A Mechanistic Perspective on the Chemistry of DNA Repair Glycosylases

James T. Stivers* and Yu Lin Jiang

Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205

Received February 11, 2003

Contents

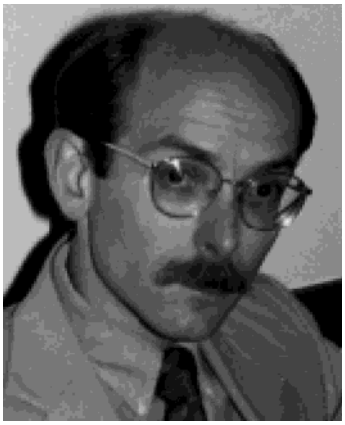
* Corresponding author. Phone: 410-502-2758. Fax: 410-955-3023. E-mail: jstivers@jhmi.edu.

I. Introduction

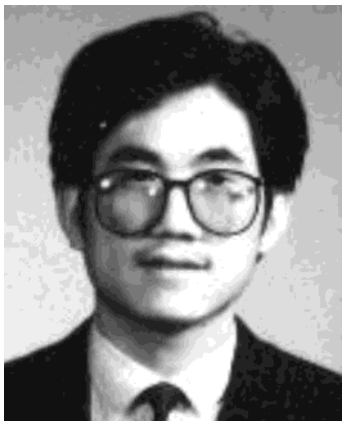
A. Overview of DNA Base Excision Repair

DNA was selected over the course of evolution as the ideal molecule for coding and storing the genetic blueprint of life. One primary driving force for selecting DNA as the repository for this critical information is the intrinsic stability of its chemical bonds, which are remarkably resistant to chemical attack by solvent and exogenous chemical agents. Despite the high stability of DNA in chemical terms, its stability with respect to ensuring the preservation of its coding content is far less than infinite. This sensitivity arises from the fact that damage to a

10.1021/cr010219b CCC: $44.00
Published on Web 05/28/2003



James T. Stivers was born in Philadelphia, PA, and raised in Anchorage, AK. After pursuing a music degree from Berklee College of Music in Boston and serving as a staff composer/arranger for the Air Force Band for several years, he received a B.S. in Microbiology and Immunology from the University of Washington in 1987. He obtained his Ph.D. in Biochemistry from Johns Hopkins University in 1992, where he became interested in enzyme catalysis. He received postdoctoral training in heteronuclear NMR and enzymology in the laboratory of Professor Albert Mildvan at Johns Hopkins Medical School, where he was an American Cancer Society postdoctoral fellow from 1993 to 1996. He then joined the faculty of the Center for Advanced Research in Biotechnology in Rockville, MD, as an Assistant Professor in 1996. During the year 2001, he moved his laboratory to the Department of Pharmacology and Molecular Sciences at Johns Hopkins Medical School, where he is now an Associate Professor. Presently, his research focuses on understanding the nature of enzyme catalysis and inhibition for a number of enzymes involved in DNA repair and recombination. When time allows, he still enjoys playing both the classical and electric guitars.



Yu Lin Jiang grew up in Zhejiang province, China. He received his B.S. and M.S. degrees at Jilin University, where he studied physical and organic chemistry. In 1989, he was made an assistant professor of Dalian University of Technology, where he taught advanced organic chemistry. He then received an Overseas Postgraduate Research Scholarship at Adelaide University from 1996 to 1999, where he studied the molecular recognition of nucleic acids bases and biotin derivatives under the supervision of Geoffrey T. Crisp. He has worked with Dr. Stivers since 1999, where he has applied novel chemical approaches to understand enzymatic DNA repair. His current interests are in the development of new small molecules to probe biological recognition and catalysis.

single codon in a genome of three billion base pairs can result in catastrophic consequences to the organism. Thus, there has been an equally strong driving force for the evolution of enzymatic machinery that efficiently recognizes and repairs damage to the information content of DNA. By necessity, these enzymes must change the covalent structure of DNA to initiate the repair process and must therefore possess considerable catalytic power to cleave the bonds that were in themselves selected for stability. It is the focus of this review to examine the chemical basis for catalysis of a subset of these important enzymes, the DNA glycosylases.

