuclease III-like glycosylase 1. SCR1 and 2: *S. cerevisiae* redoxyendonuclease 1 and 2. [e] Endonuclease III - *S. pombe.* [f] fapy glycosylase or MutM. [g] OG glycosylase. [h] MutY homologue. [i] See section IV.D.

The excision of uracil from DNA is also dependent on the DNA sequence surrounding the uracil and the base opposite uracil in duplex DNA. If the presence of uracil in DNA results primarily from cytosine deamination, one might expect that U:G base pairs would predominate in DNA and may be the logically preferred substrates for UDG. However, there is no consensus regarding the preference of UDG for U:G vs U:A base pairs in DNA. Several groups have reported higher activity of U:A over U:G base pairs,[43] while others have reported U:G to be preferred over U:A base pairs.[44,45] However, UDG also appears to be sensitive to the sequence context of uracil with a 10−15-fold variation in the activity of UDG.[43,45,46] In general, the sequence context effects have been found to be more important for determining the rate of uracil removal than whether the uracil is in a matched (U:A) or mismatched (U:G) base pair.[11] Such sequence context effects may explain the conflicting data on the relative efficiency of UDG toward U:A and U:G base pairs. In general, UDG has been found to be more efficient in A−T-rich sequences over G−C-rich sequences. The consensus for facile sequences is A/T-UA-A/T and the consensus for slowly processed UDG sequences is G/C-U-T/G/C. However, G−C base pair content was found not to be the only factor, since sequences containing a 5′-XUT-3′ were particularly poor substrates for human and *E. coli* UDG.[43,45,46] There appears to be no interpretable sequence context effects for UDG action on single-stranded DNA as there is with duplex DNA. This points to the importance of local melting around the uracil for its efficient removal by UDG.[45] Interestingly, Nilsen et al.[46] found a correlation between sequences which had been shown to have a relatively high frequency for mutations and sequence contexts which are slow substrates for UDG. This suggests that the observed sequence-dependent variation in removal rates is biologically significant and that inefficient repair in certain sequence environments may contribute to the appearance of DNA mutations.

B. Mismatch-Specific Thymine-DNA Glycosylase (TDG) and Related Double-Strand Specific Uracil DNA Glycosylase (dsUDG)

The discovery and properties of thymine-DNA glycosylase (TDG) have been recently reviewed by Lettieri and Jiricny[47] and the reader is referred to this review for more details. The discovery that 5-methylcytosine is prone to deamination, producing G:T mismatches, led to the search for an activity that restored G:T mismatches to G:C base pairs. Jiricny and co-workers discovered that such repair was occurring in mammalian cells[48] and a protein able to bind DNA containing G:T mismatches was identified.[49] By using synthetic duplexes containing a G:T mismatch, Jiricny's laboratory identified a glycosylase activity in HeLa cell extracts specific for removal of thymine in a G:T mismatch.[18] Further studies indicated that partially purified TDG protein bound tightly to G:U-containing duplexes and was able to

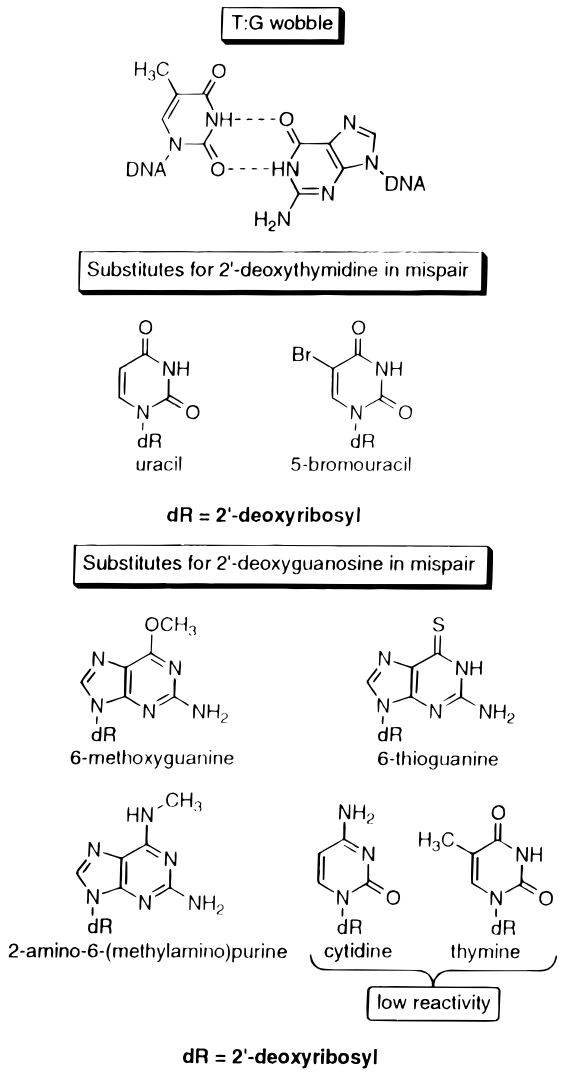

Figure 4. Thymine-DNA glycosylase (TDG) substrates. At the top is the expected T:G wobble base-pair structure. Below are respective substituting compounds which can also form mispair TDG substrates.

efficiently remove uracil from G:U base pairs.[50] Using this knowledge, Jiricny's group then purified TDG by using a G:U-derivatized affinity column.[51] Eventually, TDG from HeLa cells was cloned and isolated from bacterial and eukaryotic expression systems.[52] The sequence information indicated that TDG is a novel DNA repair protein containing no motifs found in other well-characterized BER glycosylases.[52]

Like UDG, TDG is active toward a rather limited number of substrates; however, unlike UDG, TDG is strictly mismatch-specific and has no activity toward thymine or uracil in single-stranded DNA.[47] The activity of TDG appears to be influenced by the sequence context.[53−55] Several studies have addressed the effect of the base 5′ to the G in the G:T base pair, and have found that G:T mispairs in the CpG context are processed to a 3−12-fold greater extent than TpG, GpG, and ApG contexts. However, a complete delineation of the sequence preferences for TDG has not been reported.

Surprisingly, TDG has been found to be slightly more efficient toward uracil in G:U base pairs than toward thymine in G:T mispairs.[50] Thymine from T:T and C:T base pairs can also be removed by TDG with low efficiency. TDG has also been shown to be active toward removal of 5-bromouracil (5-BrU) in 5-BrU:G mismatches, which is not surprising based on the similar sizes of methyl and bromo substituents.[50] TDG can tolerate some modifications of the guanine in the G:T mismatch. In particular, O^6-methylguanine:T, 6-thioguanine:T, and 2-amino-6-(methylamino)purine:T mispairs are substrates for TDG.[54−56] The diaminopurine:T base pair had been identified as a substrate using cell extracts.[57] This mispair was later demonstrated not to be a substrate using purified recombinant TDG.[55] Sibghat-Ullah and Day conclude that the unifying property of the substrates for TDG is that they maintain the wobble base pair hydrogen bonding scheme of the G:T mismatch. The G:T wobble base pair would be expected to cause a distortion in the double helix and this may be a recognition feature for TDG. 4-MeT:G in DNA is not a substrate for TDG, indicating a possible requirement for thymine to fit into an active site pocket.[55] This is analogous to the inability of UDG to process uracil substituted at certain positions, such as 5-bromouracil or thymine. The identified substrates for TDG are illustrated in Figure 4.

The ability to process G:T base pairs is lost upon truncation of TDG by 112 amino acids at the N-terminus. However, the truncated TDG retains the ability to remove uracil from a G:U base pair.[47,58] This surprising observation suggests the N-terminal portion modulates the properties of the active site to allow access of thymine or 5-bromouracil. This hypothesis is consistent with the discovery of bacterial genes homologous to the truncated version of TDG.[58] Isolation and expression of an *E. coli* homologue indicates that this new protein has an activity similar to the truncated TDG. These new enzymes are termed dsUDGs. They are similar to UDG in their ability to remove uracil; however, they remove uracil more slowly than UDG and they operate exclusively on G:U base pairs within double-stranded DNA. Gallinari and Jiricny[58] also identified dsUDG activity in pupating insects thought to lack the ability to remove uracil from DNA. The high specificity of dsUDGs for G:U base pairs suggests that they have arisen in situations where cytosine deamination is prevalent.

C. Alkylated Base Removal

Alkylating agents react with DNA to form numerous modified bases. A subset of these alkylated bases are removed by the BER pathway. Many DNA adducts formed with bulky alkylating agents are processed by the nucleotide excision repair pathway.[59] Prior reviews have discussed alkylation-specific BER glycosylases and the reader is directed to these reviews for additional information.[13,59,60] DNA glycosylases specific for the removal of alkylated bases were first detected in *E. coli* and have been subsequently identified in other prokaryotic (*Bacillus subtilus*),[61] yeast (*Saccharomyces cerevisiae, Schizosaccharomyces pombe*),[62−64] mammalian (rat,[65] mouse,[66] human[67−69]), and plant (*Arabidopsis thaliana*)[70] cells.

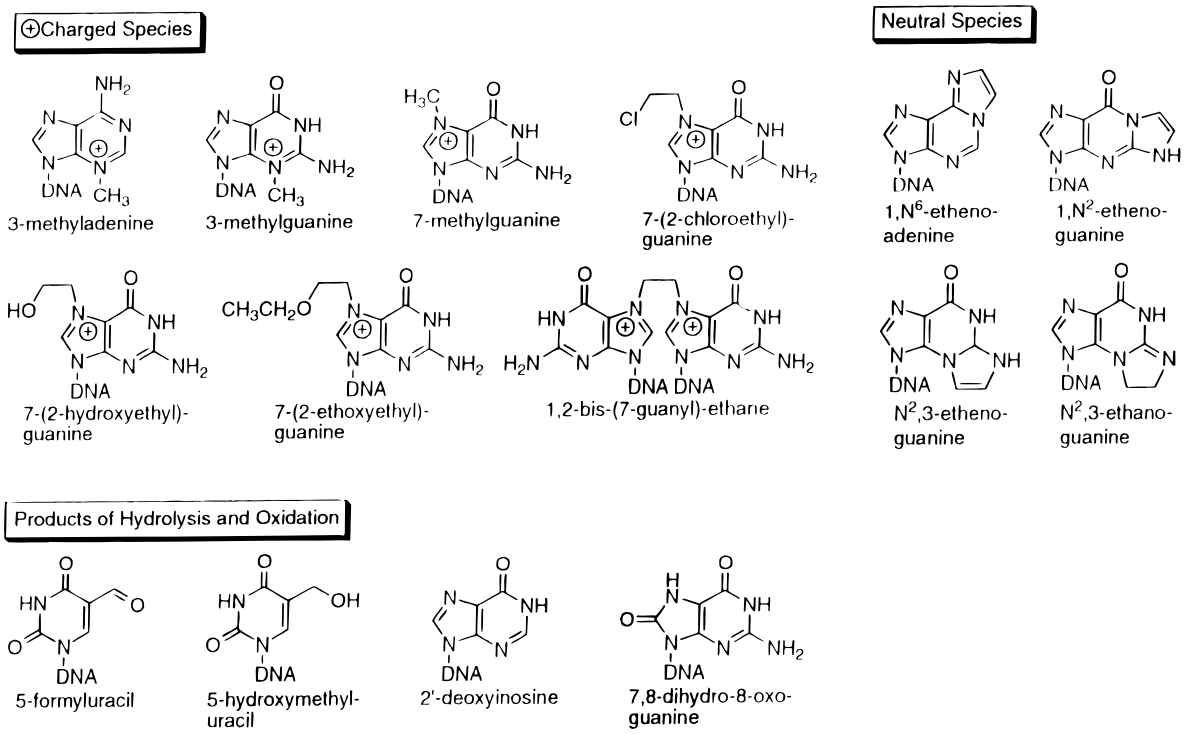

Figure 5. Representative substrates for 3-methyladenine glycosylases from *E. coli*, *S. cerevisiae*, and humans. Substrates are grouped by substrate type. Relative rates for removal of these substrates from DNA depends on enzyme source. Consult text for additional details.

In *E. coli*, the BER glycosylase AlkA (also called 3-methyladenine glycosylase II) is induced upon exposure of cells to alkylating agents. A second enzyme in *E. coli* called 3-methyladenine glycosylase I (or Tag protein) is constituitively expressed.[60] Enzymatic counterparts with substrate specificity similar to AlkA have been identified in *S. cerevisiae* (termed MAG or 3-methyladenine glycosylase)[71−73] and in mammalian cells (referred to as MPG, ANPG, or 3-methyladenine glycosylase).[65,74] Many of the eukaryotic 3-methyladenine DNA glycosylases were cloned by their ability to rescue *alkA*$^-$tag$^-$ *E. coli* from the detrimental effects of alkylating agents. The *S. cerevisiae* MAG enzyme appears to have significant sequence homology to the *E. coli* AlkA enzyme.[62,71,72,75] The mammalian enzymes have regions of homology to AlkA,[69] although the overall homology is limited.[68,76] In this review, the focus will be on the best-characterized enzymes in this class.

These enzymes are generally referred as "3-methyladenine glycosylases" since this was often the first substrate identified for the enzyme; however, many of these enzymes recognize more than one type of alkylated base. The inducible *E. coli* AlkA enzyme, the *S. cerevisiae* MAG and human 3-methyladenine glycosylase all have an extremely wide substrate specificity.[13] Representative substrates recognized by AlkA and related eukaryotic enzymes are illustrated in Figure 5. The relative activity toward the different substrates depends on the source of the 3-methyladenine glycosylase. The primary substrates are those resulting from alkylation at the N-3 and N-7 positions of adenine and guanine, respectively. Endogenous sources of methylation include *S*-adenosylmethionine which produces both 7-methylguanine (7-meG) and 3-methyladenine (3-meA).[77] Of these two, 3-meA is likely an important substrate in vivo due to its ability to block DNA synthesis and cause mutations.[59,78] Significantly, AlkA and some of its eukaryotic counterparts are also able to repair alkylation damage caused by exogenous sources such as clinically useful nitrogen mustards,[79] the chemical warfare agent mustard gas [bis(2-chloroethyl) sulfide][80] and chloroethyl ethyl sulfide.[81] Adducts formed by the reaction of haloethylnitrosourea antitumor agents with DNA (e.g. as 7-(hydroxyethyl)guanine and 7-(chloroethyl)guanine) are removed by 3-meA glycosylases, protecting cells from the toxic effects of these compounds.[73]

In addition to the repair of purines alkylated at N-7 or N-3, removal of O-alkylated pyrimidines has also been observed.[13] Related ring-expanded products that result from reaction of guanine or adenine with liver metabolites of vinyl chloride compounds are also found to be removed by the various types of 3-methyladenine glycosylases.[42] For example, the removal of N^2,3-ethenoguanine and N^2,3-ethanoguanine by AlkA has been observed, although the efficiency was estimated to be 1/20th of that for removal of 3-meA.[82,83] 1,N^6-Ethenoadenine has been shown to be removed by *E. coli*, *S. cerevisiae*, rat, and human 3-methyladenine DNA glycosylases.[84,85] Surprisingly, 1,N^6-ethenoadenine is *preferred* as a substrate for the human 3-methyladenine DNA glycosylase. This raises interesting questions about the recognition mechanism used by the human enzyme and the identity of the primary substrate in vivo.[13,86] Products of hydrolysis reactions, namely hypoxanthine,[72] and oxidative reactions such as 5-formyluracil, 5-hydroxyuracil,[87] and 8-oxoguanine[88] have also been shown to be released by the bacterial and mammalian 3-meA glycosylases, albeit with reduced ef-

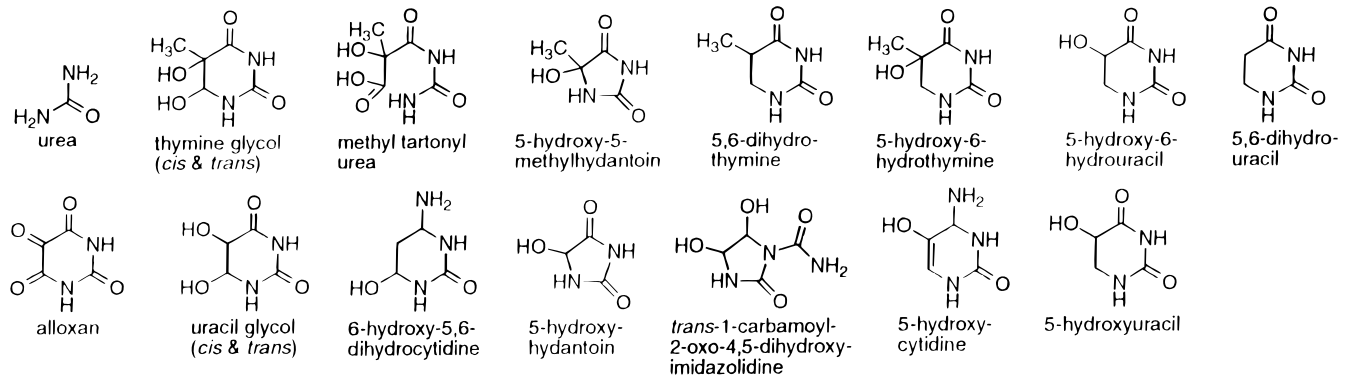

Figure 6. Bases released by *E. coli* endonuclease III. The types of bases released by endo III are products of oxidative reactions of cytosine and thymine.

ficiency. The diversity of modifications processed by these enzymes is illustrated in Figure 5. The large number of structural variations that must be accommodated by the active site of these enzymes suggests a recognition mode that capitalizes on unifying features of these substrates rather than the use of a base-specific pocket as with UDG. Considerable insight into this issue has been provided by the structural characterization of the AlkA enzyme (section III.B.3).

In contrast to AlkA, the bacterial cousin enzyme Tag removes primarily 3-methyladenine in DNA[43] with weak activity for removal of 3-meG.[89] In addition, only Tag is inhibited by the product base, 3-meA.[60] The amino acid sequences of the two *E. coli* enzymes are also quite distinct.[75,90] In an interesting twist, it has recently been shown that although both enzymes act upon 3-methyladenine in duplex DNA with similar efficiency, AlkA is 10−20 times more active upon 3-methyladenine in single-stranded DNA than Tag.[91] The authors postulate that the activity of AlkA toward 3-methyladenine in single-stranded DNA may be important for removal of alkylation damage from single-stranded DNA formed transiently during transcription and replication. Although this study showed that AlkA is more active toward removal of 3-methyladenine in single-stranded DNA than the Tag enzyme, AlkA is considerably less efficient at the removal of 3-methyladenine in single-stranded than in double-stranded DNA. This is in contrast to UDG, which is active on both single-stranded and double-stranded DNA, with higher efficiency for the single-stranded form. The results on AlkA and Tag taken together suggest that the recognition mechanism for 3-meA is likely to be significantly different for the two enzymes. AlkA may recognize features that are common to alkylated bases, while Tag may recognize 3-meA in a specific fashion.

D. Endonuclease III and Related Enzymes

Endonuclease III (endo III) was first identified as an activity in *E. coli* which introduced strand scission in X-irradiated DNA,[78] UV-irradiated DNA,[92] and OsO_4-treated DNA.[93] Since the originally identified activity was DNA strand scission, the enzyme was called *endonuclease* III or X-ray *endonuclease.* Later, endo III was found to be a *glycosylase* specific for removal of damaged thymine residues such as urea and thymine glycol, in addition to catalyzing strand scission after damaged base removal.[94−96] Endo III was also referred to as thymine glycol glycosylase or urea glycosylase before it was determined all three refer to the same enzyme. A wide variety of additional substrates have been identified for endo III, endowing it with one of the widest substrate specificities of the BER glycosylases.[41,97−101] The array of endo III substrates share the common feature of having resulted from oxidative damage to thymine or cytosine. These substrates are listed in Figure 6. Most of these substrates derive from oxidative reactions at the 5,6-double bond in thymine or cytosine, such as thymine glycol and 5-hydroxycytosine.[41] However, in some substrates for endo III, oxidation of the 5,6-double bond has resulted in bond cleavage, producing fragmentation products, such as methyltartonylurea, or ring contraction products, such as 5-hydroxyhydantoin.[98,101]

Recently, endo III has been shown to mediate DNA cleavage of an unidentified UV photoproduct of dipyrimidine, TC and CC sequences; the target lesion is distinct from known products of UV and oxidative damage at pyrimidine dimers.[102] However, this indicates that additional substrates for endo III may exist that have eluded detection due to their lability and low abundance. A common theme of the identified endo III substrates is loss of aromaticity in the pyrimidine ring resulting in loss of planarity of the base. The exceptions to this include 5-hydroxycytosine and 5-hydroxyuracil.[41] The active site of endo III must be able to accommodate this wide array of substrates (see section III.B.3).

The relative activity of endo III toward many of these substrates and their in vivo importance remains to be determined. A difficulty encountered with endo III substrates is that they are not easily incorporated into synthetic oligonucleotides due to their intrinsic lability. In a recent study by Wang and Essigmann,[103] an electrophoretic-based assay was used to determine the relative activity of endo III toward oxidized cytosine products site specifically incorporated into identical duplexes. 5-Hydroxycytosine (5-OH-C)- and 5-hydroxyuracil (5-OH-U)-containing oligonucleotides were synthesized using standard phosphoramidite chemistry; however, the lability of uracil glycol (UG) required the combined

use of chemical and enzymatic methods to effect its site-specific incorporation. By using these directly comparable substrates, the relative ranking of Ug > 5-OH-C > 5-OH−U was obtained. The kinetic values indicate that the rates for release of these products is within the ranges reported for other endo III substrates. Thus, studies such as these will be important for the quantitative determination of endo III's activity with different substrates and may shed light on the physiological relevance of these substrates for endo III. Endo III appears to be important in the repair of oxidative damage; however, *E. coli* cells lacking endo III exhibit only a weak mutator phenotype and are not sensitive to oxidizing agents.[104] This likely stems from redundant pathways for the repair of oxidative DNA damage.

Since the early discovery of endo III, similar enzymes have been identified from a wide variety of sources including yeast[105,106] and mammals.[107,108] Initial studies on the yeast and mammalian enzymes indicate that their substrate specificity is similar to that of *E. coli* endo III. In *S. cerevisiae*, cloning and expression has shown that there are two distinct yeast homologues to endo III.[109,110] One of these homologues , called NTG1 (endonuclease three-like glycosylase 1)[110] or SCR1 (*Saccharomyces cerevisiae* redoxyendonuclease 1)[109] seems to be similar in sequence to endo III but lacks the Fe/S cluster domain (section III.B.2). The enzymatic properties of SCR1/NTG1 are similar to endo III in that thymine glycol, pyrimidine photohydrates, and dihydrouracil bases are substrates;[109,110] however, SCR1/NTG1 is also able to catalyze the removal of Fapy-guanine (2,6-diamino-5-formamidopyrimidine) which is an activity not presently ascribed to endo III. The sequence of the second endo III homologue SCR2 (*Saccharomyces cerevisiae* redoxyendonuclease 2) is highly similar to endo III and unlike SCR1/NTG1 also contains the Fe/S domain.[109] Additional studies with various substrates will be required to differentiate between SCR1/NTG1 and SCR2. Targeted gene disruption of NTG1 produces mutant cells which are sensitive to hydrogen peroxide and menadione, indicating the importance of NTG1 for the repair of oxidative damage in yeast.[110] A gene homologous to *endo III* in *S. pombe* has been cloned and expressed.[111] The *S. pombe* endo III homologue (Nth-Spo) releases both thymine glycol and urea from DNA substrates and possesses AP-lyase activity. However, the *S. pombe* enzyme is more active than its bacterial counterpart toward the removal of urea and less active toward the removal of thymine glycol.[111]

Enzymes with substrate specificity similar to endo III have also been isolated from human and bovine cells.[107,108,112] Recently, a bovine homologue was identified, sequenced, and cloned by taking advantage of the bovine endo III's ability to form a stable covalent intermediate with substrate DNA in the presence of sodium borohydride (see section IV).[113] This enzyme is similar in terms of sequence and functional aspects to the *E. coli* counterpart. A human homologue (h*N*TH1) has also been cloned and sequenced,[101,102] and found to be active toward the removal of urea and thymine glycol, indicating a similar substrate specificity to that of endo III.

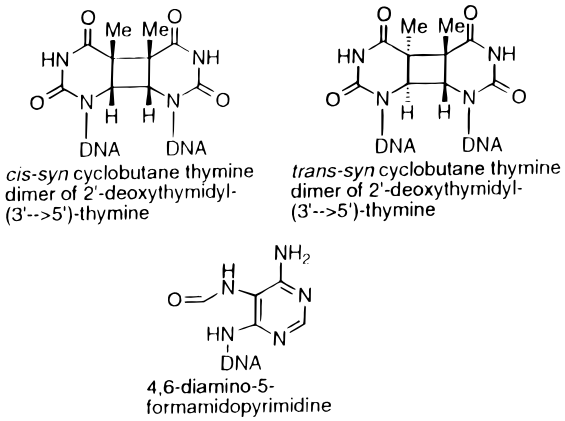

Figure 7. Narrow substrate-specificity of T4-endonuclease V.

Additional putative sequence homologues have been identified in other species, suggesting that an endo III-like protein is conserved across all phylogeny.[114]

A second enzyme, endonuclease VIII (endo VIII), has been isolated from *E. coli* and has been shown to have enzymatic properties similar to endo III.[115] This enzyme releases thymine glycol, dihydrothymine, β-ureidoisobutyric acid, and urea residues from DNA. Recently, the gene was cloned, sequenced, and overexpressed in *E. coli*.[116] These results show that the amino acid sequence of endo VIII is not similar to endo III. Indeed, endo VIII exhibits some homology to N-terminal and C-terminal portions of the FPG protein (section II.F). Interestingly, *E. coli* cells lacking functional endo III and endo VIII show a strong mutator phenotype and are hypersensitive to ionizing radiation and hydrogen peroxide, providing evidence for the role of these enzymes in the protection against oxidative DNA damage.[116]

E. Pyrimidine Dimer Glycosylases

T4-endonuclease V (T4-endo V) has been one of the most extensively studied BER enzymes and was the subject of a review in 1989.[117] This enzyme was originally identified by the observation that *E. coli* infected with bacteriophage T4 are able to release nucleotides containing pyrimidine dimers from UV-irradiated cells. This activity is the result of the bacteriophage T4 *denV* gene product, T4-endo V.[118−120,121] T4-endo V was later characterized thoroughly and shown to cleave the N-glycosyl bond of 5′-pyrimidines at dimers of adjacent pyrimidines with subsequent AP endonuclease activity to give phosphodiester bond cleavage between the dimerized pyrimidines.[122−125,110−113] Sequencing and overexpression of this enzyme have facilitated its characterization.[126] Enzymes with similar activity have also been isolated from *Micrococcus luteus* (*M. luteus* UV-endonuclease)[127−130] and *S. cerevisiae*.[131] The reports on the *S. cerevisiae* enzyme indicate that the activity is similar to T4-endo V; however, this enzyme has not been as extensively characterized.

T4-endoV and *M. luteus* UV-endonuclease are active upon all cyclobutane pyrimidine dimers; however, thymidyl(3′-5′)thymidine (T◇T) dimers (Figure 7) are the most prevalent in UV-irradiated DNA[132] and are the preferred substrates for both enzymes.[133]

Both T4-endo V and *M. luteus* UV-endonuclease are highly specific, with *cis-syn*-cyclobutane dimers as the primary substrates.[117,134] Using synthetic oligonucleotides containing structurally characterized thymidylyl(3′-5′)thymidine photoproducts, *trans-syn*-cyclobutane thymine dimers were found to be substrates for T4-endoV and were removed ~100 times slower than *cis-syn*-cyclobutane thymine dimers.[135] No activity was observed with T4-endo V or *M. luteus* UV endonuclease with oligonucleotides containing a (6−4) pyrimidine dimer or a Dewar dimer photoproduct.[134,135] A relatively new substrate for T4-endo V has been identified by monitoring the products released by T4-endo V from γ- or UV-irradiated DNA.[136] By using this method, 4,6-diamino-5-formamidopyrimidine (fapy-adenine) was shown to be removed by T4-endo V. Fapy-adenine excision by T4-endo V is ~1−3% of that for removal of UV irradiation induced *cis-syn*-cyclobutane dimers. The structures of identified substrates for T4-endo V are illustrated in Figure 7. The biological significance of the activity T4-endo V on fapy-adenine in DNA is presently unknown. An early report suggested that T4-endo V is also active upon a *purine* photoproduct, which was not fapy-adenine and may be a purine dimer.[137] However, this UV photoproduct was not fully characterized.