DNA glycosylases are the opening act in a highly conserved process for excision of damaged bases in DNA called the base excision repair pathway (BER, Figure 1).[1,2] This pathway is most simply viewed as consisting of two steps: damaged base excision by a DNA glycosylase, followed by base replacement that is catalyzed by the consecutive action of at least three enzymes. As detailed in Figure 1, DNA glycosylases are faced with the task of cleaving the carbon−nitrogen bond that connects the anomeric carbon of the sugar to the nucleobase, known as the glycosidic bond, thereby producing two products: an abasic site in the DNA and the free damaged base. As will be discussed below, glycosylase enzymes may be broadly classified as monofunctional or bifunctional.

1. Repair Initiated by Monofunctional DNA Glycosylases

Monofunctional DNA glycosylases are hydrolase enzymes that simply use a water molecule to attack the anomeric carbon of the damaged nucleotide (Figure 1). The resulting abasic product is one of the most cytotoxic lesions that can be found in DNA[3−5] and must be efficiently excised in the base replacement segment of BER. Removal of the abasic nucleotide begins through the action of an abasic site endonuclease (AP endo) that hydrolytically cleaves the 5′ phosphodiester linkage of the abasic site in a Mg^{2+}-dependent reaction,[6,7] resulting in a 3′ hydroxyl and a 5′ phosphate abasic nucleotide. The 3′ hydroxyl of the neighboring nucleotide provides the required nucleophile for template directed addition of the correct dNTP through the action of a small, bifunctional DNA repair polymerase (pol β).[8,9] This remarkable polymerase possesses an additional 9 kDa lyase domain which directs attack of a conserved lysine residue at the aldehydic carbon of the ring open form of the abasic nucleotide. The resulting Schiff base electron sink can then facilitate cleavage of the phosphodiester linkage on the 3′ side of the abasic nucleotide via a β elimination mechanism (Figure 1). This elegant reaction produces a nick in the DNA which is sealed by DNA ligase I or III in an ATP-dependent reaction that connects the 3′ hydroxyl of the newly inserted nucleotide with the 5′ phosphate of its neighbor, completing the repair process.

2. Repair Initiated by Bifunctional DNA Glycosylases

Bifunctional glycosylases differ from their monofunctional counterparts by their use of an active site amine nucleophile rather than water.[10,11] In addition to expelling the damaged base in the glycosidic bond cleavage reaction, this amine nucleophile also serves as a Schiff base electron sink to facilitate a subsequent β elimination reaction in which the 3′ phosphate of the abasic nucleotide is expelled. Thus, this secondary activity of the bifunctional glycosylases efficiently replaces the lyase function of pol β (Figure 1). Accordingly, bifunctional glycosylases only require DNA polymerases such as pol δ or ϵ, which do not have lyase domains, to complete the repair process (Figure 1).

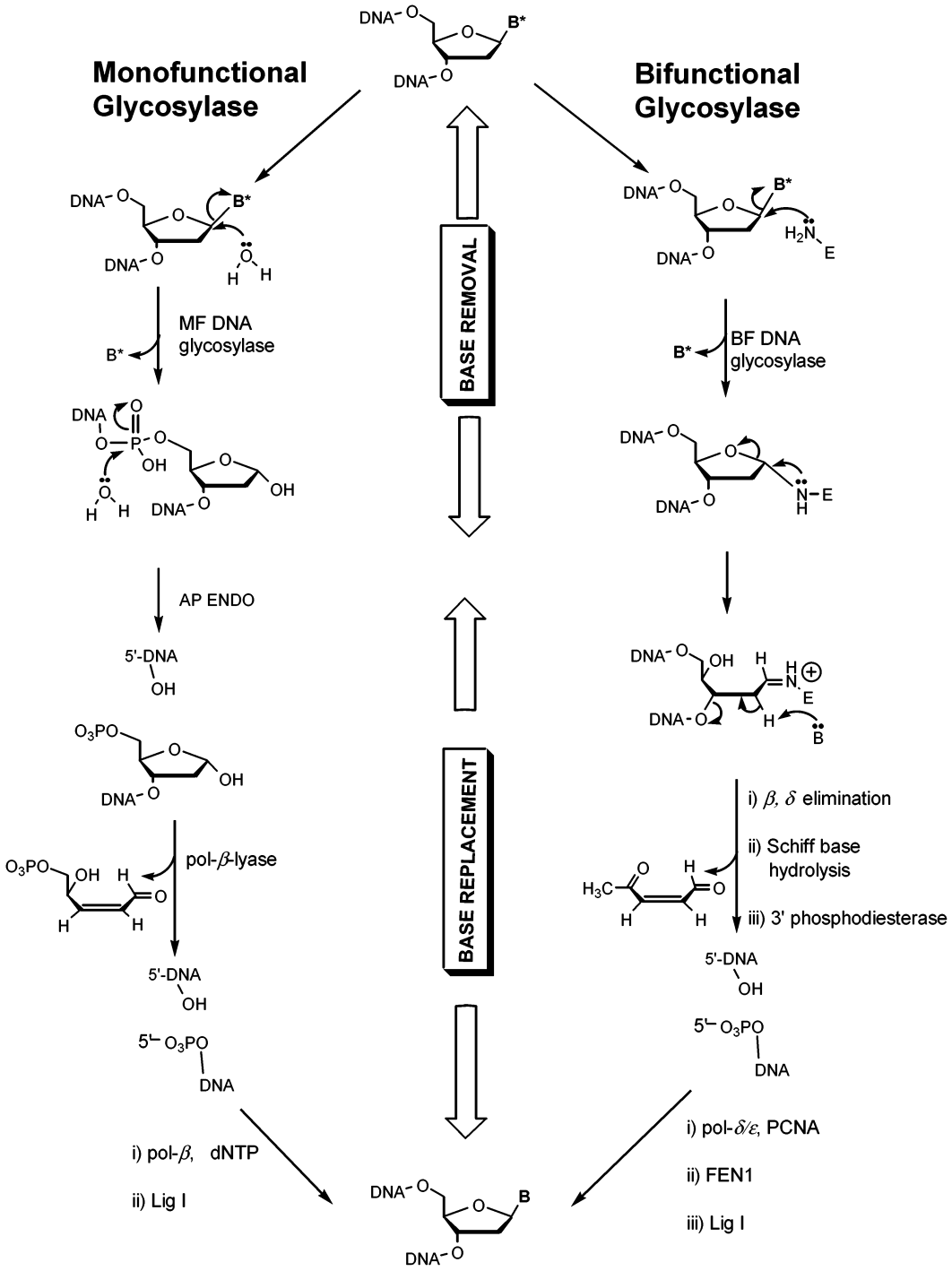
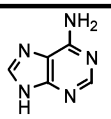
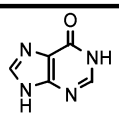
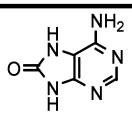
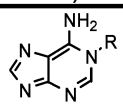
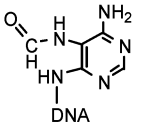
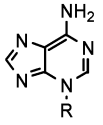
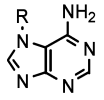
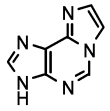
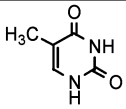
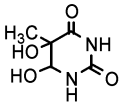
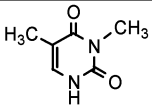
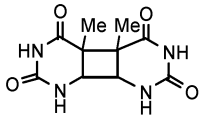
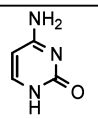
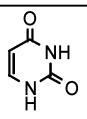
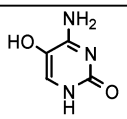
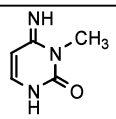
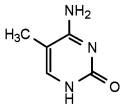
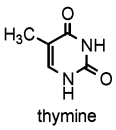
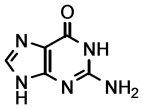
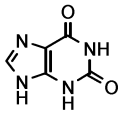
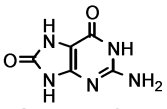
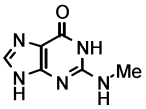
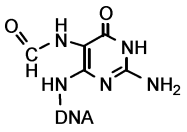
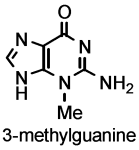
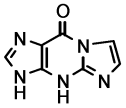


Figure 1. Two pathways for base excision repair of damaged bases in human DNA. These pathways are both comprised of an initial base removal step that involves a monofunctional (MF) or bifunctional (BF) glycosylase. Short patch repair involves the excision and replacement of the damaged nucleotide only, while long patch repair involves the removal of an additional 3 or four nucleotides 3′ of the damage site (see text). Although short patch and long patch repair are depicted for MF and BF glycosylases, respectively, these alternative replacement pathways can be operative for either type of glycosylase product. Abbreviations: AB endo, abasic site endonuclease; pol β, DNA polymerase β; pol δ, DNA polymerase δ; pol ϵ, DNA polymerase ϵ; FEN 1, FLAP endonuclease; Lig, DNA ligase.