Interestingly, despite the fact that the substrate specificity and enzymatic properties of *M. luteus* UV-endonuclease are similar to T4-endo V, the two enzymes do not share sequence homology. In addition, recent reports indicate that there are two UV-endonucleases in *M. luteus* with masses of 30.4 and 17.1 kDa.[138] The larger *M. luteus* UV-endonuclease has sequence homology to *E. coli* endo III and MutY.[134]

F. The FPG Protein (Fapy Glycosylase or MutM)

In *E. coli*, a DNA glycosylase that excised imidazole ring-opened forms of N7-methylguanine (2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine or 7-methyl-Fapy-guanine) was identified and named Fapy glycosylase.[139] 7-Methyl-fapy-guanine lesions arise in DNA from methylation at N-7 of guanine and subsequent ring opening upon treatment with base (Figure 2).[140,141] These lesions block DNA synthesis and therefore are potentially lethal lesions in vivo.[140,142] The gene coding for the Fapy glycosylase was later identified and the overexpressed protein was found to have a tightly associated nicking activity at AP sites and was therefore renamed the FPG protein to be more inclusive about its activity.[143] The FPG protein also catalyzes the removal of 5′-terminal deoxyribose phosphates from DNA (dRpase activity) which are the products remaining after the action of AP endonucleases.[144] Other imidazole ring-opened substrates for the FPG protein have been identified and include fapy-guanine (2,6-diamino-4-hydroxy-5-formamidopyrimidine) and fapy-adenine (4,6-diamino-5-formamidopyrimidine), which are produced in DNA by ionizing radiation or photosensitization.[145,146] Aflatoxin-B_1-fapy adducts[147,148] and aminofluorene adducts at C-8 of guanine[136,137] are reported substrates for the FPG protein.[137] FPG has also been shown to remove other types of adducts formed by alkylation and subsequent ring opening, such as those formed with sulfur mustard, 7-(hydroxyethyl)-thioethylguanine.[149] Representative substrates for FPG are illustrated in Figure 8.

It was later determined that the FPG protein was identical to an *E. coli* enzyme responsible for the removal of 8-oxoguanine bases in DNA. This enzyme had been appropriately named 8-oxoguanine DNA glycosylase.[150,151] 7,8-Dihydro-8-oxo-2′-deoxyguanosine (OG) residues form in DNA from reactions of 2′-deoxyguanosine with oxidizing agents and ionizing radiation.[152] OG is one of the most common oxidative lesions in DNA and its presence results in DNA mutations.[153] Genetic experiments had previously

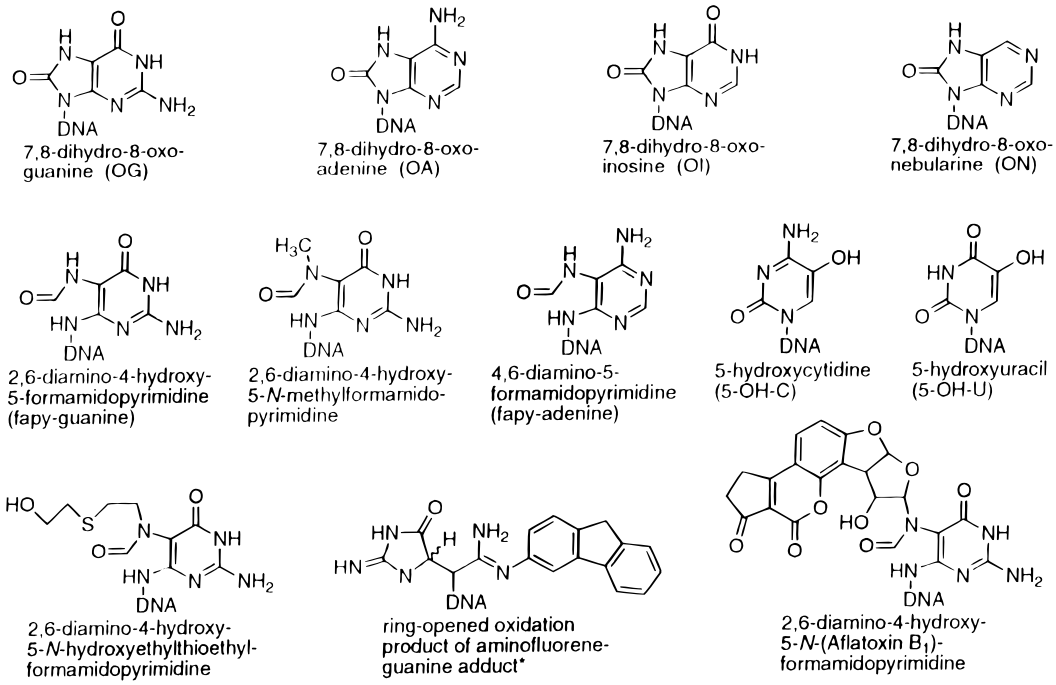

Figure 8. FPG protein substrates. *The structure of this adduct has not yet been confirmed conclusively.

identified a gene, *mutM*, which when inactivated by disruption, specifically increases G:C to T:A transversions.[154−156] Fapy lesions block DNA synthesis but would not be expected to cause the increased frequency of G:C to T:A transversions in a *mutM*$^-$ strain. However, in vivo and in vitro DNA replication of DNA containing OG lesions had shown that DNA polymerases can misincorporate adenine opposite OG,[157] which in subsequent replication events, results in conversion of a G:C base pair to a T:A base-pair. Furthermore, *E. coli* lacking a functional *mutM* gene are deficient in their ability to remove OG from DNA.[158] Cloning and sequencing of *mutM* indicated that it was identical to the *fpg* previously identified.[143,148,154,159] Thus, all four names "MutM", "Fapy glycosylase", "FPG", and "8-oxoguanine glycosylase" refer to the same protein. The importance of FPG in preventing mutations caused by oxidative stress has led to the proposal that OG is an important in vivo substrate.[151,155,156,160,161]

A variety of other substrates have been identified for the FPG protein (Figure 8); however, the importance of these substrates in vivo is unknown. Regardless, these additional substrates provide insight into the mechanism of DNA recognition by the FPG protein. For example, 5-hydroxycytosine and 5-hydroxyuracil have been found to be released by the FPG protein.[24,41] It was suggested that the FPG protein may be able to accommodate these small oxidized pyrimidines within its active site since it may be designed for larger oxidized purines. A feature of the oxidized substrates must be recognized by FPG since undamaged pyrimidines are not substrates. Another product of oxidative damage, 7,8-dihydro-8-oxo-2′-deoxyadenosine (OA), is removed by FPG, although extremely inefficiently.[151,162] Additionally, Tchou et al. synthesized a series of substrate analogues to test the requirements for recognition by FPG. This work uncovered two additional substrates for FPG, 7,8-dihydro-8-oxo-2′-deoxyinosine and 7,8-dihydro-8-oxo-2′-deoxynebularine which are both removed more efficiently than OA but significantly less efficiently than OG. The authors interpret these results as indicative of the importance of the *8-oxo* functionality in the substrate recognition by FPG. The ability of FPG to also remove fapy lesions and 5-hydroxypyrimidines may suggest that the presence of carbonyl or hydroxy functional groups can effectively mimic the 8-oxo functionality of OG.

The presence of OG in DNA can produce mutations in DNA due to misincorporation of 2′-deoxyadenosine opposite OG to form OG:A base pairs.[157] The action of FPG on OG in OG:A base pairs would increase the likelihood of mutations and therefore, it seems logical that FPG would be sensitive to the nature of the base opposite the OG. Biochemical studies have indicated that FPG is the least efficient toward OG in an OG:A base pair.[151,162,163] There is some disagreement regarding the magnitude of the rate differences between reported studies which may be due to the conditions used and the sequence context of OG. FPG also appears to bind less efficiently to OG:A base pairs than OG:C/G/T base pairs, a fact consistent with the biological role for FPG in preventing mutations caused by OG. OG:A base pairs are repaired by another DNA repair enzyme, MutY (vide infra). Very little work on FPG has addressed sequence context effects; however, one study has shown that FPG is slightly more active in G-rich sequences.[164]

G. Eukaryotic OG Glycosylases

A sequence homologue to the FPG protein in eukaryotic cells has not been identified but functional homologues have been isolated from yeast and mammalian cells. An OG glycosylase with an apparent molecular weight of 40 kDa was partially purified from cell extracts of *S. cerevisiae*.[153,165] The initial biochemistry on this enzyme was somewhat puzzling, since it showed a preference for an OG:G-containing duplex substrate rather than the corresponding OG:C duplex as had been observed with the FPG protein. Additionally, this protein appeared to act preferentially on fapy residues rather than OG in DNA. Later, two groups independently isolated, sequenced, and characterized an OG glycosylase from yeast which they both termed yOGG1.[166,167] The Boiteux laboratory used a genetic complementation method to isolate clones containing the yeast gene for a functional OG glycosylase,[166] while the Verdine research group used a chemical method which relied on the ability to trap a putative Schiff base intermediate (section IV).[167] The biochemical properties and sequence information indicate that both research groups had identified the same enzyme.[166−168] The substrate specificity of the yOGG1 is remarkably similar to the FPG protein. The substrate specificity is OG:C > OG:T ≫ OG:G > OG:A. Fapy residues are also released by the yOGG1 enzyme. The yOGG1 enzyme has also been reported to possess deoxyribophosphodiesterase (dRpase) activity that removes 2′-deoxyribose 5′-phosphate moieties at hydrolyzed 5′-AP sites in a manner similar to FPG.[169] In a Mg(II)-dependent fashion, yOGG1 also has the ability to hydrolytically remove *trans*-4-hydroxy-2-pentanal 5-phosphate from 3′-termini.[169] Sequence information on yOGG1 indicates that it does not share sequence homology to the FPG protein; rather, it bears high homology to *E. coli* endo III (see section III.B). The Verdine group also identified a second OG glycosylase (yOGG2) from *S. cerevisiae* which has a unique sequence and an unusual substrate specificity where OG:G > OG:A > OG:T ≈ OG:C.[167] The biochemical properties of yOGG2 appear to parallel those of the OG glycosylase enzyme originally isolated from *S. cerevisiae* cell extracts.

With knowledge of the yeast OG glycosylase (yOGG1) sequence, murine (mOGG1) and human (hOGG1) homologues were identified by a large number of laboratories.[5,170−176] In preliminary biochemical experiments with the mammalian OG glycosylases, the enzymes appear to be more selective for OG:C base-pair substrates over the alternative OG:N (N = T,G,A) substrates than the yeast counterpart.[5,170,173−175] The results of two laboratories indicate that the hOGG1 enzyme effects the removal of OG from OG:N base pairs where the activity is such that N = C >T ≫ G > A; however, with OG:N base pairs, where N = T, G, or A, no associated

strand cleavage due to AP lyase (β-elimination) activity was observed.[173,176] Cleavage by the hOGG1 protein at an abasic site was also found to be dependent on the opposite base.[173] There are slight discrepancies in the activity of mammalian OG glycosylases among the reported studies. Further characterization of these mammalian analogues will likely provide additional insight into these observations.

H. Mismatch-Specific Adenine Glycosylases: MutY and MYH

Mismatch repair in *E. coli* has been well-characterized and shown to be the results of *mutS*, *mutH*, and *mutL* gene products.[4,177] In this repair pathway, methylation at a GATC sequence distinguishes the template strand from the newly synthesized strand. The mismatch repair enzymes involved specifically remove the mismatched base in the unmethylated strand. However, in the late 1980s an activity was detected in *E. coli* cells which restored G:A mismatches to G:C matches and was insensitive to the methylation state of the template strand.[178,179] Additionally, Miller et al. identified a mutator locus in *E. coli* that generates G:C to T:A transversion mutations which they termed *mutY*.[180] It was later determined that the *mutY* gene product was an adenine glycosylase that removes the adenine base within G:A mismatches and thus is able to prevent mutations caused by G:A mismatches.[181−183]

In an interesting combination of genetics and biochemistry, Miller and Michaels later determined that MutY plays an important role in the prevention of DNA mutations caused by the oxidatively formed lesion OG.[155] They characterized an additional mutator locus *mutM* which encodes for the FPG protein.[154] Genetic studies indicated that the *mutY*$^-$ or *mutM*$^-$ strains have elevated rates of G:C to T:A transversion mutations; however, *mutY*$^-$*mutM*$^-$ double mutants have extremely high mutation rates which are significantly higher than the sum of the individual mutation rates.[184] This suggests that the two proteins work together to prevent G:C to T:A transversion mutations and the authors proposed that the primary function of the MutY protein in vivo may be the removal of adenine from OG:A base pairs. Biochemical studies indicated that MutY is an active adenine glycosylase toward adenine in OG:A base pairs.[160,184] The mutagenic effect of OG results from misincorporation of 2′-deoxyadenosine opposite OG by DNA polymerase to form stable OG:A base pairs within duplex DNA.[157] If left unrepaired, subsequent replication events result in G to T transversion mutations. Miller and Michaels have termed the MutY/MutM repair systems as the "GO" repair system which is illustrated in Figure 9.[144] The mutation rates for disabling the GO repair system are significantly higher than those obtained by disabling the mismatch repair system (*mutS*) or the polymerase III editing function (*mutD*). Thus, this is a critical repair system for preventing mutations caused by oxidative damage in *E. coli*.[184] A third member of the repair system to protect against OG incorporation in DNA is the dOGTPase, MutT. This enzyme eliminates dOGTP from the dNTP pool to prevent its misincorporation by DNA polymerase opposite adenine during DNA replication.[173]

MutY is rather unique among BER enzymes in exhibiting glycosylase activity toward a normal base within a mismatched base pair and this characteristic is shared only with the G:T-specific thymine glycosylase (TDG). MutY has also been shown to remove adenine from C:A mismatches in vivo[185] and in vitro.[184,186] The in vitro experiments indicated the efficiency of MutY's adenine glycosylase activity on C:A mismatches to be 5−35-fold lower than on G:A mismatches. A number of other substrates for MutY have been identified, most of which may not be important substrates in vivo, but do provide considerable insight into the recognition and enzymatic requirements of MutY.[187−189] In Figure 10, alternative bases that can be substituted into the OG:A base pair and retain activity by MutY are shown. The recognition and repair by MutY is sensitive to both bases in the damaged and/or mismatched base pair; however, in general, MutY is more sensitive to alterations in the adenine. A number of bases can substitute for OG, including OA, 8-methoxy-2′-deoxyguanosine, and inosine. The adenine can also be substituted by a variety of bases including inosine and 2-aminopurine. In all cases, MutY's ability to remove modified bases is dependent on the identity of the opposite base, with OG enhancing the glycosylase activity.

In the studies from laboratories working with MutY, there is a lack of agreement on the relative activity of MutY toward some of its substrates. For example, Lu et al.[187] and Manuel et al.[189] provide data which suggests that G:A-containing duplexes are more efficient substrates than OG:A-containing duplexes. Enzyme kinetics studies by Bulychev et al.[188] indicate the opposite trend. A possible origin for the different results may depend on the use of base treatment in the analysis of glycosylase activity. The reports from Bulychev et al.[188] use base treatment to analyze MutY's glycosylase activity while the other reports[187,189] do not. The inconsistent use of base treatment is due to the lack of agreement regarding whether MutY has an associated β-lyase activity to provide strand scission (section IV.D). If the base removal step and the strand scission reaction are *not* directly correlated for all of MutY's substrates, reliance on MutY to provide strand scission may underestimate the extent of glycosylase action. On the basis of the controversy, it seems prudent to use base treatment when monitoring the glycosylase activity of MutY to ensure strand scission. Another issue that is not taken into account when comparing MutY substrates is the possible effect of slow enzymatic turnover. MutY binds with significant affinity to the product produced after removal of the adenine from an OG:A-containing duplex.[160,190] The consequence of this is extremely slow turnover of MutY with OG:A-containing duplexes compared to G:A-containing duplexes and therefore glycosylase assays are strongly influenced by the amount of MutY enzyme used relative to the DNA duplex.[191] In addition, measurements of "k_{cat}" are dominated by the rate of enzyme

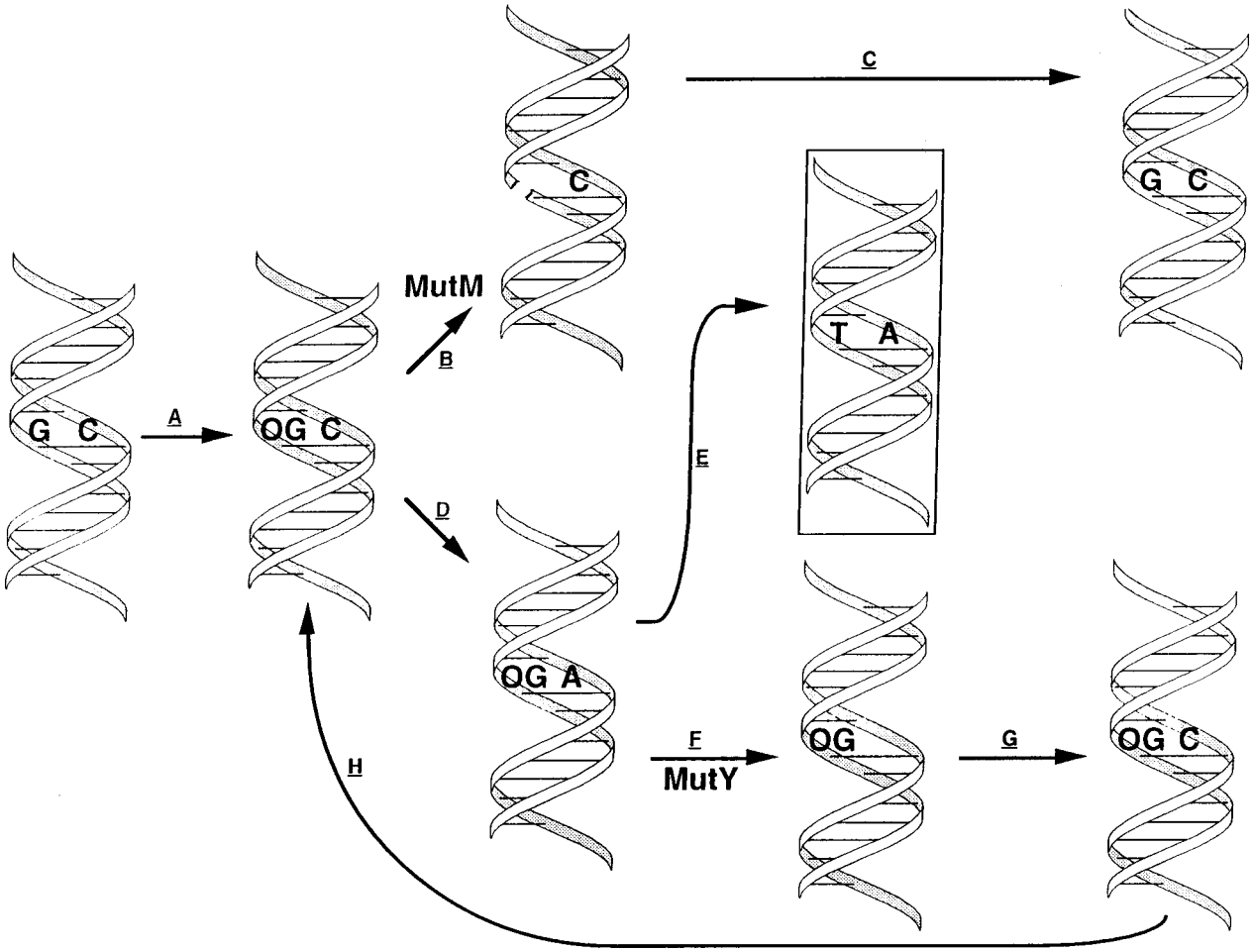

Figure 9. "GO" repair system involving MutM (FPG) and MutY. Oxidative damage to a G:C base pair results in the formation of 1,8-dihydro-8-oxo-2′-deoxyguanosine, generating an OG:C mispair (A). MutM removes the oxidized guanine base, and cleaves the DNA strand at the AP site (B). Cellular enzymatic repair (polymerases/ligases) complete the repair process by reestablishing the original G:C base pair and sealing the DNA backbone (C). If the OG:C mispair is allowed to undergo a cycle of DNA replication, adenine will be paired with the OG of the parent strand (D). The resulting OG:A mispair is a substrate for MutY, which removes the misincorporated adenine (F). Cellular repair machinery are recruited to regenerate the OG:C mispair (G), effectively recycling the duplex to MutM substrate (H) for eventual regeneration of the original G:C base pair. In the absence of MutY, the OG:A mispair can eventually lead to permanent G:C to T:A transversion mutations (E). (Adapted from ref 184.)

dissociation from the product rather than the intrinsic rate of the chemical reaction. In single-turnover experiments, the rate for MutY's adenine glycosylase activity with an OG:A-containing duplex is much faster than the corresponding G:A duplex.[191]

Qualitatively, the K_d values reported for OG:A- vs G:A-containing substrates also support OG:A-containing duplexes as the preferred substrates; however, the scatter in reported K_d values is also likely due to problems associated with enzymatic turnover and the presence of different amounts of substrate relative to product during the binding measurements.[190] In studies using a noncleavable 2′-deoxyadenosine analogue, 2′-deoxyformycin A (F), MutY was found to bind ~30-fold more tightly to OG:F than G:F, providing additional biochemical evidence for OG:A as the preferred substrate for MutY.[190]

Sequencing of the *mutY* gene indicates that the MutY protein has high sequence homology to endo III.[192] Limited proteolytic treatment of MutY with trypsin or thermolysin has been shown to produce stable truncated versions of MutY of 26 or 25 kDa, respectively.[193,194] In these preparations, the C-terminal portion which is not homologous to endo III has been removed. This proteolytic treatment provides a nearly "endo III"-sized MutY fragment. Cloning and overexpression of the 26 kDa MutY fragment (p26) has also been reported.[189] Biochemical characterization has shown that the enzymatic properties of the truncated versions are similar to native MutY. In the case of the 25 kDa sized MutY, reduced binding to an OG:A substrate was reported and it was proposed that the C-terminal portion may be responsible for specific recognition of OG. However, similar results were not reported with the p26 MutY fragment. On the basis of the inconsistent results on the substrate specificity (G:A vs OG:A) of native MutY (vide supra), these results may also be influenced by enzymatic turnover and the method of analysis as has been observed with wild-type MutY.[191]

A human homologue to MutY has also been identified,[195] and its gene was cloned and sequenced.[47,196] The human homologue, MYH, has 41% identity to

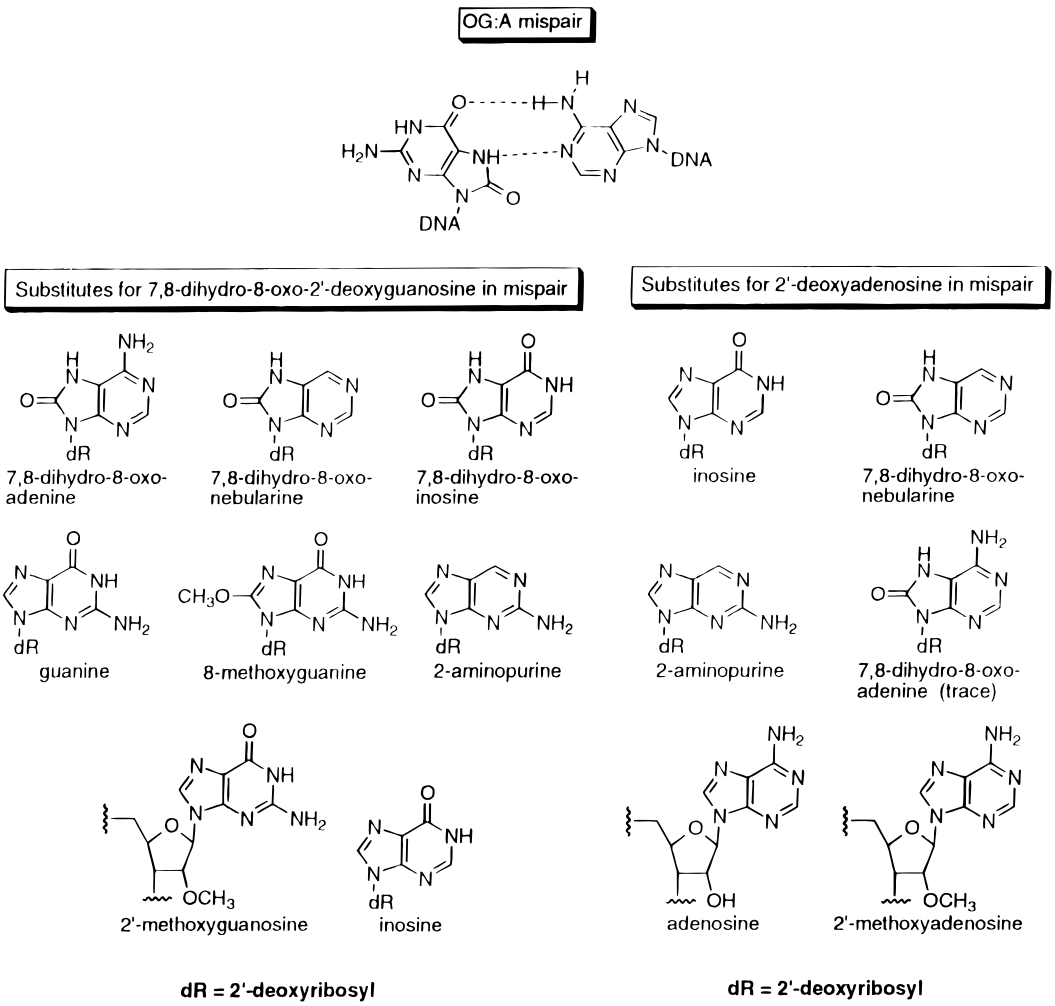

Figure 10. MutY substrate OG(syn):A(anti) mispair, and respective substituted compounds which can also form mispair MutY substrates.

the *E. coli* MutY[196] and a substrate specificity similar to the bacterial enzyme.[184,195] Interestingly, MYH was reported to be 2-fold more active on G:A mismatches than OG:A mismatches, although it reportedly binds more tightly to OG:A than to G:A mismatches.[195] These results may also be influenced by enzymatic turnover and the method of analysis as has been observed with wild-type MutY (vide supra).

III. Insights into Damage Recognition by BER Glycosylases from Structural and Biochemical Studies

A. Uracil-DNA Glycosylase: Evidence for a Nucleotide-Flipping Mechanism

Uracil-DNA glycosylases have been shown to be extremely specific for uracil and this led to the prediction that these enzymes may possess a base-specific pocket that preferentially accommodates uracil. Crystallographic studies have shown that both the human[197] and HSV-1[198] UDGs possess a specific uracil-binding pocket comprised of basic amino acid side chains highly conserved among the known UDG primary structures. The pocket was identified in these two structural studies by soaking the crystals with either 6-aminouracil or uracil. In both cases, this pocket exists in the center of a narrow channel along the enzyme surface. The channel is cone-shaped to allow for dsDNA binding at the wide end, but only with significant conformational change at the narrower end.[186] Thus for uracil to reach the binding pocket within the active site, it must adopt an extrahelical conformation. This flipping-out process allows uracil to come into direct contact with conserved functional groups positioned in the binding pocket. The binding of a uracil-containing substrate in an extrahelical fashion by UDG is also consistent with the 2−3-fold greater rate of glycosylase activity toward ssDNA over dsDNA; a significant amount of energy must be spent to disrupt the base-pairing and stacking interactions of uracil in dsDNA.