Although the above-described single nucleotide replacement pathway is highly prevalent, depending on the type of damage lesion and whether the lesion is repaired by a monofunctional or bifunctional glycosylase, other repair polymerases (i.e., pol δ or pol ϵ), or endonucleases (i.e., FLAP endonuclease) may become involved (Figure 1), leading to the excision and filling in of three additional nucleotides on the 3′ side of the original damaged site ("long patch repair").[12,13] Long patch repair may also be necessary when the abasic site has been damaged so that it is no longer a substrate for the AP lyase reaction catalyzed by pol β. The enzymes involved in short patch and long patch BER in mammals are indicated in Figure 1, along with their substrates and products. The reader is referred to other reviews for a discussion of other accessory proteins that are involved in short and long patch repair.[14]

Table 1. Common Types of Damaged Bases Found in DNA

Normal Base	Examples of Common Base Modifications			
	Deamination	Oxidation	Alkylation (R = Me)	UV
 adenine	 hypoxanthine	 8-oxoadenine	 1-methyladenine	
		 4,6-diamino-5-formamidopyrimidine	 3-methyladenine	
			 7-methyladenine	
			 1,N^6-ethenoadenine	
 thymine		 thymine glycol	 3-methylthymine	 cyclobutane thymine dimer
 cytosine	 uracil	 5-hydroxycytosine	 3-methylcytosine	
 5-Me-cytosine	 thymine			
 guanine	 xanthine	 8-oxoguanine	 N^2-methylguanine	
		 2,6-diamino-4-hydroxy 5-formamidopyrimidine	 3-methylguanine	
			 1,N^2-ethenoguanine	

3. The Range of Chemical Damage to DNA Bases

The types of damage that may be incurred to DNA bases is varied and in many cases unavoidable in the aqueous milieu of the cell (Table 1). The general categories of base damage range from hydrolytic deamination of the exocyclic amino groups of cytosine, guanine, and adenine, alkylation of the elec-

Table 2. Representative DNA Repair Glycosylases

general class of damage	enzymes	substrate	organisms[a]	monofunctional (M) or bifunctional (B)
mismatched purine	MutY	**A**:8oxoG	b, h	M
		A:G		
oxidized purine	FPG	**8oxoG**:C	b	B
	hOGG1	**8oxoG**:C	h	B
pyrimidine	UDG	**U**:A	b, h	M
		U:G		
		ss**U**		
	MUG	**U**:G, **eC**:G		M
	TDG	**U**:G, **eC**:G		M
cyclobutane dimer	PDG	**T**∧T	T4 phage	B
alkylation	AlkA	3-meA, other purines	b	M
	AAG	3-meA, other purines	h	M
	TAG	3-MeA, 3-MeG	b	M

[a] b, bacteria; h, human.

tron rich heteroatoms of purine and pyrimidine bases by endogenous or exogenous electrophiles, and oxidative reactions of both purines and pyrimidines by reactive oxygen species.[1] In addition, absorption of short wavelength UV light by the 5,6 double bond of two thymine residues adjacent to one another on the same strand of duplex DNA may lead to spontaneous formation of pyrimidine dimers in which the electrons within the 5,6 double bonds have rearranged to form a cyclobutane ring (Table 1).[15] As will be emphasized in this review, each type of base damage elicits a change in the covalent structure and electronic properties of the base that has a major impact on both the recognition and catalytic mechanisms for enzymatic repair. Indeed, the stability of the glycosidic bonds for the types of base damage shown in Table 1 vary by over 6 orders of magnitude (see below), which indicates that individual glycosylases encounter catalytic problems of immensely different scale.

B. Scope of Review

This review will emphasize developments since the last review of DNA glycosylases in this journal in 1998.[16] Rather than attempting to serve as a comprehensive encyclopedia of these enzymatic reactions, the approach here will be problem based. That is, the focus will be on the general biophysical and chemical problems that each of these enzymes must overcome, with a decided emphasis on studies that most clearly illuminate the particular mechanistic aspect under consideration. We will begin by examining the chemical problem of glycosidic bond cleavage in nucleosides with the aim of uncovering the intrinsic energetic barriers that glycosylase enzymes must overcome. Then the enigma of the damage site search and recognition mechanism will be considered, where once again, basic physical principles and the dynamic nature of the DNA substrate will be used to examine the viability of possible mechanistic alternatives. Finally, using representative enzyme examples, the two key steps in the overall enzymatic process of damaged base removal will be considered: the extrusion of the damaged base from the DNA duplex into the enzyme active site ("base flipping"), and the catalytic step of glycosidic bond breakage. Our intent is to corral observations from different studies and hopefully reveal commonalities in mechanism that are revealing of function. The reader is referred to numerous previous reviews on DNA repair glycosylases that have covered other important aspects of these enzymes ranging from structure,[14,17] biology,[18] and evolutionary relationships.[19,20] The representative DNA glycosylases that will be considered in this review are listed in Table 2. These enzymes were chosen because they catalyze reactions with substrates that cover a wide range of structure and because structural and mechanistic studies have progressed most rapidly in these systems. We apologize to our colleagues for the omission of excellent work that fell outside the scope of this review.

II. Glycosidic Bond Hydrolysis: The Intrinsic Chemical Problem

A. Estimating the Power of DNA Glycosylases

The native purine and pyrimidine nucleotides in DNA have distinct chemical properties that dictate the types of damage that each may incur (Table 1), but also limit the possible mechanisms for catalyzing glycosidic bond cleavage in solution and in the active sites of enzymes. The mechanisms for glycosidic bond cleavage of normal 2′ deoxyribonucleosides and their 2′ hydroxyl counterparts have been studied extensively by physical organic chemists over the last 50 years, and such studies provide a foundation for understanding the chemical barriers that enzymes must overcome to accelerate these reactions.[21]

It is instructive to estimate the range and magnitude of the catalytic problems that are faced by DNA glycosylases using the approach of Wolfenden [22] (Figure 2). In this analysis, the rates for the nonenzymatic hydrolysis reactions (k_{non}) are compared to the maximal rates for the corresponding enzymatic reactions (k_{enz}), and the ratio k_{enz}/k_{non} is a measure of the catalytic power of the enzyme. In free energy terms, $\Delta\Delta G^{act} = -RT \ln (k_{enz}/k_{non})$ is the difference free energy between the activation barriers for the enzymatic and nonenzymatic reactions. In other words, $\Delta\Delta G^{act}$ provides the magnitude of the activation barrier lowering by the enzyme. The lengths of the vertical lines in Figure 2 are proportional to this difference free energy. At neutral pH and 25 °C, normal purine and pyrimidine deoxynucleosides are highly resistant to hydrolytic cleavage of the glyco-

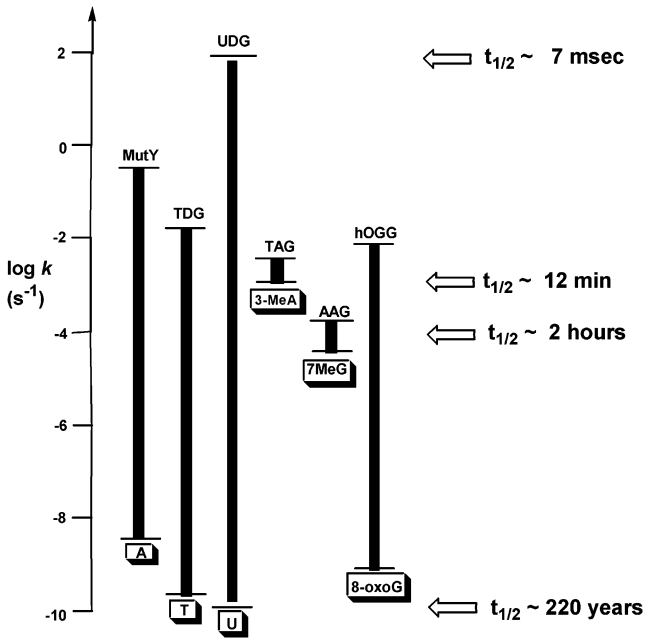


Figure 2. Catalytic powers of representative DNA glycosylases at neutral pH and ambient temperatures. The rate enhancements of DNA glycosylases may be estimated from the spontaneous rates of hydrolysis of the free deoxynucleosides and the corresponding enzymatic rate. The spontaneous hydrolysis rate constants shown here were calculated from the values reported in refs 23−26 and, if necessary, were corrected to 25° (deoxyuridine) or 37° (all others) using the Arrhenius equation and the measured activation enthalpies. The enzymatic rate constants were obtained from refs 27−31 and, with the possible exception of TAG, are single-turnover rate constants.

sidic bond with half-lives for hydrolysis of about 220 years,[23−26] and the enzymes that accelerate these reactions have catalytic powers in the range 10^7−10^{12}-fold.[27−31] In contrast, the spontaneous hydrolyses of alkylated purine nucleosides occurs with half-lives as short as 30 min.[32] The catalytic power of the enzymes that remove positively charged alkylated purine bases is 10^7−10^{11}-fold less than the enzymes that recognize neutral bases (Figure 2). It is a common observation that the catalytic power of enzymes correlates with the slowness of the spontaneous reactions, rather than the rapidity of the enzymatic reactions, which tend to fall in a rather narrow range of rates.[22] Thus, it appears that DNA repair glycosylases have evolved to remove bases as rapidly as is required by the evolutionary selection mechanism, which is likely determined by two factors: the probability of incurring a given type of chemical damage, and the average time between damage occurrence and DNA replication.