The crystal structures of the human and HSV-1 UDGs and their complexes with inhibitor bases strongly suggested a flip-out type of mechanism. Indeed, the nucleotide-flipping mechanism was clearly illustrated in the subsequent crystal structure of human uracil-DNA glycosylase bound to DNA.[199,200] In the crystal structures of human and HSV-1 UDGs, a conserved leucine residue at position 272 located above the uracil-specific pocket was implicated to play a role in uracil recognition and base flipping.[186] Using site-directed mutagenesis on the hUDG enzyme, a doubly mutant form (L272R/D145N) was

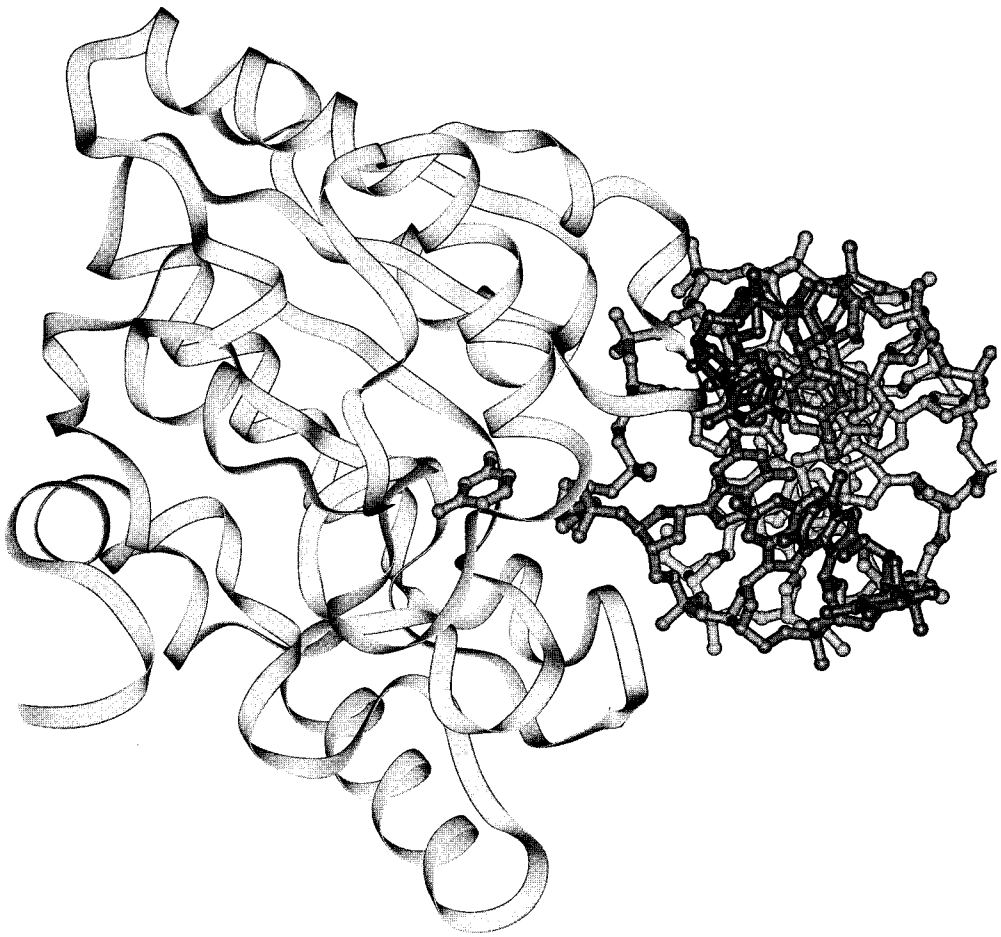

Figure 11. Representation of the L272R/D145N form of human uracil-DNA glycosylase bound to a uracil-containing 10-base-pair duplex (ref 199). In this representation of the hUDG-DNA complex, the DNA helix is viewed end-on. The Arg 272 residue penetrates the helix pushing it into an extrahelical conformation. The C1′−N glycosylic bond is cleaved in this structure; however, the uracil remains bound in a uracil-specific pocket. (The figure was generated from coordinates provided by the authors.)

created in which Leu 272 was converted to an arginine and Asp 145 to an asparagine. This mutated enzyme exhibited an increased binding affinity for DNA and reduced activity for uracil N-glycosylic bond cleavage. The L272R/D145N enzyme was cocrystallized with a 10 base pair duplex substrate containing a centrally located G:U base pair. The resulting electron-density map provided a structure representative of the enzyme−product complex since the N−C1′ glycosylic bond is cleaved; however, the uracil base remains in the base-specific pocket and the abasic site is flipped out of the DNA helix (Figure 11). Thus, it seems reasonable that this hUDG−product complex reasonably approximates the enzyme−substrate complex. The structure reveals that Arg 272 is inserted into the DNA from the minor groove and its movement has effectively "pushed" the uracil base out of the helix on the major groove side. The entire uridine unit is rotated 180° with respect to normal B-form DNA, and therefore the authors suggest that the more accurate word to describe this occurrence is "nucleotide flipping" rather than the popular term, "base flipping". Specific contacts between the enzyme and the sugar−phosphate backbone adjacent to uracil promote DNA backbone compression which also favors the extrahelical orientation of the uracil nucleotide. Importantly, there are also contacts between the enzyme and DNA which provide for specific recognition of the flipped-out 5′-phosphate, deoxyribose, and uracil. These contacts supply the "pull" to capture the flipped-out nucleotide and the authors suggest that it is the coordinated action of the "push" by the Leu 272 (Arg in the mutant) and phosphate compression, with the "pull" provided by the specific pocket which results in the efficient recognition of uracil in DNA.

In the case of both the HSV-1 and human UDG, specificity for uracil is brought about by van der Waals interaction with Tyr and Phe side groups, as well as electrostatic interactions with the polar atom arrangement specific to uracil, namely O2, O4, N3, and C5[201] as illustrated in Figure 12. In the hUDG-DNA structure, the uracil is stacked on Phe 158 and specific hydrogen bonds are made among the O4, N3, and O2 with Phe 158 NH, Asn 204 side chain, 144 and 145 backbone amides, and His 268 imidazole. The contacts between the Asn 204 side chain and the protein backbone hydrogen bonds with uracil confer specificity for uracil over cytosine. This has been demonstrated by replacement of Asn 204 in hUDG with an aspartate residue which modifies the enzyme activity of the mutated enzyme to bestow it with glycosylase activity toward cytosine in DNA.[191,202] In the hUDG complex, tyrosine 147 is tightly positioned

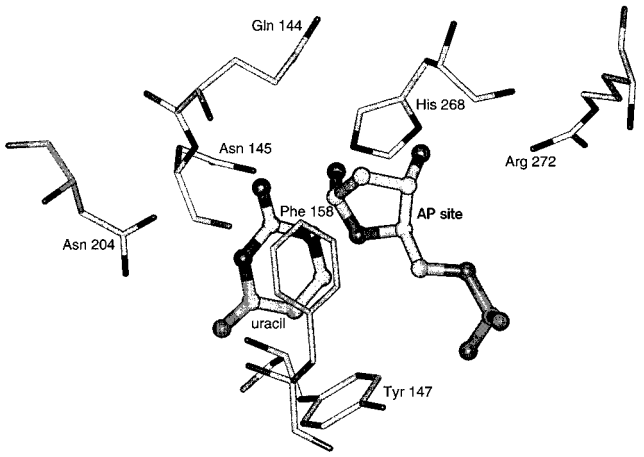

Figure 12. Active site of human uracil-DNA glycosylase containing a bound uracil (ref 199). The uracil base and nucleotide are illustrated in a ball-and-stick mode (thick lines) while the amino acid side chains are in stick (thin line) representation. Many of the protein backbone residues and some amino acid side chains have been deleted to illustrate the specific amino acids involved in uracil recognition as described in the text. (The figure was generated from coordinates provided by the authors.)

adjacent to the C5 of uracil which precludes binding of thymine. The role of Tyr 147 has also been established by its replacement with alanine, cysteine, or serine. The enzymes mutated at position 147 gain the ability to act as thymine-DNA glycosylases.[202] These structures also provide some insight into the substrate specificity of UDG and the surprising ability of UDG to remove the oxidized pyrimidines, isodialuric acid, and 5-hydroxy-uracil from DNA.[39,41] This can be rationalized on the basis of the crystallographic data, since these oxidized bases are the only ones which could reasonably fit into the uracil binding pocket.

B. DNA Recognition Motifs in a BER Superfamily

1. Helix–Hairpin–Helix and GPD Motifs

The crystallographic characterization of endo III and AlkA (3-methyladenine glycosylase II) provides exquisite details on features that provide specific recognition of their respective substrates and identifies new structural motifs that appear to be common among a large number of BER enzymes. Endo III and AlkA belong to a superfamily of BER enzymes that have a similar three-dimensional fold.[167,203−205] The hallmark of this class of BER enzymes is the presence of helix−hairpin−helix (HhH) and a Gly/Pro-rich stretch with nearby Asp (GPD) motifs. Sequence alignments of representative members of this superfamily is illustrated in Figure 13A. The HhH motif is found in a variety of DNA binding proteins and likely represents a non-sequence-specific DNA recognition motif.[206] The hairpin portion of the HhH motif was identified as the location of binding of the damaged base thymine glycol in the original structural studies on endo III.[207] With the aid of site-directed mutagenesis, Lys 120, which is positioned at the N-terminus of one of the helices in the helix−hairpin−helix motif, was identified as important for the catalytic activity of endo III. Site-directed mutagenesis also implicated Asp 138 as a critical residue for the activity which lies in the interdomain groove of endo III, but just outside of the HhH motif. The overall structural similarity between AlkA and endo III was somewhat surprising since the enzymes were not predicted to be similar on the basis of sequence homology.[203,204] AlkA has three structural domains, two of which can be overlaid onto the structure of endo III. The structural homology between the two enzymes is particularly strong in the helix−hairpin−helix motif. Residues in the interdomain cleft of AlkA and endo III are critical for enzyme activity. For example, Asp 238 in AlkA aligns with Asp 138 in endo III in sequence alignments (Figure 13A) and mutation of the aspartic acid in both enzymes abolishes enzyme activity. An aspartate residue has been found to be highly conserved at this position among the entire superfamily and led to the addition of the Gly/Pro-rich Asp motif as a critical catalytic motif for this class of BER enzymes. This BER superfamily includes MutY, *M. luteus* UV-endonuclease, and the OGG1 proteins from yeast and mammals (Figure 13A).

2. Fe–S Cluster Loop Domain of Endo III and MutY

A unique feature of endo III is the presence of a $[4Fe\text{-}4S]^{2+}$ cluster,[208] a metal site more commonly found in electron-transfer proteins.[209−211] In endo III, the observation that the cluster is resistant to oxidation or reduction has led to the suggestion of a structural role for the $[4Fe\text{-}4S]^{2+}$ center in endo III rather than a role in electron transfer.[197] Endo III has four cysteine residues with a spacing (Cys-X_6-Cys-X_2-Cys-X_5-Cys) distinct (Figure 13B) from that observed for the ligands to the cluster in other Fe−S proteins, indicating that endo III is a prototype for a new class of Fe−S proteins involved in DNA repair. In the crystal structure of endo III (Figure 14), these four cysteines coordinate the $[4Fe\text{-}4S]^{2+}$ cluster at the surface of the protein near the C-terminus to form a distinct metal cluster binding domain which has been referred to as the $[4Fe\text{-}4S]^{2+}$ cluster loop (FCL) domain.[196] Replacement of Lys 191, positioned in a solvent-exposed loop of FCL domain, by glutamic acid via site-directed mutagenesis resulted in a >100-fold increase in K_m.[194] This result provides compelling evidence that the FCL domain of endo III is involved in DNA binding. The conservation of cysteine ligand positioning and identity of amino acids in this region in MutY, suggests the presence of a $[4Fe\text{-}4S]^{2+}$ center. Indeed, a number of the BER superfamily enzymes also appear to contain the FCL domain (Figure 13B). These enzymes include *M. luteus* UV endonuclease[134] and a G:T-specific thymine glycosylase from *Methanobacterium thermoautotrophicum* THF[212] in addition to some MutY and endo III homologues. In the case of *S. cerevisiae* endo III homologues, there appears to be one homologue which contains the Fe−S binding motif and one that does not.[109] In the case of the *S. pombe* endo III homologue, the Fe−S

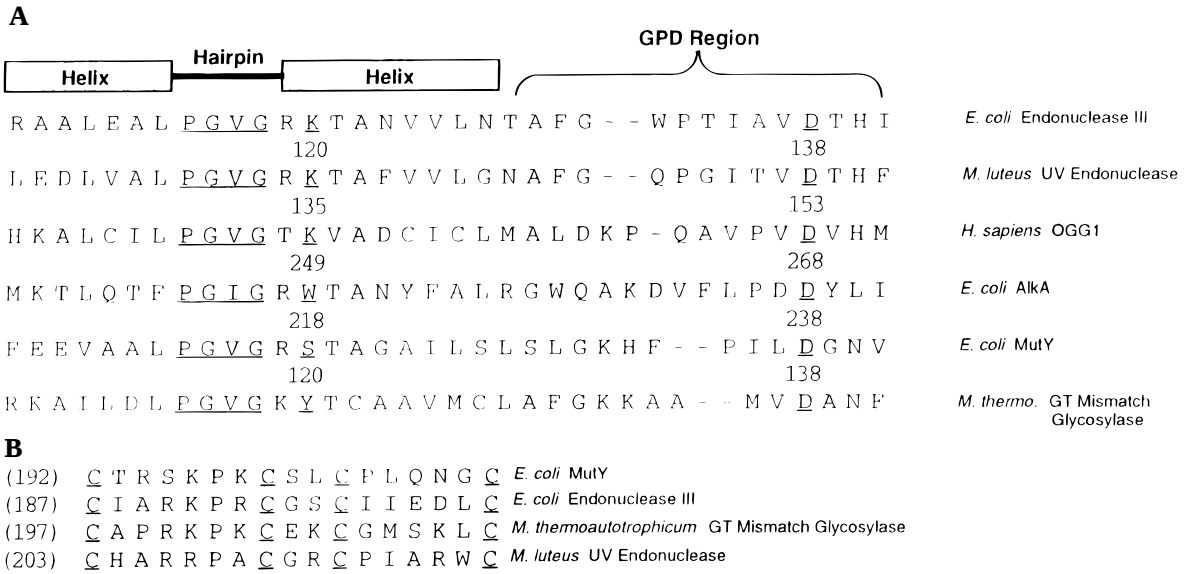

Figure 13. Sequence alignment of members of a BER superfamily (A) Active-site HhH and GPD motifs conserved among the BER superfamily DNA glycosylases. Six representative enzymes are illustrated. Alignment was made on the basis of previously published alignments indicating the presence of a BER superfamily (refs 5 and 167). (B) Alignment of FCL domains of a subset of the BER superfamily.

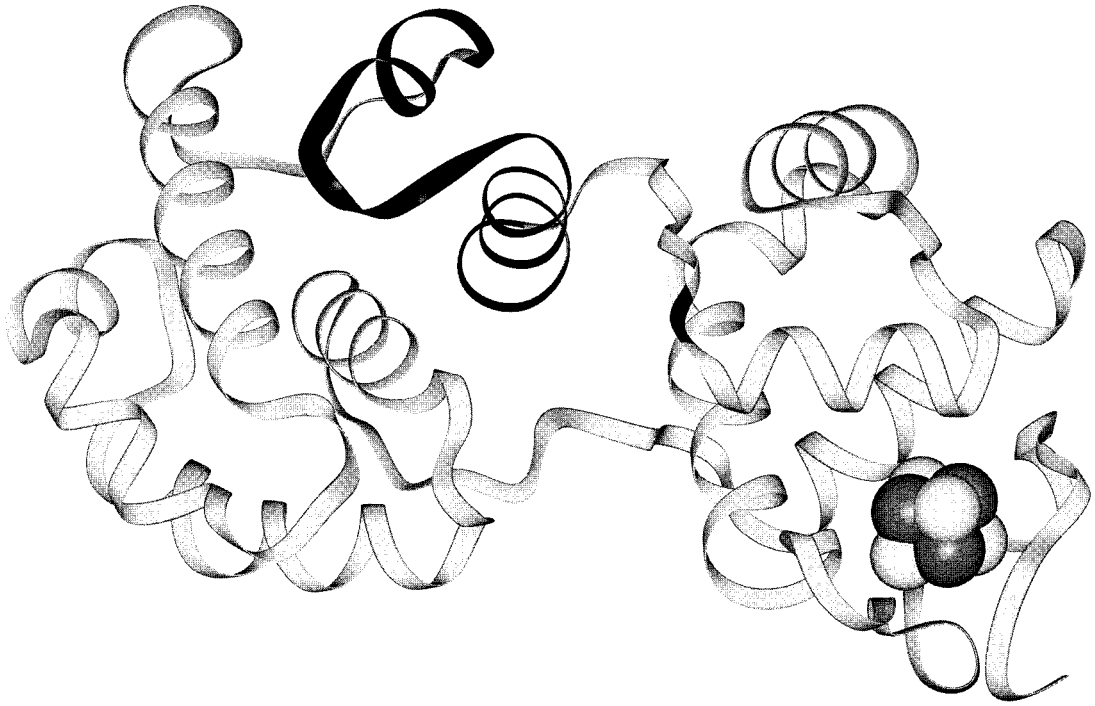

Figure 14. Three-dimensional structure of endonuclease III from X-ray crystallographic data. The iron and sulfide ions of the [4Fe-4S] cluster are depicted with space-filling representation. Random coil and turn portions of endonuclease III's polypeptide chain (including the cysteine ligands to the cluster) are depicted as ribbons. The helix−hairpin−helix and Asp 138 are shaded in black. (The figure was generated using the Brookhaven PDB file ABK2.)

motif appears to be present, although the spacing of cysteine ligands has been altered by addition of two amino acids between the last two cysteine residues (Cys-X_6-Cys-X_2-Cys-X_7-Cys).[111] Thus, there appears to be a subclass within the BER superfamily containing the FCL domain.

The importance of $[4Fe\text{-}4S]^{2+}$ cluster assembly in MutY folding has recently been investigated.[213] Denatured and metal-free MutY can refold in the presence of ferrous and sulfide ions to yield active MutY. This indicates that the cluster can self-assemble and that this process is facile in vitro. Interestingly, circular dichroism spectra of MutY refolded with and without assembly of the $[4Fe\text{-}4S]^{2+}$ are essentially identical, suggesting that assembly of the cluster is not required for initiation of global folding of MutY. In addition, the $[4Fe\text{-}4S]^{2+}$ cluster does not appear to contribute to the overall thermal stability of MutY as determined by circular dichroism thermal denaturation experiments. However, refolded forms of MutY which lack the $[4Fe\text{-}4S]^{2+}$ cluster do not possess detectable adenine glycosylase

activity. Furthermore, they do not bind specifically to substrate DNA and exhibit only weak nonspecific binding to DNA. Interestingly, these inactive folded MutY samples regain activity upon assembly of the $[4Fe\text{-}4S]^{2+}$ cluster, indicating a superficial location for the cluster. More importantly, these results show that the presence of the $[4Fe\text{-}4S]^{2+}$ cluster is critical for specific substrate recognition. The ability of the $[4Fe\text{-}4S]^{2+}$ cluster in MutY to readily self-assemble even after most of the protein has folded *may* be a property which is used in vivo to regulate the enzymatic properties of MutY. The possibility that Fe−S proteins participating in oxidatively damaged DNA repair may respond to the oxidizing conditions and metal ion concentrations of the cell by assembly or disassembly of an Fe−S cluster should not be overlooked.

The presence or absence of the FCL domain may be related to the substrate specificity of members of this BER superfamily. Cooperative interactions of the HhH-GPD motif and other DNA binding motifs (like the FCL domain) may fine-tune the active site and DNA binding surface for recognition of a specific type of damaged base or base pair. Thus, the presence of the $[4Fe\text{-}4S]^{2+}$ may be used to achieve specific substrate recognition by MutY or endo III. The lack of the $[4Fe\text{-}4S]^{2+}$ cluster domain in the AlkA enzyme may be significant in altering the access to the active site to achieve its wide substrate specificity. This may be somewhat analogous to the truncation of TDG which altered its ability to recognize thymine.[47] In the hOGG1 protein, Arai et al.[171] identified a putative Cys_2His_2-type zinc finger motif adjacent to the HhH and they suggest that this potential DNA binding motif may function in concert with the HhH in a manner analogous to the FCL motif in endo III and MutY. Notably, the FPG protein which is more distantly related to the BER superfamily[14] has a Cys_4 zinc finger motif at its C-terminus.[214,215] The presence of the zinc finger motif in the FPG protein is critical for DNA binding and the glycosylase/lyase activity. This again suggests that different DNA binding motifs may coordinate to effect DNA recognition. Further elucidation of the relationship between the various DNA binding motifs in the BER superfamily will be aided by a cocrystal structure of a BER superfamily enzyme bound to substrate DNA.

3. Substrate Recognition in the BER Superfamily: Endo III and AlkA

The structural studies on endo III[205] and AlkA[203,204] explain the disparate substrate specificity of these two enzymes that both contain the HhH and GPD motifs. Modeling with both structures suggests that the interdomain cleft in endo III and AlkA could accommodate a base if it is rotated out of the double helix. In the case of AlkA, the active site cleft is lined with electron-rich aromatic rings from amino acids Phe 18, Trp 218, Tyr 222, Trp 272, and Trp 273. These electron-rich side chains are nicely poised for π-donor/acceptor interactions with electron-deficient bases resulting from alkylation. The width (10 Å) and length (20 Å) of the active site cleft is sufficiently large to accommodate the various substrates processed by AlkA. The large surface area of aromatic rings provided by the active site may explain the affinity for extended aromatic structures, such as ethenoA, that are recognized efficiently by AlkA, in addition to those that are positively charged and electron-deficient such as 3-methyladenine. The ability of AlkA to remove products of hydrolysis and oxidation is difficult to rationalize on the basis of the structure. This may result from the presence of a few charged residues (Arg 22 and Asp 237) in the spacious active-site cleft which results in low levels of activity of AlkA toward these substrates. In contrast, the proposed active-site cleft of endo III is filled with water and contains a large number of polar residues (e.g., Arg 184, Asn 123, Arg 108, Glu 23, and His 176) which could participate in hydrogen bonds directly or with the aid of an intermediate water molecule with oxidized thymine and cytosine bases. A comparison of the two active-site structures is illustrated in Figure 15.

C. Indirect Recognition of DNA Damage: T4-EndoV Bound to a Pyrimidine Dimer Containing DNA Duplex

The idea that DNA repair enzymes recognize a flipped-out base was well-accepted before a crystal structure of a DNA repair enzyme bound to DNA was determined. The structure of a mutant T4-endo V (E32Q) bound to a pyrimidine dimer substrate was the first glimpse of a DNA repair enzyme complexed with its substrate DNA, and held some interesting surprises regarding the mechanism of DNA damage recognition.[216] The replacement of a glutamate with a glutamine in the E23Q enzyme disrupts the enzyme's ability to initiate glycosylase activity, but preserves native enzyme binding affinity for pyrimidine dimer substrates. In the structure of the E23Q enzyme with a 13 base pair thymine dimer containing duplex illustrated in Figure 16, one of the most surprising features is the large distortion in the DNA duplex at the position of the T◇T dimer. The DNA is severely kinked and the adenine base opposite the 5′-T of the T◇T dimer is flipped out of the DNA duplex and accommodated within a protein cavity. The kinking and flipped-out adenine provide for the formation of a large cavity in the DNA near the T◇T dimer, facilitating the enzyme's access to the damaged site. The protein conformation in the complex is altered minimally compared to the free enzyme structure,[217] and in particular, relative to the changes in the DNA structure. The enzyme approaches from the minor groove of the helix and there are extensive interactions between the protein and the phosphate backbone which stabilize the bent and flipped-out conformation as illustrated in Figure 17. Interestingly, there are no direct hydrogen-bonding interactions of the enzyme with the pyrimidine dimer itself; however, there are extensive interactions at the pyrimidine dimer with the deformed phosphate backbone which is shortened relative to a normal backbone.

This cocrystal structure indicates that important factors in recognition of damaged DNA are *indirect*

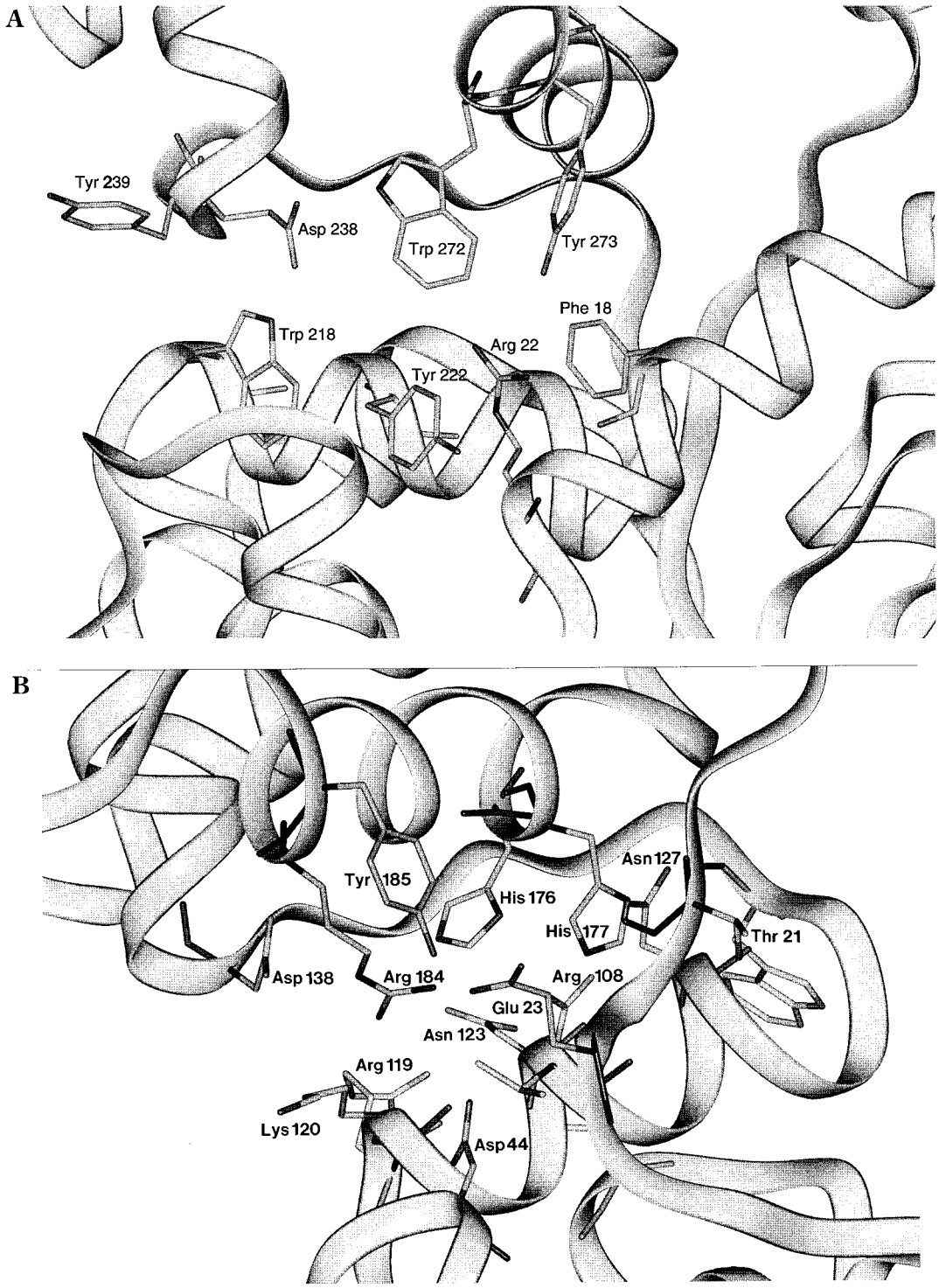

Figure 15. Comparison of proposed active-site clefts of AlkA and endo III. (A) The proposed active site cleft of AlkA indicates that the cleft is rich in electron-rich side chains poised for interactions with electron-deficient substrates through π-donor/π-acceptor interactions. This figure was generated from coordinates provided by the authors (ref 204). (B) The proposed active-site cleft of endo III. Note Asp 138 and Lys 120 at the mouth of the cleft, and the large number of charged amino acids. These charged residues are nicely poised for interactions with oxygen-rich bases. (The figure was generated using Brookhaven PDB file ABK2.)

and include the deformability of the DNA duplex at the damaged site and the presence of nonstandard DNA structural properties adjacent to the location of damaged bases. In this case, the large kink in the DNA is induced by T4-endo V binding but is likely facilitated by the presence of the poorly stacked T◇T dimer. Similarly, the ability to flip-out the adenine is facilitated by the deformation of the duplex and presence of nonoptimal Watson−Crick base-pairing between the adenines and thymines of the T◇T dimer. A surprising feature of this structure is the flipping out of the complementary base rather than the damaged base, as in the hUDG-DNA structure (Figure 12). However, in both cases, the flipped-out

Figure 16. Three-dimensional structure of T4-endonuclease V bound to a thymine dimer containing duplex (ref 216). T4-endonuclease V is shown in a dark ribbon presentation while the DNA duplex is represented in light gray stick presentation. An adenine opposite the thymine dimer is flipped out of the DNA helix and is imbedded in a T4-endonuclease V pocket. Note the large amount of bending of the DNA helix. (The figure was generated using Brookhaven PDB file VAS1.)

nucleotide serves the same role in providing access of the enzyme to the C1′−N bond. An additional feature in this structure is the extensive hydrogen bonding of the enzyme to the phosphate backbone and recognition of its deformity rather than direct interactions with the damaged base. This is in striking contrast to the hUDG-DNA structure where the enzyme specifically recognizes the damaged uracil base; instead, T4-endo V appears to recognize structural and thermodynamic properties associated with the T◇T dimer modification of the DNA helix. Thus, it appears likely that both direct and indirect mechanisms of damaged base recognition will be employed depending on the responsibilities of the specific BER enzyme.

D. Mechanisms Facilitating Location of DNA Damage

A daunting task of BER enzymes is that they must *locate* their target base-damaged site within the context of a large amount of undamaged DNA. An important feature of DNA binding proteins is a positively charged surface for favorable electrostatic interactions with the negatively charged phosphate backbone. These positively charged surfaces allow for significant affinity for nontarget DNA. Continued nonspecific association allows for sliding along the DNA helix for location of the requisite target sequence or damaged DNA base.[218,219] This sliding or scanning phenomenon is used by a number of restriction enzymes and transcription factors.[218] This method facilitates location of the target sequence or damage by reducing the dimensionality of the search to one dimension along the DNA helix, significantly decreasing the area which must be searched compared to a three-dimensional diffusive mechanism.[219] In the case of BER glycosylases, these enzymes presumably scan the helix until structural defects or deformations associated with DNA damage or mismatches are located. This type of mechanism is used by T4-endo V[220,221] and UDG [44,222] and is believed to be a general property of BER glycosylases.

The processive nature of T4-endo V was first illustrated by its ability to rapidly cause double-strand breaks in UV-irradiated plasmid DNA.[220] In these experiments, T4-endo V action resulted in the appearance of form III (linear) DNA at reaction times when significant amounts of unreacted form I (su-

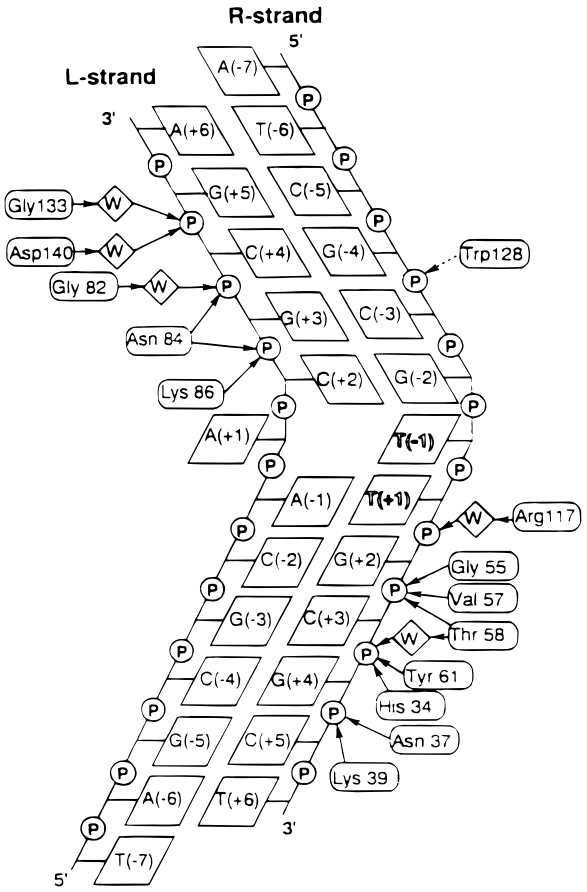

Figure 17. Schematic representation of polar interactions between the T4-endonuclease V and the DNA phosphate backbone surrounding the thymine dimer. The pyrimidine dimer strand is denoted R, while the complementary strand is denoted L. (Reproduced with permission from ref 216. Copyright 1995 Cell Press.)

percoiled) remained. This observation strongly suggested that form III DNA was produced by processive nicking events of T4 endo V, as opposed to a distributive mechanism, in which the complete conversion of form I to form II (nicked) DNA would be observed prior to formation of form III DNA. The electrostatic nature of this interaction is indicated by dependence upon the salt concentrations and the type of salt (chloride versus glutamate) used in the buffer.[221] Evidence of the important biological significance of the processive sliding mechanism for T4-endo V has been provided by using site-directed mutagenesis to produce mutated enzymes which retain native enzyme catalytic properties in vitro and in vivo but have impaired nontarget scanning ability.[223] These mutated enzyme forms, when expressed in repair-deficient *E. coli*, have a reduced ability to confer UV resistance to the cells. Thus, the wild-type enzyme's ability to function in the repair of UV-damage in vivo appears to be influenced by the ability to locate the target using an electrostatic scanning mechanism.[223]

UDG has also been shown to use a processive scanning mechanism on uracil-containing plasmid DNA.[222] However, two laboratories have obtained conflicting data on the processive versus distributive nature of UDG's action using uracil-containing oligonucleotides ligated to form concatemeric polynucleotides with a defined (but different) uracil spacing.[44,224] In these experiments, the rapid appearance of short oligonucleotide fragments can occur only if there is a processive nature to UDG's action. The most recent study suggests that the differences in the results on the polynucleotides may be dependent on the spacing of uracil residues and the sequence context.[44] The experiments which indicated a distributive mechanism for UDG used a sequence environment that was a suboptimal substrate for UDG while the other set of experiments in favor of a processive mechanism were performed on a substrate with a sequence that is optimal for UDG. The possible effects of the sequence environment on the ability of DNA repair enzymes to locate DNA damage or mismatches deserve further inquiry and may shed light on the remarkable efficiency of these enzymes.

However, even with the efficiency afforded by employing scanning mechanisms, it is still difficult to imagine how the BER glycosylases, which recognize such a diverse set of damaged bases and mismatches, do so with such proficiency. In the case of DNA base damage producing significant structural alteration, the BER enzyme may scan the DNA until a structural abnormality is encountered. However, many targets for BER enzymes do not result in significant structural distortion or thermodynamic instability. For example, the substrate OG:A base pair for MutY is stable and does not significantly distort the DNA helix.[225,226] Recently, some researchers in the field have suggested that BER glycosylases may take a more active role in detecting nondestabilizing base damage. For example, Verdine and Bruner proposed a mechanism in which the enzyme first acts by extruding a base at a location remote from the target base and then uses a processive extrusion mechanism to propagate the extrahelical base along the DNA until the target base is located.[227]

IV. Mechanistic Considerations

A. Classification of Glycosylases vs Glycosylase/Lyase Enzymes

Generally, the BER enzymes can be classified as either monofunctional glycosylases or bifunctional glycosylase/AP lyases. This is illustrated in Figure 18. The monofunctional DNA glycosylases comprise a large portion of the enzymes involved in base-excision repair. Their primary catalytic function is cleavage of the C1′−N glycosylic bond, resulting in an AP site. Such sites render the phosphodiester backbone vulnerable to scission in basic conditions via β-elimination (Figure 18). Activity in these enzymes is monitored either by treatment with base (NaOH, piperidine) and subsequent monitoring of strand scission by PAGE or by monitoring directly the release of the free base via HPLC or GC/MS. For example, the activity of UDG is often monitored by HPLC analysis of [^{3}H]uracil base release. Any DNA strand scission activity detected in enzymes classified as monofunctional DNA glycosylases is severalfold slower than the glycosylase step. This criterion distinguishes the monofunctional DNA glycosylases from those in the bifunctional glycosylase/lyase category where the removal of the base and resulting

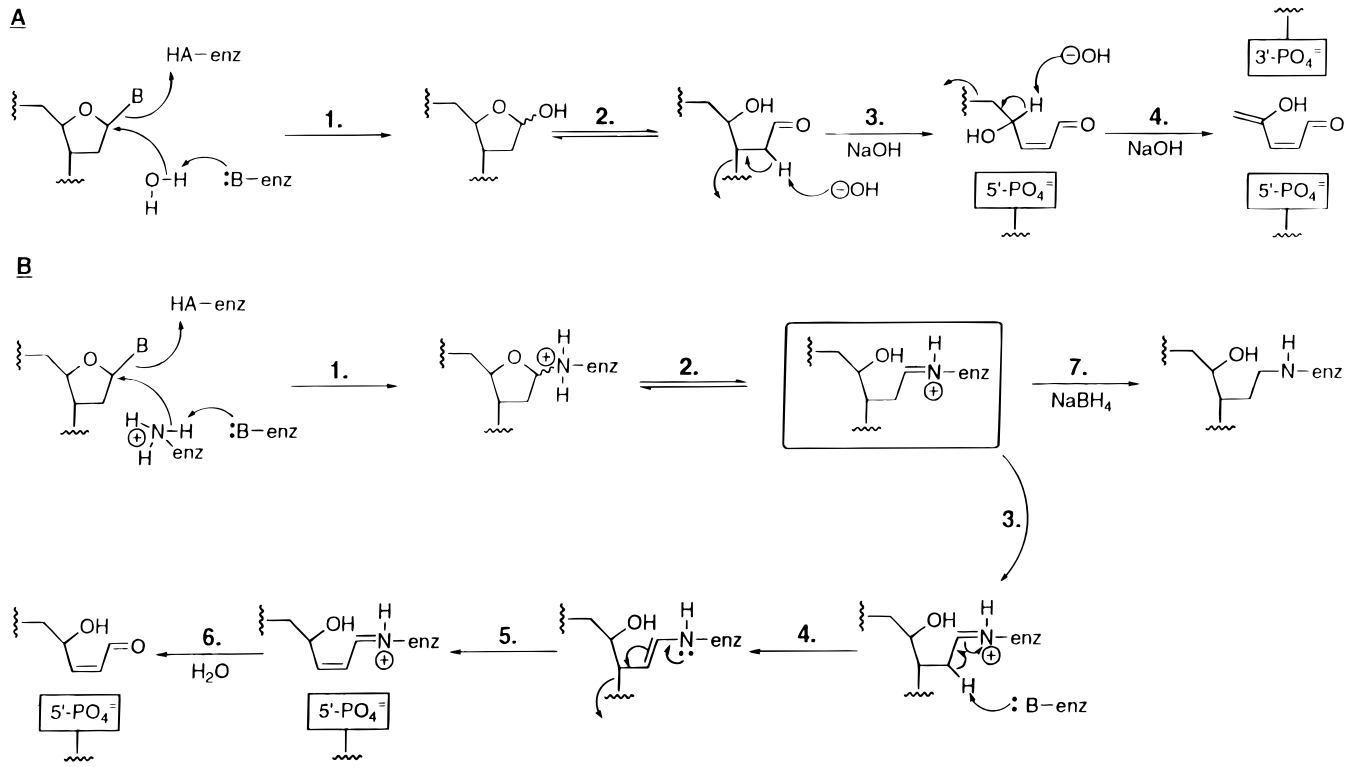

Figure 18. (A) Proposed catalytic mechanism for AP site generation by monofunctional glycosylase activity. Nucleophilic attack at the C1′ carbon by an activated water molecule is facilitated by protonation of the base [1]. The resulting AP site tautomerizes between the ring-closed and ring-open forms [2]. β- [3] and δ-eliminations [4] take place to effect DNA strand cleavage upon base treatment of the AP site. (B) Proposed catalytic mechanism for bifunctional glycosylase/AP-lyases. A side-chain amine serves as the attacking nucleophile at the C1′ carbon [1]. The resulting Schiff base undergoes open- and closed-form tautomerization [2], with the open ring form [3] facilitating proton abstraction from the 2′ position [4], effecting β-elimination [5]. Hydrolysis cleaves the covalent enzyme/DNA intermediate [6]. Sodium borohydride treatment reduces the Schiff base intermediate to generate a stable covalent enzyme/DNA complex [7].

strand scission are highly coupled, giving equal amounts of base removal and strand cleavage (Figure 18). Considerable mechanistic investigations have established that the enzyme-catalyzed strand scission reactions are a result of a β-elimination reaction rather than hydrolysis of the phosphate backbone.[228−232] The stereochemical course of the β-elimination reaction is syn, involving abstraction of the pro-*S* 2′-hydrogen and formation of a trans-α,β-unsaturated aldose product.[233,234] This is in contrast to the β-elimination reactions that occur under alkaline conditions which proceed by anti-β-elimination mechanism.[233] A subset of the enzymes in the bifunctional group catalyze a subsequent δ-elimination reaction, yielding a 3′-phosphate as the final product.[144,228,229] For example, in the case of the FPG protein, experiments using ^{18}O-labeling and identification of products by ^{31}P NMR and mass spectrometry have conclusively documented that the AP lyase activity of FPG entails successive β- and δ-elimination reactions (Figure 19).[229]

In this review, bifunctional enzymes are classified as those which catalyze both the glycosylase and lyase reactions with equal stoichiometry. The classification of BER enzymes has been previously described by Lloyd et al.[235,236] Uracil glycosylase and AlkA are well-characterized examples of the "simple" glycosylase category where any activity which results in strand scission is much slower than the glycosylase activity. T4-endo V, FPG protein, and endo III are examples of the group which catalyze the glycosylase and lyase reactions with equal rates. The classification of some BER enzymes remains controversial. For example, literature reports of *E. coli* MutY have characterized it as both a monofunctional glycosylase and bifunctional glycosylase/lyase.

Although the reactions catalyzed by the two classes of BER enzymes appear to be divergent, the only evident distinction is the identity of the nucleophile. This has been previously described by Lloyd as a "Unified Catalytic Mechanism" for BER enzymes.[235,236] In the case of monofunctional glycosylases, the nucleophile is postulated to be an activated water molecule that displaces the base and gives rise to an abasic site. In the case of the bifunctional enzymes, the nucleophile is a protein-derived amino group which displaces the base to produce a protonated Schiff base intermediate. This intermediate is poised for a β-elimination reaction and promotes subsequent strand scission (β-lyase) activity.[237] Two types of experiments are usually performed to distinguish the two classes. The first set of experiments involves investigating whether there is a difference in the amount of strand scission in the presence and absence of added base. In the case of monofunctional glycosylases which remove only the offending base, strand scission at the AP site requires the addition of base. If the enzyme is a coupled glycosylase/lyase, the amount of strand scission should be the same regardless of base treatment. The second type of experiment takes advantage of the fact that all of the characterized bifunctional enzymes proceed via nu-

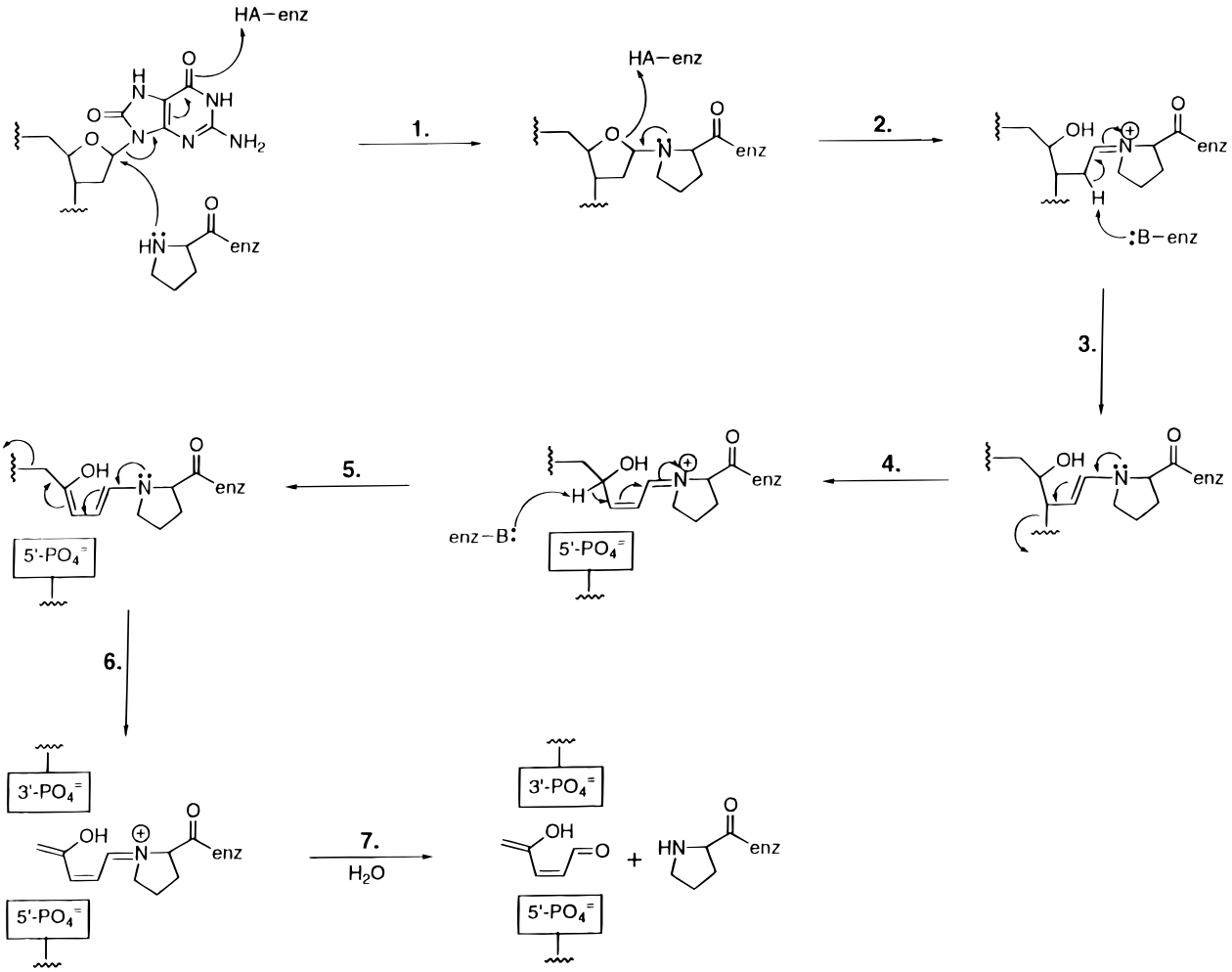

Figure 19. Enzymatic mechanism of the FPG protein illustrated on an OG (7,8-dihydro-8-oxoguanine) base in DNA. Nucleophilic attack at the C1′ carbon by the N-terminal proline residue is facilitated by protonation of the OG base [1]. The ring-opened iminium ion tautomer [2] is subject to 2′ proton abstraction [3]. The Schiff base is regenerated; conjugation effects β-elimination [4]. This is followed by abstraction of a 4′ proton [5], and the Schiff base is formed once again, resulting in a δ-elimination [6]. Hydrolysis cleaves the covalent enzyme/DNA intermediate [7].

cleophilic attack by an amino group to produce a Schiff base intermediate (Figure 18). This Schiff base intermediate can be trapped by reduction with sodium borohydride to yield a covalently linked enzyme−DNA complex which can be detected by denaturing PAGE as a slow-migrating species relative to free DNA. Another method used to support the intermediacy of an enzyme−DNA Schiff base is the inhibition of the enzymatic activity in the presence of cyanide.[238] This results from reversible addition of cyanide transiently intercepting the imine intermediate and preventing catalysis.

B. Identification of Active-Site Amino Groups in T4-Endo V and FPG

In T4-endo V, the intermediacy of a Schiff base has been well-documented and the identity of the nucleophilic amino group was determined to be the N-terminus. Reductive methylation of T4-endo V by treatment with formaldehyde and sodium borohydride inactivates the enzyme's glycosylase activity, suggesting the importance of a nucleophilic amino group.[239] Using protease digestion, the site of methylation was determined to be the α-amino group of T4-endo V.[239] Site-directed mutagenesis further characterized the importance of the amino-terminus in initiating the glycosylase/lyase activity.[240] Alterations of the N-terminal amino acid to similar amino acids had no detrimental effect on the enzymatic activity and ability to confer UV resistance on cells when expressed in *E. coli*. However, alterations in T4-endo V which extended the N-terminus by one amino acid or altered the nature of the amine by changing it to a proline inactivated the enzyme and drastically reduced the UV survival of cells expressing mutated forms of the enzyme. The intermediacy of a Schiff base was also indirectly suggested by the ability of cyanide to inhibit the activity of T4-endo V. Direct evidence for an imino intermediate (Schiff base) in the T4-endo V reaction and identification of the critical amino group was provided by trapping the intermediate in the presence of sodium borohydride.[238] The covalent enzyme−DNA intermediate was degraded using cyanogen bromide and the N-terminal peptide was isolated. The only amino group in this particular peptide was the α-amino group, thus implicating it as the locus for covalent attachment.

The identification of the nucleophilic amino group in the FPG protein was determined in a manner similar to that reported for T4-endo V.[241,242] Sodium cyanoborohydride was used to convert the transient imino intermediate formed during the enzymatic reaction into a covalent FPG−DNA complex. Subsequent treatment with cyanogen bromide resulted in identification of an N-terminal peptide fragment that contained the attached DNA.[241] After proteolytic digestion with trypsin and Pronase E, 10 and 5 amino acid N-terminal fragments were identified as containing the covalent modification using electrospray mass spectrometry techniques.[242] The inability to sequence the covalently modified fragments pointed to DNA modification of the N-terminal proline residue. In addition, modification of the N-terminal proline by replacement with glycine abolishes glycosylase activity.[241] The discovery that the N-terminal proline acts as a nucleophile with formation of a Schiff base intermediate is unique. The ensuing mechanism places a positive charge on the nitrogen facilitating the subsequent β- and δ-elimination reactions. The mechanism for the FPG protein is shown in Figure 19.

C. Identification of Imine Intermediates in the BER Superfamily

1. Endonuclease III and Identification of a Mammalian Counterpart

Early work on endo III by Kow and Wallace showed that the N-glycosylase and AP endonuclease activity reside on the same molecule and that the two activities act in concert.[243] These authors were the first to suggest that the intermediate steps of the reactions may involve formation of a Schiff base enzyme−DNA intermediate. The formation of such an intermediate is consistent with the considerable information available which indicates that the phosphodiester cleavage reaction results from an enzyme-catalyzed β-elimination reaction.[244] However, unlike FPG and T4-endo V, the identity of the catalytic amine has not been determined. On the basis of the crystallographic data of endo III, the active site is suggested to lie in a solvent-filled pocket within the interdomain groove near the HhH motif.[205,207] Lys 120 and Asp 138 are conspicuously located at the mouth of the pocket, suggesting their importance as catalytic residues. Conversion of Lys 120 to a glutamine (K120Q) residue resulted in a catalytically compromised enzyme, but did not affect the DNA binding properties. This evidence strongly implicates Lys 120 as the nucleophilic amine in endo III, but does not provide unambiguous support for this assertion.

In the presence of sodium borohydride, an irreversible cross-link can be formed between endo III and a thymine glycol containing DNA oligonucleotide.[113] This property was then extended to aid in the isolation and characterization of an endo III mammalian homologue from calf thymus.[113] By incubating a ^{32}P-labeled thymine glycol oligonucleotide with $NaCNBH_3$ and partially purified bovine enzyme, the covalently cross-linked band corresponding to the pure protein species was isolated by SDS−PAGE in quantities sufficient for primary sequencing. The sequencing information allowed for the cloning and expression of the bovine enzyme. The sequence of the bovine enzyme indicated that it is highly homologous to *E. coli* endo III.[113] With the use of the information on the bovine sequence, the isolation of cDNA and characterization of a human endo III homologue was also achieved.[114]

2. Yeast and Mammalian OG Glycosylase: Isolation, Sequencing, and Identification of a Critical Active-Site Lysine Residue

The ability to trap a Schiff base intermediate by reduction with sodium borohydride was also used by Verdine and co-workers to isolate a yeast OG glycosylase,[155] in a method analogous to that used to isolate the bovine endo III homologue. These authors also took advantage of the high affinity that many bifunctional enzymes exhibit toward synthetic reduced abasic site containing DNA (rAB) by incorporating the use of an rAB affinity column as an initial purification step. The resulting mixture of proteins was incubated with an OG:C duplex and sodium borohydride, and separated by SDS−PAGE. This allowed for the identification of two yeast OG glycosylases, yOGG1 and yOGG2. Sequencing revealed that yOGG1 had high homology to endo III and belongs to the BER superfamily.[167] By using the yeast sequence, the corresponding murine (mOGG1) and human (hOGG1) enzymes were located by a number of research groups.[5,171−173]

Lys 241 of yOGG1 and Lys 249 of the mammalian OGG1 proteins are located in the position corresponding to Lys 120 in endo III. Conversion of Lys 241 in yOGG1[5] and Lys 249 in mOGG1 and hOGG1[234] to glutamine abolishes the glycosylase activity, indicating an important catalytic role. Identification of the active-site amine in hOGG1 was accomplished by sodium borohydride trapping, protease digestion and amino acid sequencing (Figure 20).[245] These results unambiguously revealed that Lys 249 is the amine residue critical for the glycosylase activity and formation of the transient imine intermediate. In addition, these researchers showed that a mutant of hOGG1 containing a cysteine at position 249 (K249C) lacked detectable glycosylase and lyase activity; however, both activities can be restored by in situ generation of a γ-thialysine at 249 by reaction with 2-bromoethylamine as illustrated in Figure 21. Thus, these experiments provide further evidence that Lys 249 is the active-site amine in hOGG1. The strict conservation of a Lys residue at this position in the bifunctional enzymes of the BER superfamily strongly suggests that a similarly positioned amine functionality (e.g., Lys 249 in hOGG1, Lys 120 in endo III) acts as the nucleophile concomitantly forming the transient Schiff base intermediate, leading to β-elimination and strand scission.

D. Monofunctional vs Bifunctional Enzymes: The Controversy of MutY

Distinguishing between BER enzymes harboring monofunctional glycosylase activity and those having

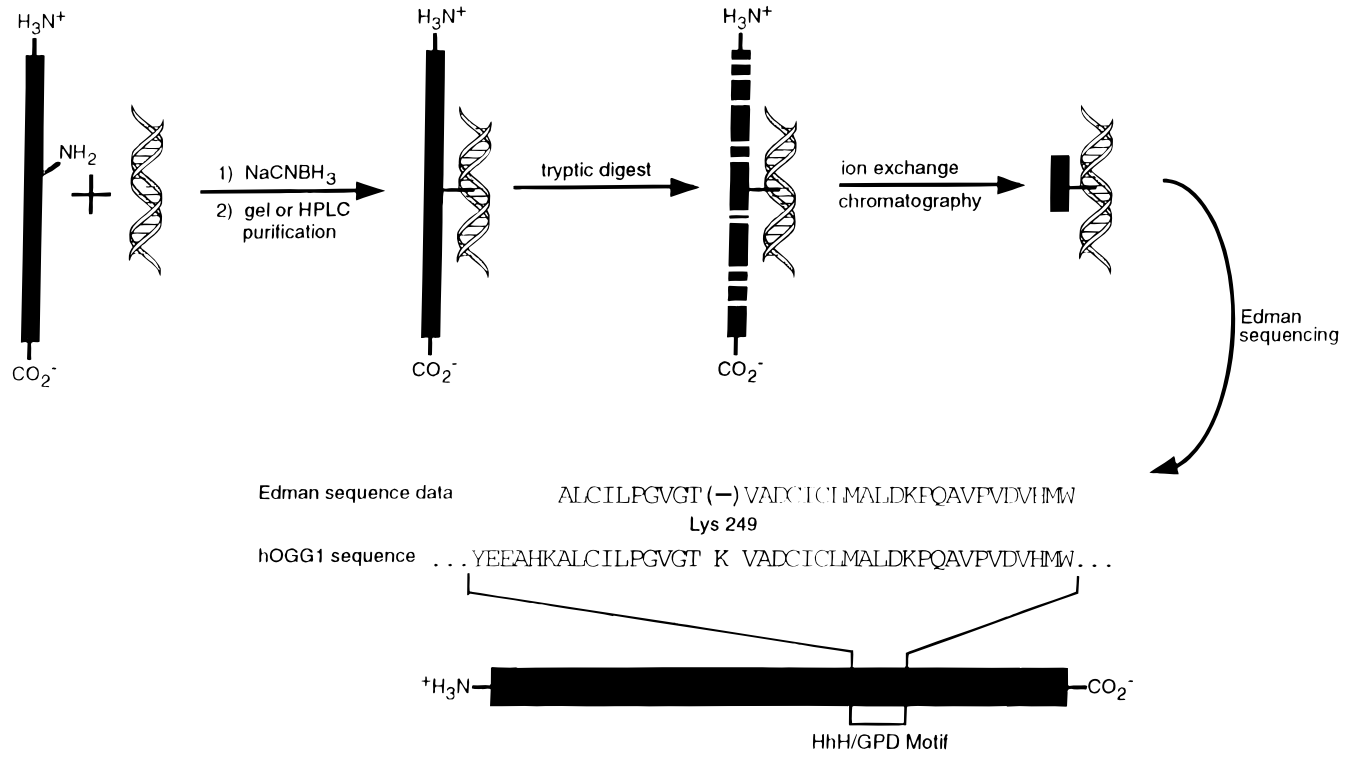

Figure 20. Identification of the hOGG1 active-site residue involved in the formation of the Schiff base intermediate during catalysis. The hOGG1 protein and an OG-containing duplex 49-mer were incubated in reducing conditions to yield a stable covalent enzyme/DNA complex. Tryptic digestion followed by anion-exchange chromatography resulted in the isolation of peptide fragments covalently bound to DNA. Forty cycles of Edman sequencing revealed no phenylthiohydantoin (PTH) amino acid derivatives at cycle 11 (Lys 249). (Adapted from ref 245.)

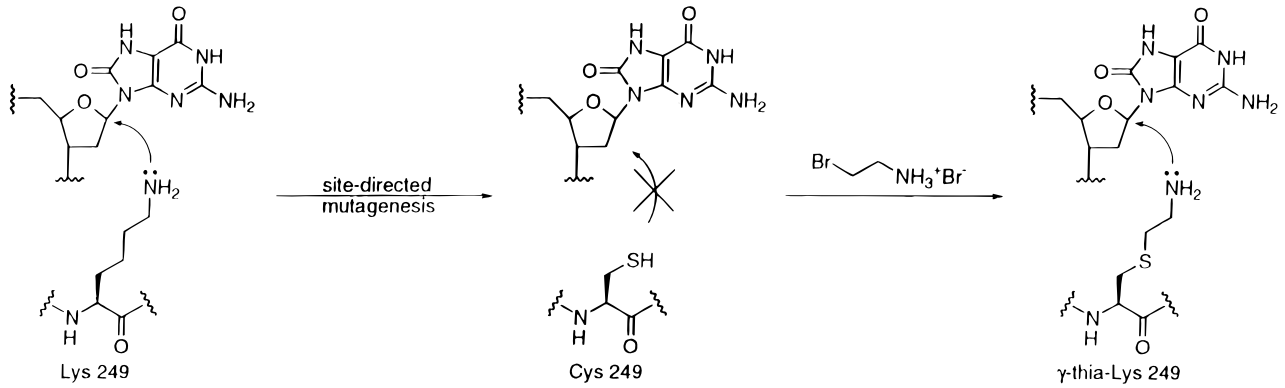

Figure 21. Active-site "rescue" of the hOGG1 glycosylase. Lys 249 of the wild-type enzyme is replaced by a Cys residue by site-directed mutagenesis. The activity of K249C mutants can be reestablished by alkylation of Cys 249 with 2-bromoethylamine hydrobromide. (Adapted from ref 245.)

bifunctional glycosylase/AP lyase activities is usually quite easy, requiring straightforward biochemical techniques. However, *E. coli* MutY has remained a difficult enzyme to characterize definitively. This fact is underscored by the conflicting data in the literature; some research groups have detected β-lyase activity associated with the glycosylase activity in their preparations of MutY, while others have not.

The past six years have seen controversy regarding this issue. Work in 1989 with MutY by Au and co-workers detected simple glycosylase activity only.[182] In 1992 however, the DNA strand nicking behavior observed in assays with MutY by Tsai-Wu et al.[186] was demonstrated to be intrinsic to the enzyme (i.e., not arising by virtue of endonuclease contamination in the final purified sample). When the two activities from samples of each purification step were quantified by densitometry, the specific activity of each increased proportionately as the purification progressed. However, the glycosylase activity for each step was shown to be slightly higher than that for strand scission.

In 1995, Lu and co-workers[187] reported strand-nicking activity in MutY toward G:A, C:A, OG:A, and G:N mispairs (where N = nebularine); this activity was not quantified with respect to AP site generation. In the same year, Sun and co-workers labeled MutY as a monofunctional glycosylase, in part by virtue of its inability to form a covalent enzyme-DNA complex in the presence of $NaBH_4$.[235] Substrate specificity studies with MutY in 1996 by Bulychev and co-workers[188] indicated no AP lyase activity in their

MutY preparations. In a recent report on the p26 catalytic domain of MutY by Manuel and Lloyd,[189] the authors report that AP lyase activity had a 1:1 rate correspondence to the glycosylase step in all nicking assays, since piperidine treatment after enzyme incubation did not increase the formation of cleaved product relative to reactions not exposed to base. This was true in the cases of both p26 and for the intact enzyme in the presence of OG:A substrate mispairs.

Preliminary results from our laboratory have shown that the generation of an AP site in OG:A mispair-containing duplexes in the presence of MutY occurs at a rate at least 10-fold higher than that of strand cleavage.[246] In quantitative activity comparisons performed between FPG, UDG, and MutY using the substrate mispairs OG:C, G:U, and OG:A, respectively, MutY behaved in a manner similar to the well-characterized monofunctional glycosylase UDG. The rate of AP site formation for both enzymes well outpaced that of strand cleavage. Furthermore, under the same conditions used for the UDG and MutY assays, FPG was shown to catalyze DNA strand cleavage at exactly the same rate as that of its glycosylase activity toward OG residues; the amount of cleaved DNA detected during the reaction was exactly the same regardless of base treatment.

In the same work, a more qualitative approach was also taken to address the issue of MutY's active-site chemistry. As mentioned above, borohydride reduction has become an emergent indicator for the formation of a Schiff base intermediate during the steps of bifunctional glycosylase/lyase enzyme catalysis. Reduction of the Schiff base forms a stable covalent enzyme−DNA complex which can be monitored apart from free DNA by its reduced migration in SPS-PAGE. The currently accepted catalytic mechanisms of monofunctional glycosylases do not involve Schiff base intermediates since the C1′-attacking nucleophile is proposed not to be a side chain amine functionality, but instead an activated water molecule.[236] As one might expect, a covalent complex was not formed between UDG and its G:U mispair-containing substrate. Covalent enzyme−DNA complexes were formed in the FPG/OG:C incubations, as would be predicted for a bifunctional DNA glycosylase eliciting strand scission concomitantly with AP site generation. Indeed, such covalent complexes were also detected in the reactions of MutY with OG:A mispairs. This would be unexpected on the basis of resemblance of the enzymatic activity to UDG in quantitative glycosylase assays.

A discrepancy exists between the results of the quantitative biochemical assays and the qualitative reduction with $NaBH_4$. MutY behaves quantitatively in a manner similar to UDG; both enzymes generate base-labile AP sites in duplex DNA at a rate which outpace any DNA backbone cleavage activity intrinsic to either enzyme preparation. It is possible that the conflict observed in the data of this work between the quantitative and qualitative assays may be at the root of the six-year-old literature discrepancy regarding the catalytic activity of MutY. Recent work has shown that for MutY, the rate for a single-turnover of an OG:A mispair substrate is faster than for a G:A mispair-containing substrate. However under conditions of turnover ([E] < [S]), at longer time periods, the percentage product conversion for G:A mispairs exceeds that for OG:A mispairs due to more efficient enzymatic turnover.[191] These kinetic data, in conjunction with dissociation constants[190] determined for substrates, products and substrate analogues suggest that MutY binding to G:A is weaker, and that enzymatic turnover is greater. Furthermore, MutY remains more tightly bound to the OG:A mispair product relative to that of the G:A mispair substrate.

Taking this information into account, it is possible to surmise that the active-site environment of MutY is more complicated than those of other BER enzymes which have been characterized without controversy surrounding the catalytic activity they harbor. One hypothesis would suggest the presence of an amine functionality (perhaps a lysine residue) in the active site of MutY close enough to the substrate mispaired 2′-deoxyadenosine to form a transient imine species which is subject to a subsequent nucleophilic attack at the C1′ carbon by an activated water molecule, as illustrated in Figure 22. Such a mechanism would explain the presence of a Schiff base intermediate in the catalytic process, as well as the fact that substrate strand scission for MutY is about 1 order of magnitude slower than glycosylase activity. Alternatively, the original nucleophilic attack at C1′ could be performed by an activated water molecule to generate an AP site. Subsequently, the C1′ of the AP site could be subject to a second nucleophilic attack by an amine functional group of the enzyme to yield the trappable Schiff base species. This may occur in vitro since MutY remains tightly bound to the product duplex.[190] In either of these models, the amine group responsible for the formation of the Schiff base intermediate would be close enough to perform a C1′ nucleophilic attack upon the substrate 2′-deoxyadenosine or AP site. However, the analogous basic residue which abstracts the 2′ proton in the current proposed mechanism for bifunctional enzymes may be in a more distal position in MutY, where water molecule activation is favored over the process of β-elimination.

Indeed a similar idea explaining the chemistry of MutY was recently suggested.[245] In such a scenario where a lysine ϵ-NH_2 group experiences occasional proximity to either the substrate for glycosylase action or the AP site, there could be an increase in strand scission events taking place separately from AP site generation. This would also explain the apparent imine species reduction in the presence of sodium borohydride. One possible way to test such a hypothesis would be to take advantage of the different rates of substrate turnover and substrate/product binding MutY has toward G:A and OG:A mispair-containing duplexes. If the formation of an imine species during MutY catalysis is the result of spurious encounters of an amine functionality with the substrate, then one might expect enzyme-mediated strand scission to be greater for OG:A than G:A mispairs, since the enzyme binds more tightly to OG:A mispairs. If the OG:A mispair substrate (or

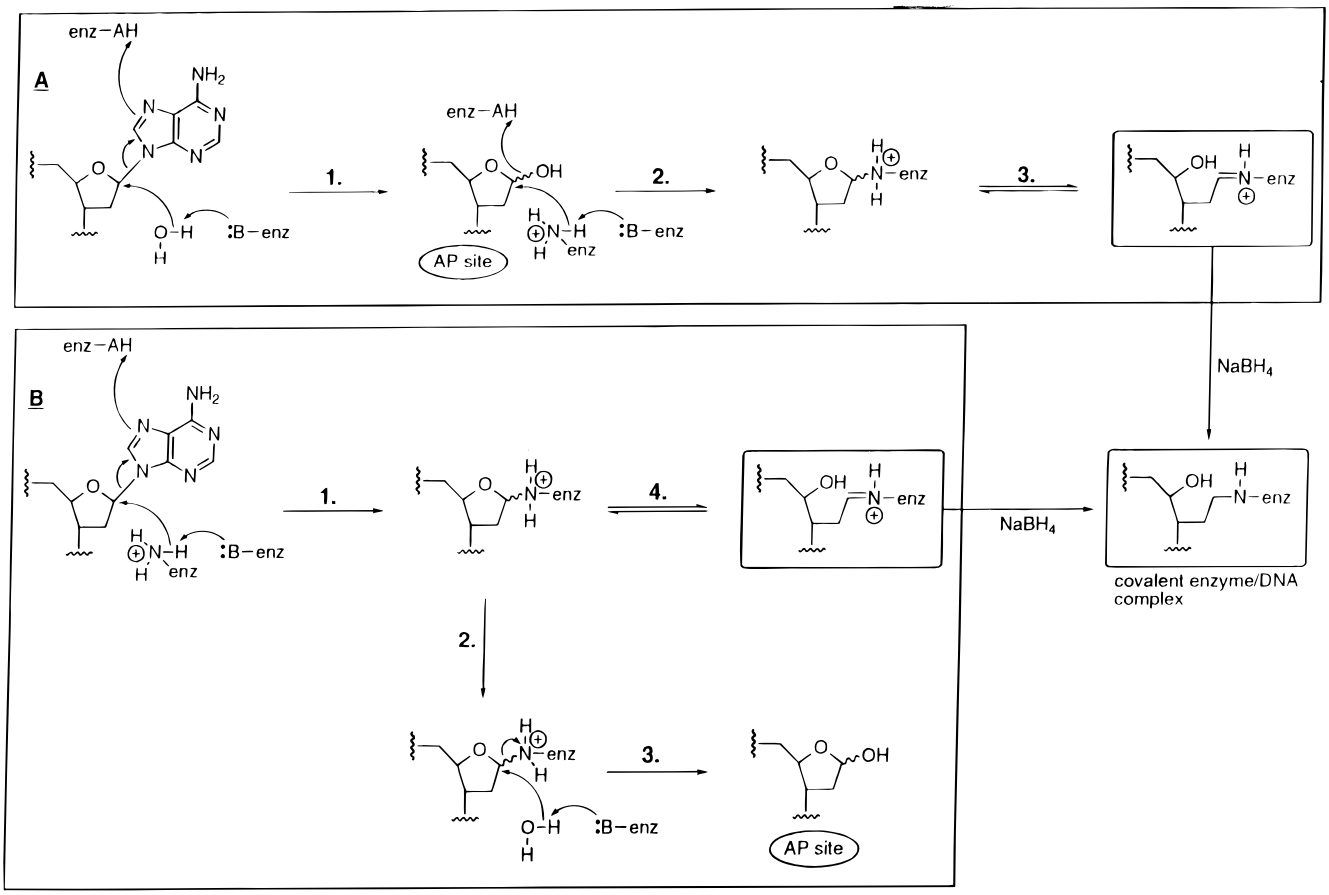

Figure 22. Possible active-site chemistry for MutY. (A) Initial nucleophilic attack at the C1′ carbon is by an activated water molecule, facilitated by protonation of the leaving adenine base [1]. The resulting AP site undergoes a second nucleophilic attack at C1′ by a side-chain amine to form the Schiff base species [2] which can be reduced to form a trapped enzyme/DNA complex [3]. (B) Initial nucleophilic attack at the C1′ carbon is by a side chain amine [1]. The closed ring Schiff base tautomer [2] undergoes another nucleophilic attack at C1′ by an activated water molecule to generate an AP site [3]. In addition, the open ring Schiff base tautomer can be reduced by sodium borohydride to form a covalent enzyme/DNA adduct.

product AP site) spends a greater amount of time in the active-site environment of MutY, the chances for an amine functionality to carry out a nucleophilic attack at C1′ would be greater and a larger amount of trappable intermediate will be observed in the presence of $NaBH_4$.

Homology modeling with the primary sequence of the MutY p26 catalytic domain and the X-ray crystal structure coordinates of endo III has located an interesting candidate lysine side chain (Lys 157) in the resulting energy-minimized MutY pseudostructure.[246] The terminal ϵ-NH_2 group lies ~3.5 Å from Ser 120, a suspected active-site residue for MutY. Indeed, Ser 120 of MutY lies in the same orientation as Lys 120 of endo III, the group proposed to be responsible for that enzyme's bifunctional glycosylase/β-lyase activity. Further credence is lent to the emergent picture of MutY's active-site pocket by virtue of sequence alignments with other related enzymes in the BER superfamily.[167,245] The endo III homologues, human OG glycosylase (hOGG1) and *M. luteus* UV endonuclease are all bifunctional glycosylase/lyases which harbor a lysine mapping to the *E. coli* endo III Lys 120 position in sequence alignments (see Figure 13). Similar alignments with monofunctional enzymes, including the MutY homologues from various species, reveal no conserved lysine residues at this position within the highly conserved HhH motif. The emergent pattern suggests that the conserved lysine residue may be important for BER superfamily enzymes eliciting AP lyase chemistry. Lysine side chains present in the monofunctional enzymes may be proximal to the active-site pocket to aid in DNA binding, but do not have the conserved lysine positioning found in bifunctional BER enzymes.

An alternative explanation for the variation in the behavior of MutY is the different protocols used for the isolation and purification of the enzyme and this has been previously suggested as the origin of the incongruous results.[189] There remains the possibility that some procedures preserve the strand scission activity of MutY while others contain steps which reduce or abolish it. This may be expected for a protein which contains a cofactor which may be lost upon purification; however, this seems unlikely with the known mechanistic properties of the AP lyase reaction. Other possible explanations for the differing results may lie with the intrinsic lability of abasic sites such that reaction conditions and handling of the duplex may result in different amounts of observed cleavage. This underscores the usefulness of comparison with well-behaved monofunctional and bifunctional enzymes in difficult cases such as MutY.

E. The Importance of an Aspartic or Glutamic Acid Residue in the Catalytic Activity of BER Glycosylases

A recurrent theme of BER enzymes of both monofunctional and bifunctional categories is the presence of a critical catalytic aspartic or glutamic acid residue. As mentioned previously with UDG and T4-endo V, the ability to produce catalytically impaired enzymes lacking an important carboxyl functionality has been exploited to provide structural snapshots of these enzymes bound to substrate DNA duplexes. Site-directed mutagenesis has implicated an important catalytic Asp or Glu in a variety of BER enzymes, including endo III, AlkA, T4-endo V, and UDG. A number of distinct roles have been postulated for this critical carboxylate residue.

The widest variety of roles have been suggested for Glu 23 in T4-endo V. An initial report indicated that replacement of Glu 23 with glutamine completely abolished the glycosylase activity, and it was suggested that the carboxyl group could be acting as a general acid by donating a proton to the base to ease its removal.[236] Additional site-directed mutagenesis experiments by Lloyd and co-workers converted Glu 23 to glutamine, aspartic acid, cysteine, and histidine, forming E23Q, E23D, E23C, and E23H mutants, respectively.[247a] The E23D mutant retained approximately 1% of the glycosylase activity while the E23H, E23C and E23Q mutants were completely inactive. Clearly these data indicate that the presence and *positioning* of a carboxyl moiety is important for the glycosylase activity of T4-endo V. The authors suggest that Glu 23 may participate as a general acid for protonation of the pyrimidine ring and/or as a general base for increasing the nucleophilicity of the α-amino group. This is consistent with the crystal structure of T4-endo V in which Glu 23 is located in close proximity to the α-amino group of the N-terminus.[247b] Interestingly, the E23D mutant still had considerable lyase activity (60% of the native enzyme) toward abasic sites in DNA. The other mutants, E23H, E23C, and E23Q, exhibited no detectable lyase activity. This was consistent with earlier work that showed that the E23D mutant retained lyase activity.[238] The importance of a carboxyl moiety for the lyase activity of T4-endo V prompted Lloyd and co-workers to suggest the Glu 23 may also play an important role in stabilizing the positive charge of the protonated imine intermediate or may be the general base that abstracts the 2′H in the β-elimination reaction. The participation of Glu 23 as a general base in catalyzing the AP-lyase activity is also consistent with the pH optimum of 5.5 for the lyase catalysis.[248] Glu 23 has also been shown to be necessary for the subsequent δ-elimination reactions catalyzed by T4-endo V. The replacement of Glu 23 with other amino acids does not adversely affect specific DNA binding, further supporting its general role in catalysis rather than damage recognition.[249]

In the cocrystal structure of the E23Q mutant T4-endo V bound to a pyrimidine dimer-containing duplex,[216] the α-amino group lies approximately 3.8 Å from the C1′ atom of the 5′-deoxyribose of the thymine dimer making it capable of participating as the nucleophile. The amide of Gln 23 makes three polar contacts with the base and sugar of the 5′-T of the T◇T dimer, suggesting that the corresponding carboxyl group (Glu 23) of the native enzyme is positioned favorably for participation in the glycosylase activity. Modeling based on the crystal structure suggests that Glu 23 could also participate in the lyase activity by abstracting the pro-*S* hydrogen at the C2′ position.

In UDG from human cells and HSV-1, crystallography and mutagenesis have shown a critical role for an aspartic acid residue in the enzyme's action.[186−188] Conversion of aspartic acid 145 to asparagine (D145N) or glutamic acid (D145E) in the human UDG results in significantly reduced enzymatic activity, but does not completely abolish it.[197] The D145N and D145E mutants retain 0.04 and 0.08% of the activity of the wild-type enzyme, respectively, while the wild-type affinity and specificity is maintained. Mutagenesis experiments also implicated His 268 in the human enzyme as important for catalysis.[197] In the two crystal structures of the human[197] and HSV-1 enzymes,[198] both the implicated His (268 and 210, respectively) and Asp (145 and 88, respectively) residues appear to be positioned such that they could participate in the enzymatic action. In the HSV-1 enzyme,[198] Savva et al. proposed that Asp 88 (Asp 145 in hUDG) may be the general base that deprotonates water for attack at C1′ to displace the uracil base, while Mol et al.[197] postulated that His 268 in hUDG (His 210 in HSV-1) may participate in an analogous fashion, or directly as the nucleophile. In the crystal structure of the hUDG-DNA complex,[199] His 268 is positioned such that it is hydrogen-bonded to the uracil base O2. In the active site, Asn 145 is rotated toward the bound uracil, suggesting that the corresponding Asp residue in the native enzyme may activate a water nucleophile for in-line attack at C1′. The His 268 residue likely assists in the release of the uracil base by transfer of a proton to the uracil O2, thus enhancing the ease of elimination/displacement of the uracil base as illustrated in Figure 23.

In the prototypes for the BER superfamily, endo III and AlkA, an important aspartic acid residue has also been identified. In the crystal structure of endo III,[205] Asp 138 and Lys 120 lie at the mouth of the suggested active-site pocket. Site-directed mutagenesis experiments on endo III,[205] in which aspartic acid 138 was converted to a glutamine (D138Q), resulted in a 100-fold reduction in k_{cat}, but only a 4-fold increase in K_m. Inspection of the endo III structure suggests that Asp 138 is well-positioned to deprotonate Lys 120. The corresponding position in AlkA harbors Asp 238 which protrudes into the center of the active-site cleft in the crystal structure.[204] Mutation of Asp 238 to asparagine (D238N) eliminates detectable enzymatic activity but does not affect DNA binding.[203,204] The D238N mutant has reduced affinity for a positively charged pyrrolidine transition state analogue but not a neutral abasic site analogue, suggesting that the negatively charged Asp residue participates in the displacement of the base.[204]

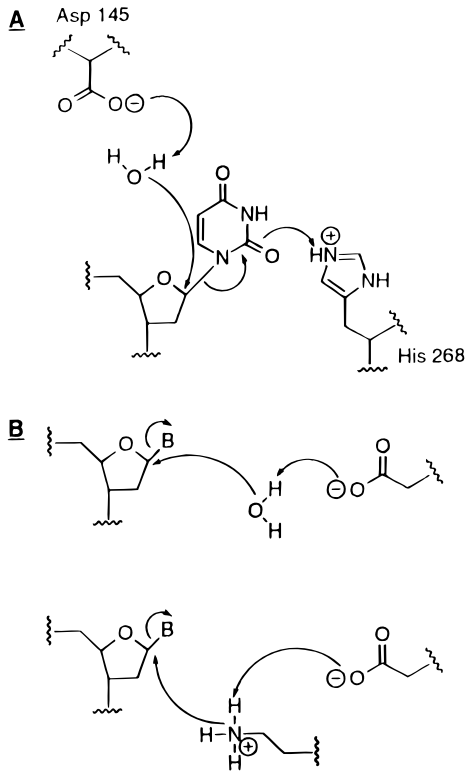

Figure 23. (A) Proposed mechanism for human uracil-DNA Glycosylase (hUDG). Asp 145 activates a water molecule nucleophile for C1′ attack, facilitated by uracil base protonation by His 268. (Adapted from refs 198 and 199.) (B) The general role of the highly conserved Asp residue in BER enzyme active sites is likely to be activation of the nucleophile in the cases of both monofunctional enzymes (water) and bifunctional glycosylase/lyases (amine).

Furthermore, D238N mutant-containing cells are unable to survive the detrimental effects of the methylating agent, methyl methanesulfonate (MMS).[203] Consistent with the spacious and nonspecific active-site cleft of AlkA, replacement of Asp 238 with a glutamic acid does not completely abolish the enzymatic activity and allows for some resistance to MMS in vivo.[203]

The presence of an Asp residue in analogous positions in both endo III and AlkA suggests that the role of this Asp residue is highly conserved in the BER superfamily. Indeed, the location and presence of an Asp residue is strictly conserved in members of the BER superfamily of both the monofunctional and bifunctional class. Likely, the Asp plays a similar role in both classes and that the mechanistic differences between the two are subtle. On the basis of the unified catalytic mechanism proposed by Lloyd,[236] and the structural work on AlkA,[204] a proposed role of the Asp (or Glu) as a general base for activating the nucleophile (water or amino group) has been suggested which is shown in Figure 23b. This suggests that the basic chemistry of the BER superfamily enzymes is similar and that slight modifications in the enzymes may be present depending on the ease of removal and recognition of the damaged/mismatched base. The fact that BER enzymes with different three-dimensional structures also contain a critically conserved Glu or Asp and have similar mechanistic features indicates that the base-excision *chemistry* in the entire class is highly conserved.

F. The Use of Synthetic DNA Analogues

In the structural and biochemical studies discussed thus far, an important theme has been the use of site-directed mutagenesis to alter BER enzymes. This has proven exceedingly profitable in the case of crystallographic characterization of inactive BER enzymes bound to their cognate DNA damage. The advances in the field of automated DNA synthesis using phosphoramidite or phosphonate chemistry has made synthetic oligonucleotides readily available as an additional tool for studying BER enzymes. Additionally, many modified phosphoramidite or phosphonate monomers are now either commercially available or readily synthesized using established chemistry. A relatively new approach to providing insight into the mechanism and recognition aspects of BER enzymes involves using synthetic analogues to the BER enzyme substrates. A powerful example of this approach has been in the preparation of substrate and transition-state analogues that bind tightly to BER enzymes, yet are resistant to their enzymatic action. These molecules provide a stable and specific BER enzyme−DNA complex amenable to structural and biochemical characterization.

Verdine and co-workers have used the transition-state destabilization approach to prepare a general class of BER enzyme inhibitors by incorporation of an electron-withdrawing fluoro substituent at C2′ of the deoxyribose ring.[250] These inhibitors are based on a proposed transition state (Figure 24) in which substantial positive charge is expected to accumulate in the 2′-deoxyribose ring between the O-1′ and C-1′ positions. The electron-withdrawing fluorine substituent at C2′ was incorporated to destabilize the transition state by increasing the positive charge density at C1′. Destabilization of the transition state translates into a decreased rate for the glycosylase reaction. A 2′-deoxy-2′-fluoroethenoadenine-containing oligonucleotide (Figure 24) is completely resistant to the glycosylase activity of human alkyl *N*-purine DNA glycosylase (ANPG or MPG). However, the dissociation constant of ANPG with the modified duplex is similar to that with the native substrate. The generalizable nature of this approach was further illustrated by the use of oligonucleotide duplexes containing 2′-deoxy-2′-fluorouridine opposite guanine as substrate analogues for the mismatch-specific thymine glycosylase (TDG).[251] In this case, both α and β, and difluoro substitution at the C2′ position were investigated and duplexes containing all three types of 2′-fluoro-substituted uridine were found to be resistant to the glycosylase activity of TDG while retaining high binding affinity. DNase I footprinting and DMS interference footprinting further demonstrated the specific nature of the interaction of TDG with the 2′-fluoro derivatives.

Iwai et al. used a similar approach by incorporation of a fluoro substituent at the 2′-(α)-position of the 5′-sugar of a *cis-syn*-thymine dimer (Figure 24).[252] However, in this case, T4-endo V was able to catalyze

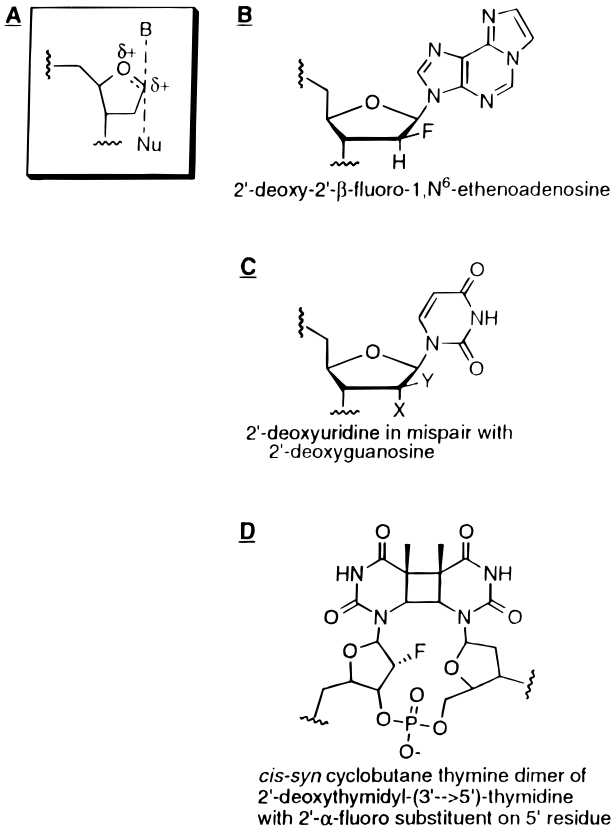

Figure 24. Fluorinated substrate analogues for BER enzymes. (A) Proposed transition state for nucleophilic attack at the C1′ in base excision. (B) Substrate analogue inhibitor for human alkyl *N*-purine DNA glycosylase (ANPG). (C) Substrate analogue U:G mispair employed in assays of human thymine DNA glycosylase (TDG). The 2′ position substitutions are as follows: X = F, Y = H, 2′-deoxy-2′-α-fluorouridine; X = H, Y = F, 2′-deoxy-2′-*β*-fluorouridine; X = F, Y = F, 2′-deoxy-2′,2′-difluorouridine. (D) Substrate analogue inhibitor for T4-endonuclease V.

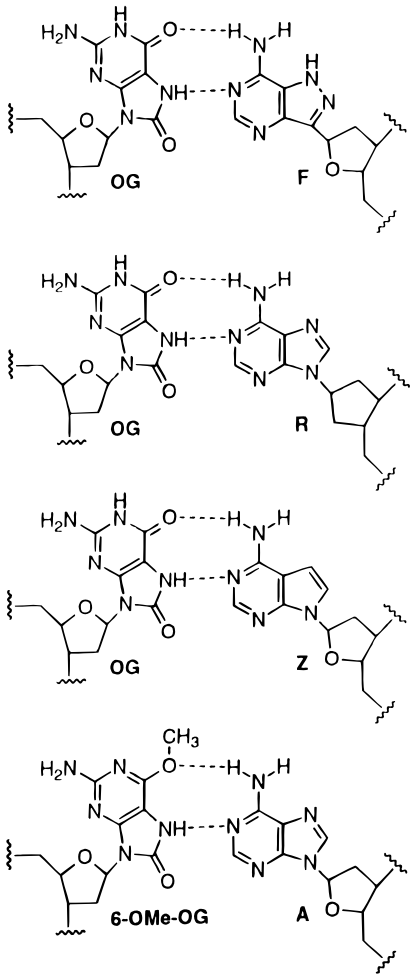

Figure 25. Substrate analogue mispairs for MutY. F,2′-deoxyformycin-A; Z-7-deazaadenosine; R,2′-deoxyaristeromycin; OG, 7,8-dihydro-8-oxo-2′-deoxyguanosine; 6-OMe-OG, 6-methoxy-7,8-dihydro-8-oxo-2′-deoxyguanosine.

the removal of the 5′-thymine of the dimer, albeit at a significantly reduced rate. The modified substrate was completely resistant to the subsequent enzyme-catalyzed *β*-elimination reaction. In an analogous fashion, fluoro substitution at the 2′-(α)-position of 2′-deoxyuracil has also been employed to make a modified substrate for hUDG.[253] In this case, the binding of hUDG to the fluoro-substituted sugar was reduced significantly (100−200-fold) and the ability to remove the uracil base was completely abolished. These authors suggest that the 3′-endo conformation of 2′-fluoro-substituted sugars may be influencing the binding of hUDG to these modified substrates. These results are consistent with the crystal structure of hUDG-DNA complex in which the movement of His 268 into position for interaction with the uracil O2 would be inhibited by the presence of the 3′-endo sugar configuration.[199] These two studies suggest that the ability to retain native binding affinity and inhibit the glycosylase activity by fluoro substitution at the 2′-position of the target nucleotide deoxyribose may depend on the nature of the base to be removed and the specific BER enzyme(s) involved.

Several types of substrate analogues have been prepared for *E. coli* MutY, possibly due to its recognition of a base pair (OG:A, G:A) and glycosylase action on the normal base (adenine) within this base pair, providing a plethora of potential targets. Many substrate analogues result in reduced glycosylase activity (see section II.H). Notably, however, many of these substrate analogues are resistant to the glycosylase action of MutY and therefore may serve as useful replacements for the substrates in structural studies. For example, Bulychev et al.[188] have found that MutY can bind with nanomolar affinity to a MutY substrate containing a methoxy substituent at the 6-position of OG (Figure 25). The mechanism of the resistance of this analogue to the adenine glycosylase activity of MutY is unclear; however, these results point to the importance of the entire base pair in the activity of MutY. They also reported that a carbocyclic analogue of 2′-deoxyadenosine (2′-deoxyaristeromycin) was resistant to the glycosylase activity of MutY, but maintained the nanomolar binding affinity.

The approach for the preparation of suitable substrate analogues taken by Porello et al.[190] used natural products that resemble the ribonucleoside adenosine as the inspiration for DNA duplex-based substrate analogues. In particular, the ribonucleosides tubercidin (7-deazaadenosine) and formycin-A are structurally similar to adenosine and have been

shown to mimic adenosine in enzymatic reactions in vitro. The presence of a ribosyl carbon linkage in formycin A and the lack of N-7 in tubercidin suggested that these analogues would be resistant to enzymatic removal of the base. Conversion of these ribonucleosides to the corresponding 2′-deoxyphosphoramidite monomers allows for incorporation into a DNA duplex (opposite G or OG) using standard phosphoramidite chemistry. The resultant duplexes were shown to be resistant to the glycosylase activity of MutY and retain high affinity and specific binding by MutY. In particular, 2′-deoxyformycin (F)-containing duplexes bound with subnanomolar affinity, suggesting that they efficiently mimic the substrate-containing duplexes. A duplex containing a central OG:F base-pair was used in methidium-propyl-EDTA-Fe(II) hydroxyl radical footprinting experiments and indicated a region of protection surrounding the OG:F base pair.[190] Additional footprinting experiments using dimethyl sulfate (DMS) showed that the guanine bases located on the OG-containing strand are protected from reaction with DMS in the presence of MutY while guanine bases located on the F-containing strand are hyperreactive to DMS in the presence of MutY.[191] Thus, MutY binding to the OG:F base pair may be accompanied by DNA distortion (such as flipping out of the F or OG nucleotide), resulting in increased accessibility of DMS to guanines proximal to the mismatch. Thus, these experiments provide evidence that this approach can provide insight into the properties of a BER enzyme−DNA complex. Similar experiments are not possible using the substrate due to rapid enzymatic processing.

Extremely tight-binding inhibitors have been prepared by Verdine and co-workers using transition-state mimicry.[254,255] On the basis of the proposed transition state for the glycosylase reaction, a pyrrolidine-based DNA inhibitor was prepared which mimics the positive charge that accumulates at O1′ in the proposed transition state as illustrated in Figure 26.[254] Previously, Schramm et al.[256−258] had identified 1,4-dideoxy-1,4-iminoribitols (e.g., pyrrolidine) as potent transition-state inhibitors of nucleoside hydrolase which suggested that an analogous approach could be applied to BER glycosylase enzymes. A pyrrolidine-based DNA duplex inhibitor binds tightly to AlkA, with an impressively high affinity (picomolar K_d). A duplex containing a tetrahydrofuran derivative which lacks the positive charge exhibited minimal binding affinity for AlkA. This approach was further elaborated by appending an adenine base to the pyrrolidine-base DNA inhibitor via a CH_2 linker (referred to as phA, Figure 26).[255] A duplex containing a central phA:OG base pair exhibited a *sub-picomolar* dissociation constant for MutY. The enhancement in binding affinity by adding the base moiety to the pyrrolidine was at least 60-fold. This is a general approach to obtain extremely tight-binding inhibitors for BER enzymes by appending different base units to the pyrrolidine analogue. This approach will be useful in forming stable enzyme−DNA complexes for structural studies and promises to provide insight into the properties of these enzymes' transition states and mechanisms of action.

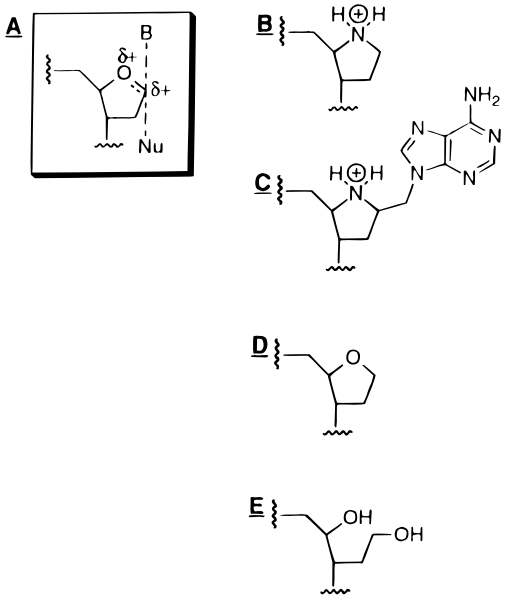

Figure 26. Transition-state and mechanism-based inhibitors: (A) proposed transition state for nucleophilic attack at the C1′ in base excision; (B) pyrrolidine inhibitor (pyrr); (C) pyrrolidine plus appended base inhibitor (phA); (D) tetrahydrofuran inhibitor (THF); and (E) reduced abasic site inhibitor (rAB).

The pyrrolidine-based inhibitor appears not to bind with as high affinity to all BER enzymes. For example, T4-endo V binds 1000-fold less tightly to a pyrrolidine duplex than AlkA.[259] T4-endo V binds with slightly higher affinity to a reduced abasic site (rAB) inhibitor than the pyrrolidine inhibitor or a neutral tetrahydrofuran (THF) moiety (Figure 26). Endo III, MutY, and FPG also exhibit high affinity for a reduced abasic site.[189,260,261] This observation for FPG was exploited in the isolation of functional yeast homologue, yOGG1, by utilizing an rAB-affinity matrix to enrich yeast extracts in the desired protein (section IV.C.2).[167] Significant affinity of many glycosylases for DNA containing tetrahydrofuran as an abasic site mimic has also been observed.[162,188,189,259] The preference in binding of these analogues to different BER enzymes may provide insight into the subtle differences in their transition states and may be expected in the future to provide insight into the mechanistic details of the monofunctional glycosylase and bifunctional glycosylase/lyase reactions.

V. AP Endonucleases Associated with BER

A. Classification of AP Endonucleases

Abasic sites in DNA arise from the action of BER glycosylases as described previously, or via spontaneous or mutagen-induced hydrolysis of the N-glycosylic bond.[1,252,253] AP sites are potentially lethal due to halting of DNA polymerase or mutagenic due to the lack of information for DNA polymerase to insert the proper opposite base.[262] A large number of enzymes have been classified as AP endonucleases, including enzymes that are both 5′- and 3′-AP endonucleases.[263] The class I-type AP endonucleases are 3′-AP endo-

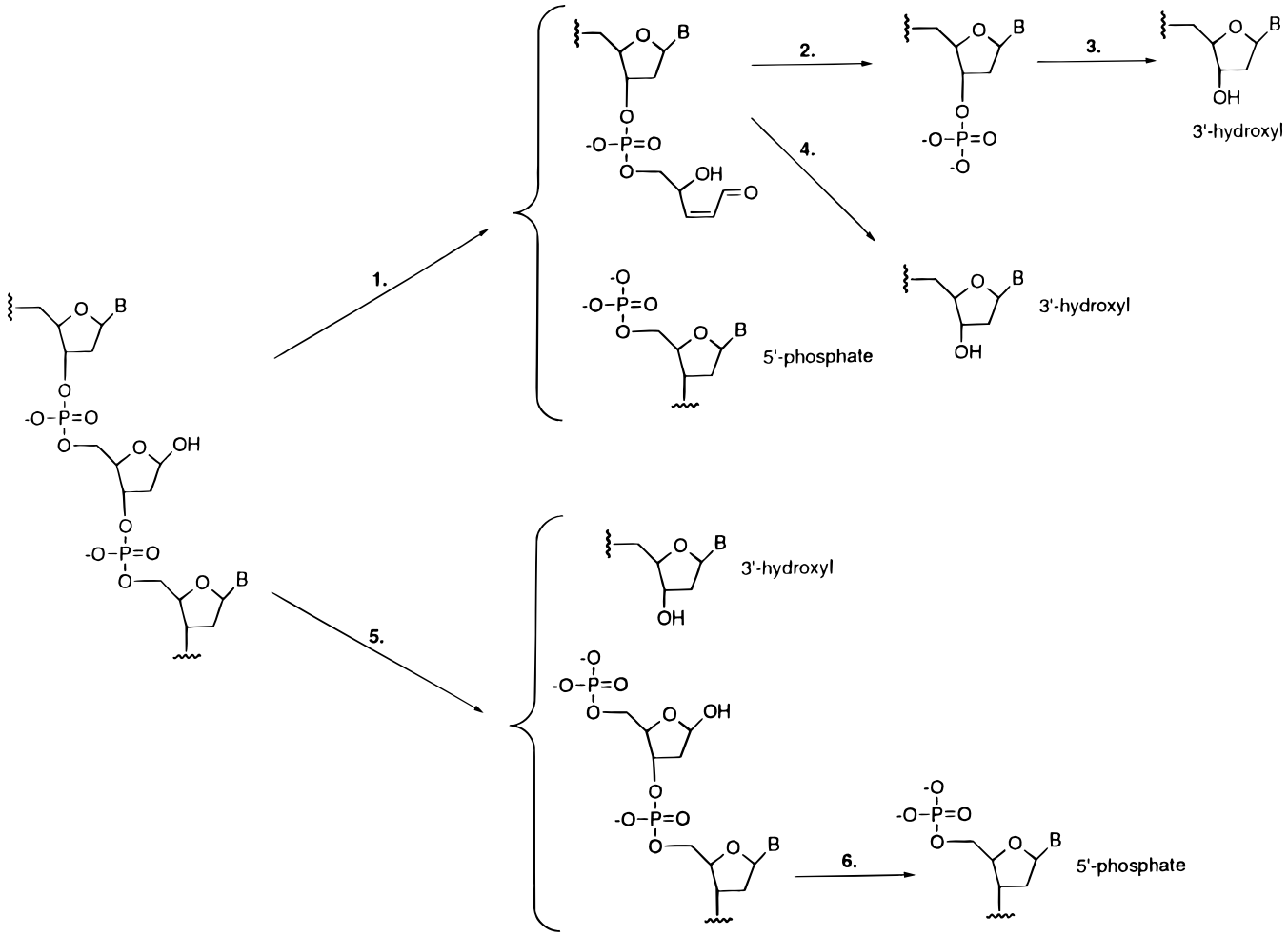

Figure 27. Processing of abasic sites in DNA by AP endonucleases. For simplification, only a single strand of DNA is shown. A class I AP endonuclease or AP lyase (usually associated with a glycosylase) generates an α,β-unsaturated aldehyde at the 3′-end via β-elimination (1), and subsequent δ-elimination (2) produces a 3′-phosphate end. The action of a 3′-phosphatase (3) or 3′-phosphodiesterase (4) produces a 3′-hydroxyl-end which is needed for the replacement chemistry of DNA polymerase. Alternatively, the AP site is processed by a class II AP endonuclease (5′-AP endonuclease) which hydrolyzes the 5′-phosphodiester bond (5). Subsequent action of a DNA deoxyribophosphodiesterase (6) generates a one-nucleotide gap with proper 3′-hydroxyl and 5′-phosphate ends for DNA polymerase. Consult text for more details. (Adapted from ref 263.)

nucleases which catalyze β-lyase (β-elimination) reactions resulting in a α,β-unsaturated aldehyde end and a 5′-phosphate end (Figure 27). This activity is usually associated with glycosylase activity and has been discussed extensively in the previous sections. A subsequent δ-elimination reaction can yield a 3′-phosphate end. However, a 3′-hydroxyl (3′-OH) end is required for DNA polymerase to fill in the missing nucleotide, and therefore the action of additional enzymes are often required prior to recruiting DNA polymerase. Alternatively, AP sites left after the action of a BER glycosylase can be the substrates for class II AP endonucleases (5′-AP endonucleases) which hydrolyze the phosphodiester backbone 5′ to the abasic site. Subsequent action of DNA deoxyribophosphodiesterase (dRPase) generates a one-nucleotide gap with the requisite 3′-OH and 5′-phosphate ends for DNA polymerase.[144,264] The processing of AP sites is illustrated in Figure 27. A large number of class II AP endonucleases are known. Those that have been extensively characterized include the *E. coli* enzymes, exonuclease III and endonuclease IV, the human AP endonuclease HAP1 (also called Ape, APEX and REF1), and the *S. cerevisiae* Apn1 protein.[263]

B. Substrate Specificity

Exonuclease III (exo III) is the major AP endonuclease in *E. coli* and is highly homologous to the human AP endonuclease HAP1.[265] Exo III was first identified as a 3′-5′-exonuclease which degraded double-stranded DNA. Indeed, this property of exo III has been exploited to identify sites of stable DNA modifications such as platination sites of cis-platin[267] since the 3′-5′-exonuclease activity halts at the site of modification. Exo III is able to catalyze phosphate ester hydrolysis reactions on either phosphomonoesters or phosphodiesters, as is illustrated by its activity as a 3′-phosphatase, 3′-phosphodiesterase, and 5′-AP endonuclease (Figure 28).[264] Interestingly, exo III also exhibits RNase H activity.[264] Eukaryotic homologues to exo III, such as HAP1, have similar enzymatic properties; however, most reports indicate that the eukaryotic homologues lack the 3′-5′ exonuclease activity.[264,268−270] Endonuclease IV (endo IV) was originally found in *E. coli* mutants lacking exo

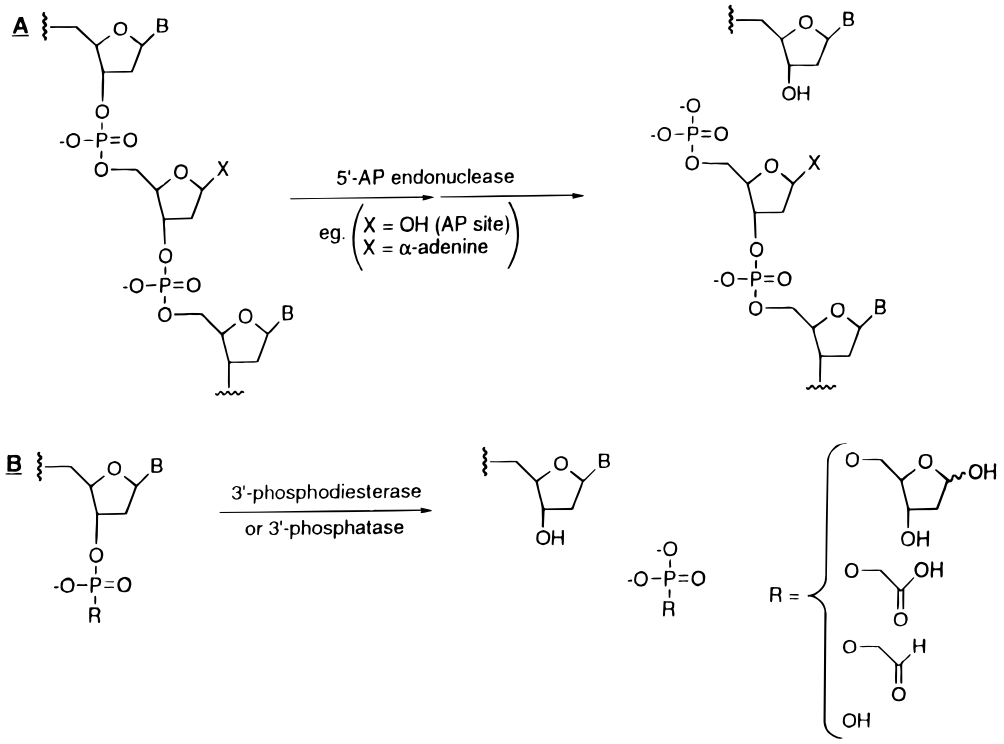

Figure 28. Representative hydrolysis reactions catalyzed by exonuclease III on phosphodiesters and phosphomonoesters. Exo III catalyzes the 5′-phosphodiester cleavage on AP sites and certain modified base adducts (A). Exo III also removes modified 3′-ends using 3′-phosphodiesterase or 3′-phosphatase activity (B). (Adapted from ref 13 and 263.)

III and its resting concentration is approximately 5% of that of AP endonucleases found in *E. coli*.[263] However, the endo IV protein is induced in response to oxidative stress as part of the SOXRS regulon.[269b] Endo IV exhibits 5′-AP endonuclease activities similar to that of exo III but lacks both the RNase H and 3′-5′-exonuclease activities. Apn1 is the major AP endonuclease in *S. cerevisiase* and is highly homolgous to *E. coli* endo IV.[263] The enzymatic properties of Apn1 are also similar to those of endo IV.[263]

Cells lacking exo III and endo IV are sensitive to oxidizing agents, suggesting that these enzymes may be important for the repair of abasic sites produced from oxidative damage to the sugar−phosphate backbone.[263,270,271] Consistent with this proposal, the class II AP endonucleases have been shown to be active toward a variety of abasic sites (Figure 29). In particular, oxidation of the C4′ and C1′ positions by Fe^{III}-bleomycin and Cu^{II}-phenanthroline in the presence of hydrogen peroxide results in the production of substrates that are recognized by exo III and endo IV. However, exo III is considerably less active (at least 400-fold) on 4′-oxidized abasic sites compared to unoxidized abasic sites.[272] Synthetic abasic site mimics such as tetrahydrofuran and propanediol are also relatively efficient substrates for exo III and endo IV.[273] The lack of a base or base-like moiety does not appear to be an important recognition feature of these enzymes, since 5′-cleavage is observed on substrates containing urea, hydroxylamine and α-adenine (Figure 29).[13] Another surprising observation was the ability of these enzymes to catalyze 5′-cleavage of substrates containing bases modified by *p*-benzoquinone (Figure 29).[13,274] The hydrolytic reaction 5′ to these bulky lesions, in addition to abasic sites, suggests that the recognition feature is not the absence of a base moiety but rather a structural distortion that is created either by the bulky adduct or the presence of an abasic site.[13]

A number of studies have examined the influence the abasic site environment has on the 5′-phosphodiester cleavage activity of exo III, HAP-1, and endo IV. These studies provide additional insights into the factors that may be important for substrate recognition. For example, phosphorothioate substitution 5′ to an abasic site mimic (tetrahydrofuran) inhibits both exo III and endo IV cleavage; however, only endo IV retains the ability to bind to an Rp-phosphorothioate-containing isomer.[273] Phosphodiester cleavage of both exo III and endo IV is insensitive to the nature of the base opposite the tetrahydrofuranyl abasic site. The activity of endo IV appears to be more sensitive to the presence of mismatched base pairs adjacent to the abasic site, particularly those located on the 5′-side of the abasic site.[273] Generally, this class of AP endonucleases is specific for duplex DNA; however, incision by exo III has been observed at an abasic site within single stranded DNA that contained a second abasic site on its 5′-side.[275] The importance of duplex around an abasic site was illustrated[276] by investigation of the 3′-phosphodiesterase activity of HAP-1 on duplexes containing 3′-phosphoglycolate ends at internal gaps or at the ends of duplexes. These results indicate that HAP-1 can remove phosphoglycolate ends from the 3′-end of an internal one base gap. In contrast, HAP-1 is sluggish toward removal of phosphoglycolate at a double-stranded end with a blunt, recessed or overhanging 3′-end.[276] The presence of two nearby abasic sites has been found to influence the activity of HAP-1 and exo

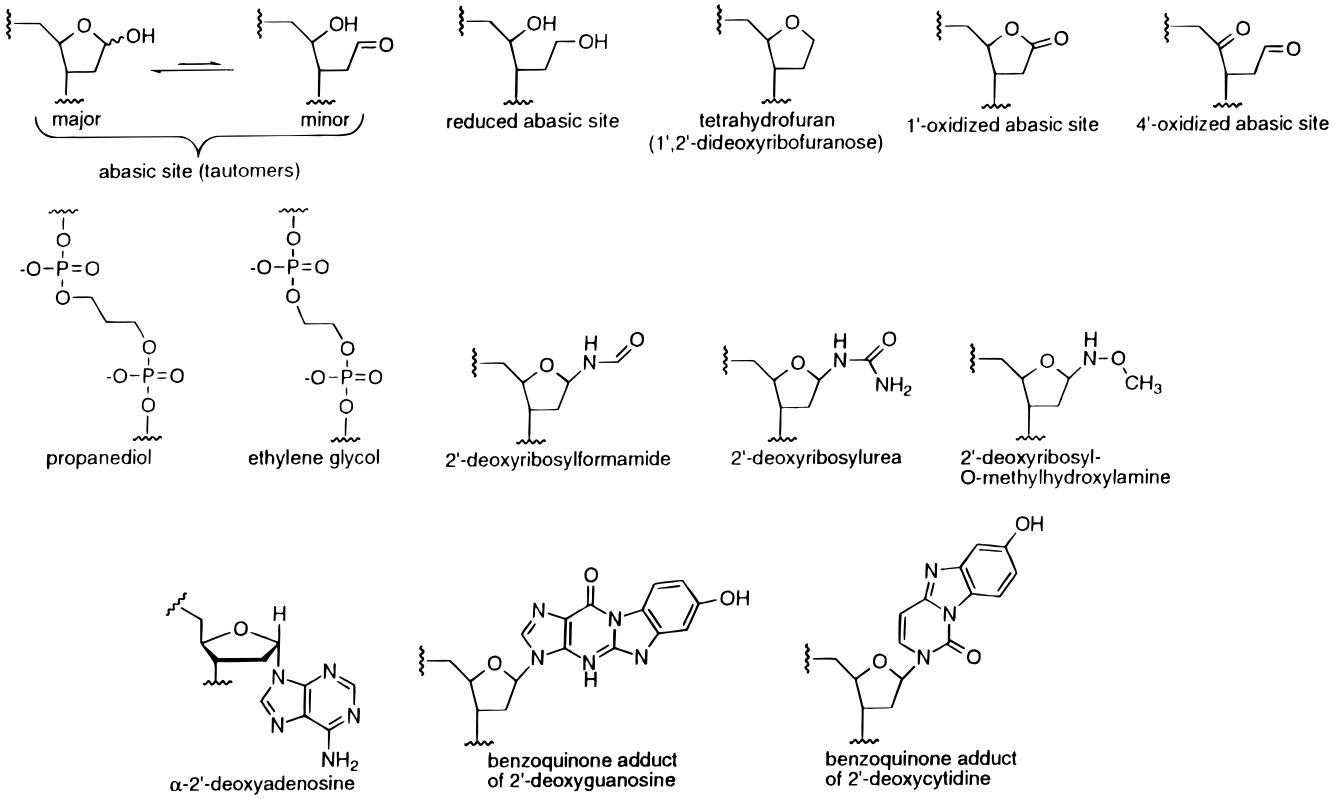

Figure 29. Substrates for class II AP endonucleases. (Modified from ref 13.)

III and this may be relevant to the production of double-strand breaks in DNA.[277] These studies indicate that 5′-phosphodiester cleavage at abasic sites in paired lesions is slower with respect to isolated abasic sites. However, cleavage at abasic sites positioned 3′ with respect to each other can occur, producing a double-strand break. If the abasic sites were located one base apart with a 5′ orientation, excision of only one strand was observed. If, however, the two were located three base pairs apart, neither strand was cleaved. Together, these results suggest the importance of a pocket formed at the abasic site and the structural properties of the duplex surrounding the abasic site. These recognition features are also suggested by the structural studies (vide infra).

C. Mechanistic Properties of Class II AP Endonucleases

Class II AP endonucleases are metalloenzymes that have slightly different metal requirements.[263,278] Endo IV and Apn1 both contain three distinct metal binding sites and the metal ions in these sites exhibit different levels of sensitivity to metal-chelating agents. In the case of endo IV, the cast of characters includes two structural Zn(II) ions and one catalytic Mn(II) ion.[279] The catalytic Mn(II) can be substituted with Co(II) or Ni(II), but not Zn(II). Unlike endo IV, catalytically competent Apn1 can be reconstituted with Zn(II), Co(II), or Mn(II) after removal of endogenous metal ions. Atomic absorption revealed that the native metal constituency of isolated Apn1 is three Zn(II) ions.[279] Both enzymes are resistant to the effects of chelating agents in the presence of DNA, indicating that the critical metal sites may be protected in the endonuclease−DNA complex.[263]

Exo III and HAP1 are also metalloenzymes which appear to contain one tightly associated metal ion.[278] The identity of this metal ion in vivo is most likely Mg(II). However, in vitro Mn(II), and to a lesser extent Ni(II), can restore enzymatic activity to demetalated exo III and HAP1. Structural studies and site-directed mutagenesis of exo III[280] and HAP1[281,282] provide considerable insight into the role of the single metal ion. In the structural studies of a ternary complex of Mn(II), dCMP and exo III,[280] a Mn(II) ion was coordinated to Glu 34 Oϵ1 at the bottom of the proposed active-site groove. Other potential metal ion ligands were not obvious from the structural work; however, the authors suggest that Asp 258 may be able to participate in metal coordination with slight movement. In the crystallographic characterization of HAP1,[282] a Sm(II) ion was found coordinated to Glu 96 (analogous to Glu 34 in exo III). Site-directed mutagenesis on HAP1 indicates that Glu 96 plays a critical role in the catalytic activity of HAP1.[281] A mutated HAP1 form, E96A, displays a 400-fold reduced catalytic activity and requires an 80-fold higher concentration of Mg(II). This suggests that Glu 96 is involved in tight Mg(II) binding and participates indirectly in catalysis. Replacement of Asp 308 in HAP1 (Asp 258 in exo III) with alanine results in a slight reduction of activity and a switch in the preference for Mn(II) rather than Mg(II).[281] This is consistent with Asp 308 playing a minor role in metal ion binding, perhaps by indirect (water-mediated) metal ion binding. These results along with the sensitivity of the enzymatic activity of exo III and HAP1 to EDTA indicate that the Mg(II) (or

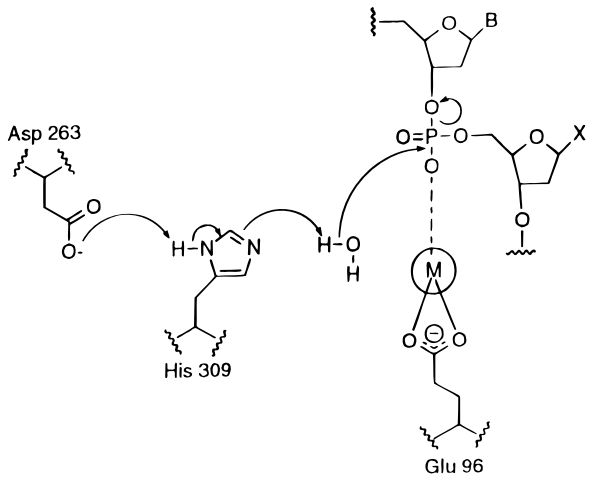

Figure 30. Proposed mechanism for 5′-phosphodiester hydrolysis catalyzed by HAP1 based on crystallographic and biochemical studies. (Adapted from ref 282.)

suitable replacement) plays an important enzymatic role. The catalytic activity is reduced but not abolished when removing the important glutamic acid residue, suggesting that the metal ion need not be tightly coordinated for catalysis.[278] The catalytic Mg(II) ion in exo III can be replaced with substitutionally inert $[Cr(NH_3)_{6-x}(H_2O)_x]^{3+}$ complexes.[283] Only the Cr(III) complexes which have a significant number of water ligands promote the phosphodiester cleavage activity of exo III. These results suggest that the metal cofactor may activate the substrate by outer-sphere stabilization of the transition state.[283]

Additional active-site residues of the two enzymes have been identified on the basis of the crystal structures and site-direct mutagenesis. In HAP1, mutation of four residues, Asp 210, Asn 212, Asp 283, and His 309, nearly eliminates all enzymatic activity, implicating these residues as vital players in catalysis.[278] Kinetic analysis also indicates that Asp 210, Asp 283, and His 309 play catalytic roles;[278] however, Asn 212 participates in DNA binding.[284] On the basis of the structural studies of DNase I[285] and exo III,[282] an Asp/His pair was suggested to participate in the activation of a water molecule for in-line attack on the phosphodiester bond. Within the HAP1 active site,[282] the imidazole ring of His 309 interacts with the carboxylate of Asp 283, which in turn interacts with Thr 265 This architecture is consistent with the suggested role of the Asp/His pair. In the proposed mechanism illustrated in Figure 30,[282] the hydrogen-bonding of Asp 283 with His 309 increases its basicity thereby facilitating its role as the general base for deprotonating an active-site water nucleophile. The metal ion coordinated to Glu 96 is proposed to participate by stabilizing the developing negative charge on the phosphate group in the transition state. Asp 210 has also been suggested to be in the vicinity of a modeled AP site, and may participate in protonating or stabilizing the departing phosphate monoester.[282] This proposed mechanism is reminiscent of the mechanism for serine proteases involving an Asp-His-Ser catalytic triad.[286]

D. Insight into Recognition of Abasic Sites by Exo III and HAP1

The substrate specificity of class II AP endonucleases has provided insight into some of the required recognition features. These features suggest that a pocket or space created by a baseless site, or a particular base modification within a DNA duplex, is recognized by these enzymes. Site-directed mutagenesis of Asn 212 in HAP1 has implicated this residue in AP site recognition.[268,274] The authors took advantage of the ability of metal-free HAP1 to bind to an AP site-containing duplex without enzymatic turnover to determine the binding properties of the mutant enzymes, N212A, N212Q, and N212D. These results indicated that all three mutants lost the ability to bind to substrate abasic site DNA, implicating the critical role of Asn 212. These results were also consistent with the crystallographic results of Mol et al.[280] on exo III in which Asn 153 (analogous to Asn 212 in HAP1) was found to hydrogen bond to the nucleotide O3′ and the 5′-phosphate group of dCMP in the ternary complex.

On the basis of the exo III structure, Mol et al.[280] proposed that recognition features of an AP site may include the extrahelical base or distortions of the sugar−phosphate backbone. In the recent structural studies of HAP1, its similarity to the three-dimensional structures of DNase I[285] was exploited to model a structure of a HAP1/AP site-containing DNA complex.[282] The cocrystal structure of DNase I with noncleaved and cleaved DNA fragments had been previously determined.[287,288] The structural studies indicated that the DNA−protein interactions are dominated by electrostatic interactions with the phosphate backbone and nonspecific hydrophobic contacts. Furthermore, DNase I binds in the minor groove of DNA promoting DNA bending and local perturbations at the abasic site in the DNaseI/abasic site complex.[287] Thus, HAP1 binding to an abasic site-containing duplex was modeled on the basis of the DNase I−DNA complex using conserved residues between the two enzymes which are likely to participate in similar contacts. A number of residues were found to be capable of making analogous phosphate contacts and stacking interactions with base pairs surrounding the abasic site. The modeled HAP1−DNA complex was also consistent with the methylation and ethylation interference experiments which indicate contacts of HAP1 with phosphate groups of 2−3 nucleotides flanking both sides of the abasic site.[289] On the basis of this model, the authors suggest that recognition of an extrahelical base *opposite* the abasic site seems unlikely. Instead, they propose that the enzyme specifically recognizes the abasic site itself. In the modeled HAP1−DNA complex, the AP deoxyribose could rotate to an extrahelical conformation and stack against Phe 266. A similar stacking was observed between Trp 212 (in analogous position of Phe 266 in HAP1) and the deoxyribose of dCMP in the exo III−dCMP complex X-ray structure. Interestingly, DNase I, which does not recognize abasic sites, does not contain an aromatic residue at this position. Consistent with this proposal, indirect evidence for DNA distortion

at the abasic site has been observed in Cu^{II}-1,10-phenanthroline (Cu-phen) footprinting experiments in which the 5′-phosphodiester bond of the abasic site is hyperreactive to oxidative Cu-phen cleavage in the HAP1−DNA complex.[289] The structural and biochemical studies on HAP1 and exo III suggest a number of features that are strikingly similar to those of the BER glycosylase enzymes discussed previously. In particular, the features of a flipped-out nucleotide (with or without an appended base) and DNA deformation emerge as important recognition features which may facilitate rapid location of DNA damage.

VI. Prospectus for the Future

In this review, we have focused on the literature available regarding the chemical properties of DNA repair enzymes. The increased sophistication in understanding of their biological properties will also greatly facilitate the understanding of their chemistry. To delineate the biological significance of BER enzymes, recent efforts have been directed toward preparing mice lacking specific DNA repair activities. Inactivation of the gene for the major AP endonuclease (analogous to HAP1) in mice results in defective embryogenesis such that no viable mice are produced.[290] These unexpected results suggest that this AP endonuclease may have additional roles in the cell that are critical for embryonic development. In contrast, recent inactivation of the gene for 3-methyladenine glycosylase (MPG) results in normally developed mice.[291] Cell extracts from these mice are devoid in repair activity toward 3-methyladenine, 1,N^6-ethenoadenine, and hypoxanthine, and exhibit hypersensitivity to alkylating agents. These results show that MPG is the major repair glycosylase for alkylation damage in mice. The viability of these repair-deficient mice provides an exciting opportunity to investigate the in vivo biological properties of repair, and how these properties affect normal development of an organism as well as disease states and aging. The outcomes of such studies and the development of animal models for other repair enzymes is eagerly anticipated. An increased understanding of their biological properties will raise new questions relating to the fundamental properties of these enzymes.

Insight into the chemical properties of BER glycosylases and endonucleases has been provided by the structural information now available for these enzymes. It is anticipated that as more cocrystal structures become available, our understanding of the recognition of DNA damage and repair mechanisms will be significantly enhanced. The time is right in this field for an increasing role of chemical approaches to studying these enzymes. These approaches will also provide insight into the biological roles of these enzymes. As discussed in this review, the preparation of substrate and transition-state analogues has been initiated, and structural information for these molecules bound to their respective DNA repair glycosylase/endonucleases is also eagerly awaited. In addition, increased efforts are needed in quantitatively defining the enzymatic properties of BER enzymes toward different damaged substrates. Such studies may shed light on the role the intrinsic chemistry plays in determining the biological properties of a given type of DNA damage. For example, correlations between the in vitro and in vivo repair activities on specific substrates may help in determining the relative importance of repair of specific types of DNA damage in carcinogenesis and aging.

It is clear that exciting new results in this area will be forthcoming and that advancement will be made from researchers hailing from a wide variety of biological and chemical fields. As new pieces of the BER puzzle are found and the fascinating story of DNA repair continues to unfold, we feel fortunate to be both participants and fans.

VII. Abbreviations

AP	apurinic-apyrimidinic
ANPG, MPG	alkyl-N-purine glycosylase or 3-methyladenine glycosylase (mammals)
B. subtilis	*Bacillus subtilis*
BER	base-excision repair
DMS	dimethyl sulfate
dsUDG	double-strand specific uracil-DNA glycosylases
E. coli	*Escherichia coli*
endo III	endonuclease III
endo IV	endonuclease IV
endo VIII	endonuclease VIII
exo III	exonuclease III
Fapy-adenine	4,6-diamino-5-formamidopyrimidine
Fapy-guanine	2,6-diamino-4-hydroxy-5-formamidopyrimidine
HAP1	human AP endonuclease
hOGG1	human OG glycosylase 1
M. thermoautotrophicum	*Methanobacterium thermoautotrophicum*
M. luteus	*Micrococcus luteus*
MYH	MutY homologue
N. crassa	*Neurospora crassa*
NTG1	endonuclease III-like glycosylase 1
Nth-spo	endonuclease III-*S. pombe*
OA	7,8-dihydro-8-oxo-2′-deoxyadenosine
OG	7,8-dihydro-8-oxo-2′-deoxyguanosine
PAGE	polyacrylamide gel electrophoresis
S. cerevisiae	*Saccharomyces cerevisiae*
SCR1	*S. cerevisiae* redoxyendonuclease 1
SCR2	*S. cerevisiae* redoxyendonuclease 2
SDS	sodium dodecyl sulfate
S. pombe	*Schizosaccharomyces pombe*
T4-endo V	T4-endonuclease V
TDG	mismatch specific thymine glycosylase
UDG	uracil-DNA glycosylase

VIII. Acknowledgments

We are extremely appreciative of Nik Chmiel and Cindy Lou Chepanoske for helping with the figures

for this review. We also greatly value the comments and suggestions of Dr. Jim Müller, Professor Cindy Burrows, Professor Peter Beal, and Ms. Selena Boothroyd. Drs. J. Tainer and C. Mol kindly provided the coordinates of the hUDG-DNA complex. Professor Tom Ellenberger graciously provided the coordinates of the AlkA structure. The contributions of members of the David laboratory in proof reading the manuscript are also gratefully acknowledged. The work in S.S.D.'s laboratory on DNA repair is supported by the National Institutes of Health CA67985 and the University of Utah Research Foundation. S.S.D.'s start in this area was facilitated by an Arnold and Mabel Beckman Foundation Young Investigator Award (1993-1996). S.S.D. is presently an Alfred P. Sloan Research Fellow (1998-2000).

IX. References

(1) Lindahl, T. *Nature* **1993**, *362*, 709−715.
(2) Lehmann, A. R. *Trends Biochem. Sci.* **1995**, *20*, 402−405.
(3) Friedberg, E. C.; Walker, G. C.; Siede, W. *DNA Repair and Mutagenesis*; ASM Press: Washington, DC, 1995.
(4) Modrich, P.; Lahue, R. *Annu. Rev. Biochem.* **1996**, *65*, 101−133.
(5) Lu, R.; Nash, H. M.; Verdine, G. L. *Curr. Biol.* **1997**, *7*, 397−407.
(6) Bowman, K. K.; Sidik, K.; Smith, C. A.; Taylor, J.-S.; Doetsch, P. W.; Freyer, G. A. *Nucleic Acids Res.* **1994**, *22*, 3026−3032.
(7) Doetsch, P. W. *Trends Biochem. Sci.* **1995**, *20*, 384−386.
(8) Freyer, G. A.; Davey, S.; Ferrer, J. V.; Martin, A. M.; Beach, D.; Doetsch, P. W. *Mol. Cell Biol.* **1995**, *15*, 4572−4577.
(9) Yajima, H.; Takao, M.; Yasuhira, S.; Zhao, J. H.; Ishi, C.; Inoue, H.; Yasui, A. *EMBO J.* **1995**, *14*, 2393−2399.
(10) Takao, M.; Yonemasu, R.; Yamamoto, K.; Yasui, A. *Nucleic Acids Res.* **1996**, *24*, 1267−1271.
(11) Yonemasu, R.; McCready, S.; Murray, J. M.; Osman, F.; Takao, M.; Yamamoto, K.; Lehmann, A. R.; Yasui, A. *Nucleic Acids Res.* **1997**, *25*, 1553−1558.
(12) Sancar, A. *Biochemistry* **1994**, *33*, 2−9.
(13) Singer, B.; Hang, B. *Chem. Res. Toxicol.* **1997**, *10*, 713−732.
(14) Krokan, H. E.; Standal, R.; Slupphaug, G. *Biochem. J.* **1997**, *325*, 1−16.
(15) Singhal, R. K.; Prasad, R.; Wilson, S. H. *J. Biol. Chem.* **1995**, *270*, 949−957.
(16) Dianov, G.; Lindahl, T. *Curr. Biol.* **1994**, *4*, 1069−1076.
(17) Dianov, G.; Price, A.; Lindahl, T. *Mol. Cell. Biol.* **1992**, *12*, 1605−1612.
(18) Wiebauer, K.; Jiricny, J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 5842−5845.
(19) Wood, R. D.; Shviji, M. K. K. *Carcinogenesis* **1997**, *18*, 605−610.
(20) Nealon, K.; Nicholl, I. D.; Kenny, M. K. *Nucleic Acids Res.* **1996**, *24*, 3763−3770.
(21) Bennett, R. O.; Wilson, D. M.; Wong, D.; Demple, B. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7166−7169.
(22) Caldecott, K. W.; Aoufouchi, S.; Johnson, P.; Shall, S. *Nucleic Acids Res.* **1996**, *24*, 4387−4394.
(23) Kubota, Y.; Nash, R. A.; Klungland, A.; Schar, P.; Barnes, D.; Lindahl, T. *EMBO J.* **1996**, *15*, 6662−6670.
(24) Frosina, G.; Fortini, P.; Rossi, O.; Carrozzino, F.; Raspaglio, G.; Cox, L. S.; Lane, D. P.; Abbondandol, A.; Dogliotti, E. *J. Biol. Chem.* **1996**, *271*, 9573−9578.
(25) Matsumoto, Y.; Kim, K.; Bogenhagen, D. F. *Mol. Cell Biol.* **1994**, *14*, 6187−6197.
(26) Klungland, A.; Lindahl, T. *EMBO J.* **1997**, *16*, 3341−3348.
(27) Lindahl, T.; Nyberg, B. *Biochemistry* **1974**, *13*, 3405−3410.
(28) Krokan, H. E.; Standal, R.; Bharati, S.; Otterlei, M.; Haug, T.; Slupphaug, G.; Skorpen, F. In *Base Excision Repair of DNA Damage*; Hickson, I. D., Ed.; Landes Bioscience: Austin, TX, 1997; pp 45−66.
(29) Lindahl, T. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3649−3653.
(30) Tomlin, N. V.; Aprelikova, O. N. *Int. Rev. Cytol.* **1989**, *114*, 125−179.
(31) Fix, D. F.; Glickman, B. W. *Mutat. Res.* **1986**, *175*, 41−45.
(32) Olsen, L. C.; Aasland, R.; Wittwer, C. U.; Krokan, H. E.; Helland, D. E. *EMBO J* **1989**, *8*, 3121−3125.
(33) Kumar, N. V.; Varshney, U. *Nucleic Acids Res.* **1994**, *22*, 3737−3741.
(34) Kumar, N. V.; Varshney, U. *Nucleic Acids Res.* **1997**, *25*, 2336−2343.
(35) Lindahl, T.; Ljungquist, S.; Siegert, W.; Nyberg, B.; Sperens, B. *J. Biol. Chem.* **1977**, *252*, 3286−3294.
(36) Varshney, U.; van de Sande, J. H. *Biochemistry* **1991**, *30*, 4055−4060.
(37) Mauro, D. J.; De Riel, J. K.; Tallarida, R. J.; Sirover, M. A. *Mol. Pharmacol.* **1993**, *43*, 854−857.
(38) Wurzer, J. C.; Tallarida, R. J.; Sirover, M. A. *J. Pharm. Exp. Ther.* **1994**, *269*, 39−43.
(39) Dizdaroglu, M.; Karakaya, A.; Jaruga, P.; Slupphaug, G.; Krokan, H. E. *Nucleic Acids Res.* **1996**, *24*, 418−422.
(40) Zastawny, T. H.; Doetsch, P. W.; Dizdaroglu, M. *FEBS Lett.* **1995**, *364*, 255−258.
(41) Hatahet, Z.; Kow, Y. W.; Purmal, A.; Cunningham, R. P.; Wallace, S. S. *J. Biol. Chem.***1994**, *269*, 18814−18820.
(42) Krokan, H.; Wittwer, C. U. *Nucleic Acids Res.* **1981**, *9*, 2598−2613.
(43) Eftedal, I.; Guddal, P. H.; Slupphaug, G.; Volden, G.; Krokan, H. E. *Nucleic Acids Res.* **1993**, *21*, 2095−2101.
(44) Bennett, S. E.; Sanderson, R. J. *Biochemistry* **1995**, *34*, 6109−6119.
(45) Slupphaug, G.; Eftedal, I.; Kavli, B.; Bharati, S.; Helle, N. M.; Haug, T.; Levine, D. W.; Krokan, H. E. *Biochemistry* **1995**, *34*, 128−138.
(46) Nilsen, H.; Yazdankhah, S. P.; Eftedal, I.; Krokan, H. E. *FEBS Lett.* **1995**, *362*, 205−209.
(47) Lettieri, T.; Jiricny, J. In *Base Excision Repair of DNA Damage*; Hickson, I. D., Ed.; Landes Bioscience: Austin, TX, 1997; pp 45−66.
(48) Brown, T. C.; Jiricny, J. *Cell* **1987**, *50*, 945−950.
(49) Jiricny, J.; Hughes, M.; Corman, N.; Rudkin, B. B. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8860−8864.
(50) Neddermann, P.; Jiricny, J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1642−1646.
(51) Nedderman, P.; Jiricny, J. *J. Biol. Chem.* **1993**, *268*, 21218−21224.
(52) Neddermann, P.; Gallinari, P.; Lettieri, T.; Schmid, D.; Truong, O.; Hsuan, J. J.; Wiebauer, K.; Jiricny, J. *J. Biol. Chem.* **1996**, *271*, 12767−12774.
(53) Sibghat-Ullah; Day, R. S. I. *Nucleic Acids Res.* **1993**, *21*, 1281−1287.
(54) Sibghat-Ullah; Day, R. S. I. *Biochemistry* **1995**, *34*, 6869−6875.
(55) Sibghat-Ullah; Gallinari, P.; Xu, Y.-Z.; Goodman, M. F.; Bloom, L. E.; Jiricny, J.; Day, R. S. *Biochemistry* **1996**, *35*, 12926−12932.
(56) Griffin, S.; Branch, P.; Xu, Y.-Z.; Karran, P. *Biochemistry* **1994**, *33*, 4787−4793.
(57) Sibghat-Ullah; Xu, Y.-Z.; Day, R. S. I. *Biochemistry* **1995**, *34*, 7438−7442.
(58) Gallinari, P.; Jiricny, J. *Nature* **1996**, *383*, 735−738.
(59) Seeberg, E.; Berdal, K. G. In *Base Excision Repair of DNA Damage*; Hickson, I. D., Ed.; Landes Bioscience: Austin, TX, 1997; pp 151−168.
(60) Sakumi, K.; Sekiguchi, M. *Mutat. Res.* **1990**, *236*, 161−172.
(61) Morohoshi, F.; Hayashi, K.; Munakata, N. *J. Bacteriology* **1993**, *175*, 6010−6017.
(62) Bjørås, M.; Klungland, A.; Johansen, R. F.; Seeberg, E. *Biochemistry* **1995**, *34*, 4577−4582.
(63) Goth-Goldstein, R.; Johnson, P. L. *Mol. Gen. Genet.* **1990**, *221*, 353−357.
(64) Memisoglu, A.; Samson, L. *Gene* **1996**, *177*, 229−235.
(65) Margison, G. P.; Pegg, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 861−865.
(66) Engelward, B. P.; Boosalis, M. S.; Chen, B. J.; Deng, Z.; Siciliano, M. J.; Samson, L. D. *Carcinogenesis* **1993**, *14*, 175−181.
(67) Singer, B.; Brent, T. P. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 856−860.
(68) Samson, L.; Derfler, B.; Boosalis, M.; Call, K. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9127−9131.
(69) O'Connor, T. R.; Laval, F. *EMBO J.* **1990**, *9*, 3337−3342.
(70) Santerre, A.; Britt, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2240−2244.
(71) Berdal, K. G.; Bjørås, M.; Bjelland, S.; Seeberg, E. *EMBO J.* **1990**, *9*, 4563−4568.
(72) Chen, J.; Derfler, B.; Samson, L. *EMBO J.* **1990**, *9*, 4569−4575.
(73) Matijasevic, Z.; Boosalis, M.; MacKay, W.; Samson, L.; Ludlum, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11855−11859.
(74) O'Connor, T. R. *Nucleic Acids Res.* **1993**, *21*, 5561−5569.
(75) Nakabeppu, Y.; Miyata, T.; Kondo, H.; Iwanaga, S.; Sekiguchi, M. *J. Biol. Chem.* **1984**, 13736−13790.
(76) Vickers, M. A.; Vyas, P.; Harris, P. C.; Simmons, D. L.; Higgs, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3437−3441.
(77) Rydberg, B.; Lindahl, T. *EMBO J.* **1982**, *1*, 211−216.
(78) Boiteux, S.; Huisman, O.; Laval, J. *EMBO J.* **1984**, *3*, 2569−2573.
(79) Mattes, W. B.; Lee, C.-S.; Laval, J.; O'Connor, T. R. *Carcinogenesis* **1996**, *17*, 643−648.
(80) Matijasevic, Z.; Stering, A.; Niu, T.-Q.; Austin-Ritchie, P.; Ludlum, D. B. *Carcinogenesis* **1996**, *17*, 2249−2252.
(81) Habraken, Y.; Ludlum, D. B. *Carcinogenesis* **1989**, *10*, 489−492.
(82) Habraken, Y.; Carter, C. A.; Sekiguchi, M.; Ludlum, D. B. *Carcinogenesis* **1991**, *12*, 1971−1973.

(83) Matijavesic, Z.; Sekiguchi, M.; Ludlum, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9331−9334.
(84) Singer, B.; Antoccia, A.; Basu, A. K.; Dosanjh, M. K.; Fraenkel-Conrat, H.; Gallagher, P. E.; Kusmierek, J. T.; Qiu, Z.-H.; Rydberg, B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9386−9390.
(85) Saparbaev, M.; Kleibl, K.; Laval, J. *Nucleic Acids Res.* **1995**, *23*, 3750−3755.
(86) Dosanjh, M. K.; Roy, R.; Mitra, S.; Singer, B. *Biochemistry* **1994**, *33*, 1624−1628.
(87) Bjelland, S.; Birkeland, N.-K.; Benneche, T.; Volden, G.; Seeberg, E. *J. Biol. Chem.* **1994**, *269*, 30489−30495.
(88) Bessho, T.; Roy, R.; Yamamoto, K.; Kasai, H.; Nishimura, S.; Tano, K.; Mitra, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8901−8904.
(89) Bjelland, S.; Bjørås, M.; Seeberg, E. *Nucelic Acids Res.* **1993**, *21*, 2045−2049.
(90) Sakumi, K.; Nakabeppu, Y.; Yamamoto, Y.; Kawabata, S.-i.; Iwanaga, S.; Sekiguchi, M. *J. Biol. Chem.* **1986**, *261*, 15761−15766.
(91) Bjelland, S.; Seeberg, E. *FEBS Lett.* **1996**, *397*, 127−129.
(92) Radman, M. *J. Biol. Chem.***1976**, *251*, 1438−1445.
(93) Armel, P. R.; Wallace, S. S. *Nucleic Acids Res.* **1977**, *5*, 3347−3356.
(94) Breimer, L.; Lindahl, T. *Nucleic Acids Res.* **1980**, *8*, 6199−6211.
(95) Katcher, H. L.; Wallace, S. S. *Biochemistry* **1983**, *22*, 4072−4081.
(96) Demple, B.; Linn, S. *Nature* **1980**, *287*, 203−208.
(97) Boorstein, R. J.; Hilbert, T. P.; Cadet, J.; Cunningham, R. P.; Teebor, G. W. *Biochemistry* **1989**, *28*, 6164−6170.
(98) Breimer, L. H.; Lindahl, T. *J. Biol. Chem.* **1984**, *259*, 5543−5548.
(99) Dizdaroglu, M.; Laval, J.; Boiteux, S. *Biochemistry* **1993**, *32*, 12105−12111.
(100) Wagner, J. R.; Blount, B. C.; Weinfeld, M. *Anal. Biochem.* **1996**, *233*, 76−86.
(101) Breimer, L. H.; Lindahl, T. *Biochemistry* **1985**, *24*, 4018−4022.
(102) Jen, J.; Mitchell, D. L.; Cunningham, R. P.; Smith, C. A.; Taylor, J.-S.; Cleaver, J. E. *Photochem. Photobiol.* **1997**, *65*, 323−329.
(103) Wang, D.; Essigmann, J. M. *Biochemistry* **1997**, *36*, 8628−8633.
(104) Cunningham, R. P.; Weiss, B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 474−478.
(105) Gossett, J.; Lee, J.; Cunningham, R. P.; Doetsch, P. W. *Biochemistry* **1988**, *27*, 2629−2634.
(106) Augeri, L.; Hamilton, K. K.; Martin, A. M.; Yohannes, P.; Doetsch, P. W. *Methods Enzymol.* **1994**, *234*, 102−114.
(107) Doestch, P. W.; Henner, W. D.; Cunningham, R. P.; Toney, J.; Helland, D. E. *Mol. Cell. Biol.* **1987**, *7*, 26−32.
(108) Higgins, S. A.; Frenkel, K.; Cummings, A.; Teebor, G. W. *Biochemistry* **1987**, *26*, 1683−1688.
(109) Augeri, L.; Lee, Y.-M.; Barton, A. B.; Doetsch, P. W. *Biochemistry* **1997**, *36*, 721−729.
(110) Eide, L.; Bjørås, M.; Pirovano, M.; Alseth, I.; Berdal, K. G.; Seeberg, E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10735−10740.
(111) Roldan-Arjona, T.; Anselmino, C.; Lindahl, T. *Nucleic Acids Res.* **1996**, *24*, 3307−3312.
(112) Doestch, P. W.; Helland, D. E.; Haseltine, W. A. *Biochemistry* **1986**, *25*, 2212−2220.
(113) Hilbert, T. P.; Boostein, R. J.; Kung, H. C.; Bolton, P. H.; Xing, D.; Cunningham, R. P.; Teebor, G. W. *Biochemistry* **1996**, *35*, 2505−2511.
(114) Hilbert, T. P.; Chaung, W.; Boorstein, R. J.; Cunningham, R. P.; Teebor, G. W. *J. Biol. Chem.* **1997**, *272*, 6733−6740.
(115) Melamede, R. J.; Hatahet, Z.; Kow, Y. W.; Ide, H.; Wallace, S. S. *Biochemistry* **1994**, *33*, 1255−1264.
(116) Jiang, D.; Hatahet, Z.; Blaisdell, J. O.; Melamede, R. J.; Wallace, S. S. *J. Bacteriol.* **1997**, *179*, 3773−3782.
(117) Dodson, M. L.; Lloyd, R. S. *Mutat. Res.* **1989**, *218*, 49−65.
(118) Yasuda, S.; Sekiguchi, M. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *67*, 1839−1845.
(119) Minton, K.; Durphy, M.; Taylor, R.; Friedberg, E. C. *J. Biol. Chem.* **1975**, *250*, 2823−2829.
(120) Friedberg, E. C.; King, J. J. *Biochem. Biophys. Res. Commun.* **1969**, *37*, 646−651.
(121) Friedberg, E. C. *Mutat. Res.* **1972**, *15*, 113−123.
(122) Radany, E. H.; Friedberg, E. C. *Nature* **1980**, *286*, 182−185.
(123) McMillan, S.; Edenberg, H. J.; Radany, E. H.; Friedberg, R. C.; Friedberg, E. C. *J. Virol.* **1981**, *40*, 211−223.
(124) Nakabeppu, Y.; Sekiguchi, M. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2742−2746.
(125) Seawell, P. C.; Smith, C. A.; Ganesan, A. K. *J. Virol.* **1980**, *35*, 790−797.
(126) Valerie, K.; Henderson, E. E.; deRiel, J. K. *Nucleic Acids Res.* **1984**, *12*, 8085−8097.
(127) Riazuddin, S.; Grossman, L. *J. Biol. Chem.* **1977**, *252*, 6280−6286.
(128) Riazuddin, S.; Grossman, L. *J. Biol. Chem.* **1977**, *252*, 6287−6293.
(129) Haseltine, W. A.; Gordon, L. K.; Lindan, C. P.; Grafstrom, R. H.; Shaper, N. L.; Grossman, L. *Nature* **1980**, *285*, 634−641.
(130) Grafstrom, R. H.; Park, L.; Grossman, L. *J. Biol. Chem.* **1982**, *257*, 13465−13474.
(131) Hamilton, K. K.; Kim, P. M. H.; Doetsch, P. W. *Nature* **1992**, *356*, 725−728.
(132) Patrick, M. H.; Rahn, R. O. In *Photochemistry and Photobiology of Nucleic Acids*; Academic Press: New York, 1976; Vol. 2, pp 35−95.
(133) Gordon, L. K.; Haseltine, W. A. *J. Biol. Chem.* **1980**, *255*, 12047−12050.
(134) Piersen, C. E.; Prince, M. A.; Augustine, M. L.; Dodson, M. L.; Lloyd, R. S. *J. Biol. Chem.* **1995**, *270*, 23475−23484.
(135) Smith, C. A. *J. Biol. Chem.* **1993**, *268*, 11143−11151.
(136) Dizdaroglu, M.; Zastawny, T.; Carmical, J. R.; Lloyd, R. S. *Mutat. Res.* **1996**, *362*, 1−8.
(137) Gallagher, P. E.; Duker, N. J. *Mol. Cell Biol.* **1986**, *6*, 707−709.
(138) Shiota, S.; Nakayama, H. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 593−598.
(139) Chetsanga, C. J.; Lindahl, T. *Nucleic Acids Res.* **1979**, *6*, 3673−3683.
(140) Laval, J.; Boiteux, S.; O'Connor, T. R. *Mutat. Res.* **1990**, *233*, 73−79.
(141) Boiteux, S.; Belleney, J.; Roques, B. P.; Laval, J. *Nucleic Acids Res.* **1984**, *12*, 5429−5439.
(142) O'Connor, T. R.; Boiteux, S.; Laval, J. *Nucleic Acids Res.* **1988**, *16*, 5879−5894.
(143) Boiteux, S.; O'Connor, T. R.; Lederer, F.; Gouyette, A.; Laval, J. *J. Biol. Chem.***1990**, *265*, 3916−3922.
(144) Graves, R. J.; Felzenszwalb, I.; Laval, J.; O'Connor, T. R. *J. Biol. Chem.***1992**, *267*, 14429−14435.
(145) Boiteux, S.; Gajewski, E.; Laval, J.; Dizdaroglu, M. *Biochemistry* **1992**, *31*, 106−110.
(146) Karakaya, A.; Jaruga, P.; Bohr, V. A.; Grollman, A. P.; Dizdaroglu, M. *Nucleic Acids Res.* **1997**, *25*, 474−479.
(147) Chetsanga, C. J.; Frenette, G. P. *Carcinogenesis* **1983**, *4*, 997−1000.
(148) Oleykowski, C. A.; Mayernik, J. A.; Lim, S. E.; Groopman, J. D.; Grossman, L.; Wogan, G. N.; Yeung, A. T. *J. Biol. Chem.* **1993**, *268*, 7990−8002.
(149) Li, Q.; Laval, J.; Ludlum, D. B. *Carcinogenesis* **1997**, *18*, 1035−1038.
(150) Chung, M. H.; Kasai, H.; Jones, D. S.; Inoue, H.; Ishikawa, H.; Ohtsuka, E.; Nishimura, S. *Mutat. Res.* **1991**, *254*, 1−12.
(151) Tchou, J.; Kasai, H.; Shibutani, S.; Chung, M.-H.; Laval, J.; Grollman, A. P.; Nishimura, S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 4690−4694.
(152) Burrows, C. M.; Muller, J. G. *Chem. Rev.* **1998**, *98*, 1109−1151.
(153) Grollman, A. P.; Moriya, M. *Trends. Genet.* **1993**, *9*, 246−249.
(154) Michaels, M. L.; Pham, L.; Cruz, C.; Miller, J. H. *Nucleic Acids Res.* **1991**, *19*, 3629−3632.
(155) Michaels, M. L.; Miller, J. H. *J. Bacteriol.* **1992**, *174*, 6321−6325.
(156) Tchou, J.; Grollman, A. P. *Mutat. Res.* **1993**, *299*, 277−287.
(157) Shibutani, S.; Takeshita, M.; Grollman, A. P. *Nature* **1991**, *349*, 431−434.
(158) Besho, T.; Tano, K.; Kasai, H.; Nishimura, S. *Biochem. Biophys. Res. Commun.* **1992**, *188*, 372−378.
(159) Boiteux, S.; O'Connor, T. R.; Laval, J. *EMBO J.* **1987**, *6*, 3177−3183.
(160) Michaels, M. L.; Tchou, J.; Grollman, A. P.; Miller, J. H. *Biochemistry* **1992**, *31*, 10964−10968.
(161) Boiteux, S. *J. Photochem. Photobiol.* **1993**, *19*, 87−96.
(162) Tchou, J.; Bodepudi, V.; Shibutani, S.; Antoscheckin, I.; Miller, J.; Grollman, A. P.; Johnson, F. P. *J. Biol. Chem.* **1994**, *269*, 15318−15324.
(163) Castaing, B.; Geiger, A.; Seliger, H.; Nehls, P.; Laval, J.; Zelwer, C.; Boiteux, S. *Nucleic Acids Res.* **1993**, *21*, 2899−2905.
(164) Graves, R.; Laval, J.; Pegg, A. E. *Carcinogenesis* **1992**, *13*, 1455−1459.
(165) de Oliviera, R.; van der Kemp, P. A.; Thomas, D.; Geiger, A.; Nehls, P.; Boiteux, S. *Nucleic Acids Res.* **1994**, *22*, 3760−3764.
(166) van Der Kemp, P. A.; Thomas, D.; Barbey, R.; De Oliveira, R.; Boiteux, S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5197−5202.
(167) Nash, H. M.; Bruner, S. D.; Shärer, O. D.; Kawate, T.; Addona, T. A.; Spooner, E.; Lane, W. S.; Verdine, G. L. *Curr. Biol.* **1996**, *6*, 1230−1233.
(168) Girard, P.-M.; Guibourt, N.; Boiteux, S. *Nucleic Acids Res.* **1997**, *25*, 3204−3211.
(169) Sandigursky, M.; Yacoub, A.; Kelley, M. R.; Xu, Y.; Franklin, W. A.; Deutsch, W. A. *Nucleic Acids Res.* **1997**, *25*, 4557−4561.
(170) Radicella, J. P.; Dherin, C.; Desmaze, C.; Fox, M. S.; Boiteux, S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8010−8015.
(171) Arai, K.; Morishita, K.; Shinmura, K.; Kohno, T.; Kim, S.-R.; Nohmi, T.; Taniwaki, M.; Ohwada, S.; Yokota, J. *Oncogene* **1997**, *14*, 2857−2861.
(172) Aburatani, H.; Hippo, Y.; Ishida, T.; Takashima, R.; Matsuba, C.; Kodama, T.; Takao, M.; Yasui, A.; Yamamoto, K.; Asano, M.; Fukusawa, K.; Yoshinari, T.; Inoue, H.; Ohtsuka, E.; Nishimura, S. *Cancer Res.* **1997**, *57*, 2151−2157.

(173) Bjørås, M.; Luna, L.; Johnsen, B.; Hoff, E.; Haug, T.; Tognes, T.; Seeberg, E. *EMBO J.* **1997**, *16*, 6314−6322.
(174) Roldan-Arjona, T.; Wei, F.; Carter, K. C.; Klungland, A.; Anselmino, C.; Wang, R.-P.; Augustus, M.; Lindahl, T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8016−8020.
(175) Rosenquist, T. A.; Zharkov, D. O.; Grollman, A. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7429−7434.
(176) Shinmura, K.; Kasai, H.; Sasaki, A.; Sugimura, H.; Yokota, J. *Mutat. Res.* **1997**, *385*, 75−82.
(177) Modrich, P. *Annu. Rev. Genet.* **1991**, *25*, 229−253.
(178) Su, S.-S.; Lahue, R. S.; Au, K. G.; Modrich, P. *J. Biol. Chem.* **1988**, *263*, 6829−6835.
(179) Lu, A.-L.; Chang, D.-Y. *Genetics* **1988**, *118*, 593−600.
(180) Nghiem, Y.; Cabrera, M.; Cupples, C. G.; Miller, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2709−2713.
(181) Au, K. G.; Cabrera, M.; Miller, J. H.; Modrich, P. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9163−9166.
(182) Au, K. G.; Clark, S.; Miller, J. H.; Modrich, P. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8877−8881.
(183) Lu, A.-L.; Chang, D.-Y. *Cell* **1988**, *54*, 805−812.
(184) Michaels, M. L.; Cruz, C.; Grollman, A. P.; Miller, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7022−7025.
(185) Radicella, J. P.; Clark, E. A.; Fox, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9674−9678.
(186) Tsai-Wu, J.-J.; Liu, H.-F.; Lu, A.-L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8779−8783.
(187) Lu, A.-L.; Tsai-Wu, J.-J.; Cillo, J. *J. Biol. Chem.* **1995**, *270*, 23582−23588.
(188) Bulychev, N. V.; Varaprasad, C. V.; Dorman, G.; Miller, J. H.; Eisenberg, M.; Grollman, A. P.; Johnson, F. *Biochemistry* **1996**, *35*, 13147−13156.
(189) Manuel, R. C.; Lloyd, R. S. *Biochemistry* **1997**, *36*, 11140−11152.
(190) Porello, S. L.; Williams, S. D.; Kuhn, H.; Michaels, M. L.; David, S. S. *J. Am. Chem. Soc.* **1996**, *118*, 10684−10692.
(191) Porello, S. L.; David, S. S. Unpublished results.
(192) Michaels, M. L.; Pham, L.; Nghiem, Y.; Cruz, C.; Miller, J. H. *Nucleic Acids Res.* **1990**, *18*, 3841−3845.
(193) Gogos, A.; Cillo, J.; Clarke, N. D.; Lu, A.-L. *Biochemistry* **1996**, *35*, 16665−16671.
(194) Manuel, R. C.; Czerwinski, E. W.; Lloyd, R. S. *J. Biol. Chem.* **1996**, *271*, 16218−16226.
(195) McGoldrick, J. P.; Yeh, Y.-C.; Solomon, M.; Essigmann, J. M.; Lu, A.-L. *Mol. Cell Biol.* **1995**, *15*, 989−996.
(196) Slupska, M. M.; Baikalov, C.; Luther, W. M.; Chiang, J.-H.; Wei, Y.-F.; Miller, J. H. *J. Bacteriology* **1996**, *178*, 3885−3892.
(197) Mol, C. D.; Arvai, A. S.; Slupphaug, G.; Kavli, B.; Alseth, I.; Krokan, H. E.; Tainer, J. A. *Cell* **1995**, *80*, 1−20.
(198) Savva, R.; McAuley-Hecht, K.; Brown, T.; Pearl, L. *Nature* **1995**, *373*, 487−493.
(199) Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A. *Nature* **1996**, *384*, 87−92.
(200) Vassylyev, D. G.; Morikawa, K. *Structure* **1996**, *4*, 1381−1385.
(201) Ellenberger, T. *Chem. Biol.* **1995**, *2*, 351−354.
(202) Kavli, B.; Slupphaug, G.; Mol, C. D.; Arvai, A. S.; Peterson, S.; Tainer, J. A.; Krokan, H. E. *EMBO J.* **1996**, *15*, 3442−3447.
(203) Yamagata, Y.; Kato, M.; Odawara, K.; Tokuno, Y.; Nakashima, Y.; Matsushima, N.; Yasumura, K.; Tomita, K.-i.; Ihara, K.; Fujii, Y.; Nakabeppu, Y.; Sekiguchi, M.; Fujii, S. *Cell* **1996**, *86*, 311−319.
(204) Labahn, J.; Schärer, O. D.; Long, A.; Ezaz-Nikpay, K.; Verdine, G. L.; Ellenberger, T. E. *Cell* **1996**, *86*, 321−329.
(205) Thayer, M. M.; Ahern, H.; Xing, D.; Cunningham, R. P.; Tainer, J. A. *EMBO J.* **1995**, *14*, 4108−4120.
(206) Doherty, A. J.; Serpell, L. C.; Ponting, C. P. *Nucleic Acids Res.* **1996**, *24*, 2488−2498.
(207) Kuo, C.-F.; McRee, D. E.; Fisher, C. L.; O'Handley, S. F.; Cunningham, R. P.; Tainer, J. A. *Science* **1992**, *258*, 434−440.
(208) Cunningham, R. P.; Asahara, H.; Bank, J. F.; Scholes, C. P.; Salerno, J. C.; Surerus, K.; Munck, E.; McCracken, J.; Peisach, J.; Emptage, M. H. *Biochemistry* **1989**, *28*, 4450−4455.
(209) Johnson, M. K. *Encylcopedia of Inorganic Chemistry*; Wiley: New York, 1994.
(210) Lovenberg, W. *Iron−Sulfur Proteins*; Academic Press: New York, 1973−1977; Vols. I and II.
(211) Spiro, T. G., Ed. *Iron−Sulfur Proteins*; Wiley: New York, 1982.
(212) Horst, J.-P.; Fritz, H.-J. *EMBO J.* **1996**, *15*, 5459−5469.
(213) Porello, S. L.; Cannon, M. J.; David, S. S. *Biochemistry* **1998**, in press.
(214) O'Connor, T. R.; Graves, R. J.; de Murcia, G.; Castaing, B.; Laval, J. *J. Biol. Chem.* **1993**, *268*, 9063−9070.
(215) Tchou, J.; Michaels, M. L.; Miller, J. H.; Grollman, A. P. *J. Biol. Chem.* **1993**, *268*, 26738−26744.
(216) Vassylyev, D. G.; Kashiwagi, T.; Mikami, Y.; Ariyoshi, M.; Iwai, S.; Ohtsuka, E.; Morikawa, K. *Cell* **1995**, *83*, 773−782.
(217) Morikawa, K.; Matsumoto, O.; Tsujimoto, M.; Katayanagi, K.; Ariyoshi, M.; Doi, T.; Ikehara, M.; Inaoka, T.; Ohtsuka, E. *Science* **1992**, *256*, 523−526.
(218) von Hippel, P. H.; Berg, O. G. *J. Biol. Chem.* **1989**, *264*, 675−678.
(219) Lohman, T. M. *CRC Crit. Rev. Biochem.* **1986**, *19*, 191−245.
(220) Lloyd, R. S.; Hanawalt, P. C.; Dodson, M. L. *Nucleic Acids Res.* **1980**, *8*, 5113−5127.
(221) Dowd, D. R.; Lloyd, R. S. *J. Mol. Biol.* **1989**, *208*, 701−707.
(222) Higley, M.; Lloyd, R. S. *Mutat. Res. DNA Repair* **1993**, *294*, 109−116.
(223) Dowd, D. R.; Lloyd, R. S. *J. Biol. Chem.***1990**, *265*, 3424−3431.
(224) Purmal, A. A.; Lampman, G. W.; Pourmal, E. I.; Melamede, R. J.; Wallace, S. S.; Kow, Y. W. *J. Biol. Chem.* **1994**, *269*, 22046−22053.
(225) Kouchakdijian, M.; Bodepudi, V.; Shibutani, S.; Eisenberg, M.; Johnson, F.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1991**, *30*, 1403−1412.
(226) McAuley-Hecht, K. E.; Leonard, G. A.; Gibson, N. J.; Thomson, J. B.; Watson, W. P.; Hunter, W. N.; Brown, T. *Biochemistry* **1994**, *33*, 10266−10270.
(227) Verdine, G. L.; Bruner, S. D. *Chem. Biol.* **1997**, *4*, 329−334.
(228) Bailly, V.; Verly, W. G.; O'Connor, T.; Laval, J. *Biochem. J.* **1989**, *262*, 581−589.
(229) Bhagwat, M.; Gerlt, J. A. *Biochemistry* **1996**, *35*, 659−665.
(230) Latham, K. A.; Lloyd, R. S. *Biochemistry* **1995**, *34*, 8796−8803.
(231) Manoharan, M.; Ransom, S. C.; Mazumder, A.; Gerlt, J. A. *Nucleosides Nucleotides* **1989**, *8*, 879−883.
(232) Kim, J.; Linn, S. *Nucleic Acids Res.* **1988**, *16*, 1135−1141.
(233) Mazumder, A.; Gerlt, J. A.; Absalon, M. J.; Stubbe, J.; Cunningham, R. P.; Withka, J.; Bolton, P. H. *Biochemistry* **1991**, *30*, 1119−1126.
(234) Mazumder, A.; Gerlt, J. A.; Rabow, L.; Absalon, M. J.; Stubbe, J.; Bolton, P. H. *J. Am. Chem. Soc.* **1989**, *111*, 8029−8030.
(235) Sun, B.; Latham, K. A.; Dodson, M. L.; Lloyd, R. S. *J. Biol. Chem.* **1995**, *270*, 19501−19508.
(236) Dodson, M. L.; Michaels, M. L.; Lloyd, R. S. *J. Biol. Chem.***1994**, *269*, 32709−32712.
(237) Weiss, B.; Grossman, L. *Adv. Enzymol.* **1987**, *60*, 1−34.
(238) Dodson, M. L.; Schrock, R. D. I.; Lloyd, R. S. *Biochemistry* **1993**, *32*, 8284−8290.
(239) Schrock, R. D. I.; Lloyd, R. S. *J. Biol. Chem.* **1991**, *266*, 17631−17639.
(240) Schrock, R. D., III; Lloyd, R. S. *J. Biol. Chem.* **1993**, *268*, 880−886.
(241) Tchou, J.; Grollman, A. P. *J. Biol. Chem.* **1995**, *270*, 11671−11677.
(242) Zharkov, D. O.; Rieger, R. A.; Iden, C., R.; Grollman, A. P. *J. Biol. Chem.* **1997**, *272*, 5335−5342.
(243) Kow, Y. W.; Wallace, S. S. *Biochemistry* **1987**, *26*, 8200−8206.
(244) Bailly, V.; Verly, W. G. *Biochem. J.* **1987**, *242*, 562−572.
(245) Nash, H. M.; Lu, R.; Lane, W. S.; Verdine, G. L. *Chem. Biol.* **1997**, *4*, 693−702.
(246) Williams, S. D.; David, S. S. Unpublished results.
(247) (a) Manuel, R. C.; Latham, K. A.; Dodson, M. L.; Lloyd, R. S. *J. Biol. Chem.* **1995**, *270*, 2652−2661. (b) Morikawa, K.; Matsumoto, D.; Tsujimoto, M.; Katayanagi, K.; Ariyoshi, M.; Doi, T.; Ikehara, M.; Inaoka, T.; Ohtsuka, E. *Science* **1992**, *256*, 523−526.
(248) Hori, N.; Doi, T.; Karaki, Y.; Kikuchi, M.; Ikehara, M.; Ohtsuka, E. *Nucleic Acids Res.* **1992**, *20*, 4761−4764.
(249) Latham, K. A.; Manuel, R. C.; Lloyd, R. S. *J. Bacteriol.* **1995**, *177*, 5166−5168.
(250) Schärer, O. D.; Verdine, G. L. *J. Am. Chem. Soc.* **1995**, *117*, 10781−10782.
(251) Schärer, O. D.; Kawate, T.; Gallinari, P.; Jiricny, J.; Verdine, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4878−4883.
(252) Iwai, S.; Maeda, M.; Shirai, M.; Shimada, Y.; Osafune, T.; Murata, T.; Ohtsuka, E. *Biochemistry* **1995**, *34*, 4601−4609.
(253) Kubareva, E. A.; Volkov, E. M.; Vinogradova, N. L.; Kanevsky, I. A.; Oretskaya, T. S.; Kuznetsova, S. A.; Brevnov, M. G.; Gromova, E. S.; Nevinsky, G. A.; Shabarova, Z. A. *Gene* **1995**, *157*, 167−171.
(254) Schärer, O. D.; Ortholand, J.-Y.; Ganesan, A.; Ezaz-Nikpay, K.; Verdine, G. L. *J. Am. Chem. Soc.* **1995**, *117*, 6623−6624.
(255) Deng, L.; Sharer, O. D.; Verdine, G. L. *J. Am. Chem. Soc.* **1997**, *119*, 7865−7866.
(256) Horenstein, B. A.; Schramm, V. L. *Biochemistry* **1993**, *32*, 7089−7097.
(257) Horenstein, B. A.; Schramm, V. L. *Biochemistry* **1993**, *32*, 9917−9925.
(258) Horenstein, B. A.; Zabinski, R. F.; Schramm, V. L. *Tetrahedron Lett.* **1993**, *34*, 7213−7216.
(259) McCullough, A. K.; Schärer, O.; Verdine, G. L.; Lloyd, R. S. *J. Biol. Chem.* **1996**, *271*, 32147−32152.
(260) O'Handley, S.; Scholes, C. P.; Cunningham, R. P. *Biochemistry* **1995**, *34*, 2528−2536.
(261) Castaing, B.; Boiteux, S.; Zelwer, C. *Nucleic Acids Res.* **1992**, *20*, 389−394.
(262) Barzilay, G.; Hickson, I. D. *BioEssays* **1995**, *17*, 713−719.
(263) Demple, B.; Harrison, L. *Annu. Rev. Biochem.* **1994**, *63*, 915−48.
(264) Price, A.; Lindahl, T. *Biochemistry* **1991**, *30*, 8631−8637.

(265) Robson, C. N.; Hickson, I. D. *Nucleic Acids Res.* **1991**, *19*, 5519−5523.
(266) Royer-Pokora, B.; Gordon, L. K.; Haseltine, W. A. *Nucleic Acids Res.* **1981**, *9*, 4595−4609.
(267) Chen, D. S.; Herman, T.; Demple, B. *Nucleic Acids Res.* **1991**, *19*, 5907−5914.
(268) Wood, R. D. *Annu. Rev. Biochem.* **1996**, *65*, 67−135.
(269) (a) Barzilay, G.; Walker, L. J.; Robson, C. N.; Hickson, I. D. *Nucleic Acids Res.* **1995**, *23*, 1544−1550. (b) Demple, B. *Gene* **1996**, *179*, 53−57.
(270) Doestch, P. W.; Cunningham, R. P. *Mutation Res.* **1990**, *236*, 173−201.
(271) Agnez, L. F.; Costa de Oliveira, R.; Di Mascio, P.; Menck, C. F. M. *Carcinogenesis* **1996**, *17*, 1183−1185.
(272) Haring, M.; Rudiger, H.; Demple, B.; Boiteux, S.; Epe, B. *Nucleic Acids Res.* **1994**, *22*, 2010−2015.
(273) Takeuchi, M.; Lillis, R.; Demple, B.; Takeshita, M. *J. Biol. Chem.* **1994**, *269*, 21907−21914.
(274) Hang, B.; Chenna, A.; Fraenkel-Conrat, H.; Singer, B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13737−13741.
(275) Shida, T.; Noda, M.; Sekiguchi, J. *Nucleic Acids Res.* **1996**, *24*, 4572−4576.
(276) Suh, D.; Wilson III, D. M.; Povrik, L. F. *Nucleic Acids Res.* **1997**, *25*, 2495−2500.
(277) Chaudhry, M. A.; Weinfeld, M. *J. Biol. Chem.* **1997**, *272*, 15650−15655.
(278) Rothwell, D. G.; Hickson, I. D. In *Base Excision Repair of DNA Damage*; Hickson, I. D., Ed.; Landes Bioscience: Austin, TX, 1997; pp 67−80.
(279) Levin, J. D.; Shapiro, R.; Demple, B. *J. Biol. Chem.* **1991**, *266*, 22893−22898.
(280) Mol, C. D.; Kuo, C.-F.; Thayer, M. M.; Cunningham, R. P.; Tainer, J. A. *Nature* **1995**, *374*, 381−386.
(281) Barzilay, G.; Mol, C. D.; Robson, C. N.; Walker, L. J.; Cunningham, R. P.; Tainer, J. A.; Hickson, I. D. *Nature Struct. Biol.* **1995**, *2*, 561−568.
(282) Gorman, M. A.; Morera, S.; Rothwell, D.; de La Forteller, E.; Mol, C. D.; Tainer, J. A.; Hickson, I. D.; Freemont, P. S. *EMBO J.* **1997**, *16*, 6548−6558.
(283) Black, C. H.; Cowan, J. A. *Eur. J. Biochem.* **1997**, *243*, 684−689.
(284) Rothwell, D. G.; Hickson, I. D. *Nucleic Acids Res.* **1996**, *24*, 4217−4221.
(285) Suck, D.; Oefner, C. *Nature* **1986**, *321*, 620−625.
(286) Fersht, A. *Enzyme Structure and Mechanism*; W. H. Freeman: New York, 1985.
(287) Lahm, A.; Suck, D. *J. Mol. Biol.* **1991**, *221*, 645−667.
(288) Weston, S. A.; Lahm, A.; Suck, D. *J. Mol. Biol.* **1991**, *226*, 1237−1256.
(289) Wilson, D. M.; Takeshita, M.; Demple, B. *Nucleic Acids Res.* **1997**, *25*, 933−939.
(290) Wilson, D. M., III; Thompson, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12754−12757.
(291) Engelward, B. P.; Weeda, G.; Wyatt, M. D.; Broekhof, J. L.; De Wit, J.; Donker, I.; Allan, J. M.; Gold, B.; Hoeijimakers, J. H. J.; Samson, L. D. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13087−13092.

CR980321H