B. Nonenzymatic Glycosidic Bond Hydrolysis in Deoxynucleosides

1. The Role of the Leaving Group Base

The mechanisms and reactivities of purine and pyrimidine nucleosides to glycosidic bond hydrolysis are very dependent on the relative acidities of various functional groups on the leaving group bases (Figure 3A). This dependence reflects the profound activating effect of base protonation, which makes the leaving

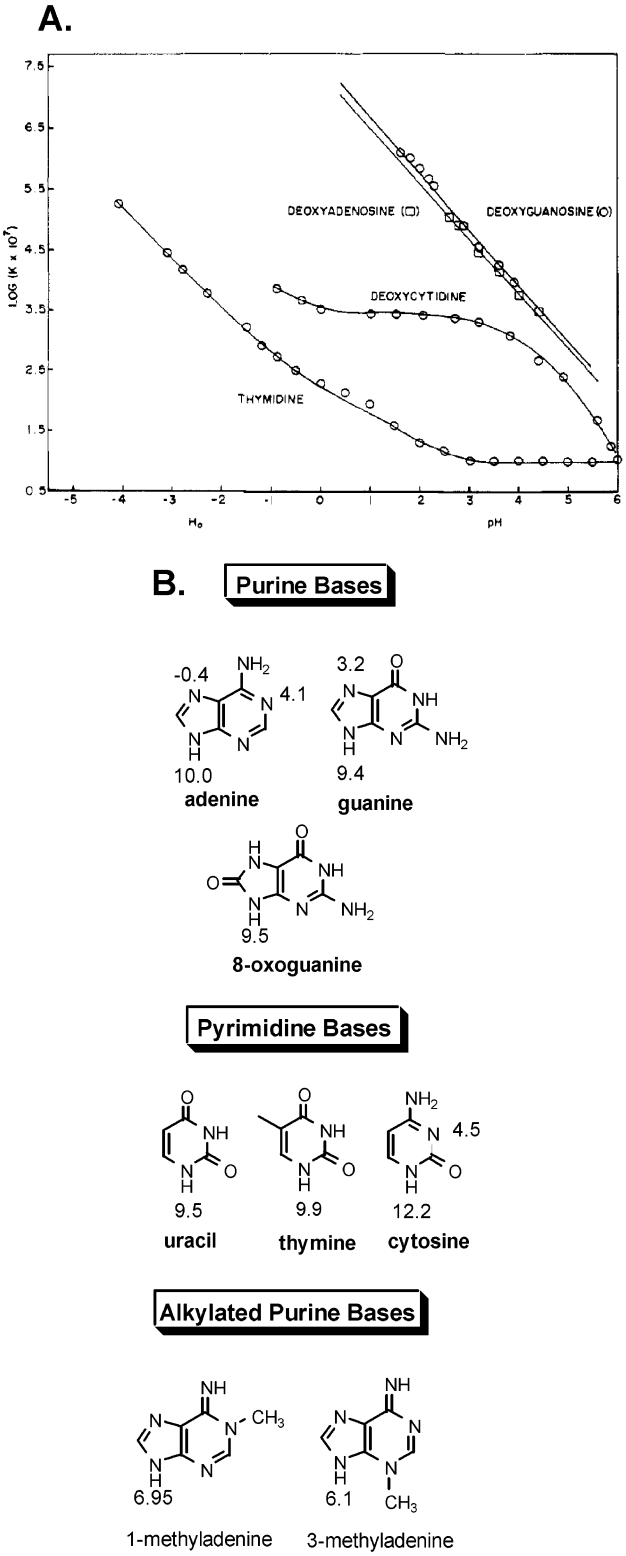


Figure 3. Catalysis of nonezymatic glycosidic bond cleavage and the acidities of pyrimidine and purine leaving groups. (A) pH dependence of the nonenzymatic glycosidic bond hydrolysis reactions of the four naturally occurring DNA nucleosides in water (reprinted from ref 24, copyright 1969 American Chemical Society. (B) pK_a values of proton acceptor sites on purine and pyrimidine bases. pK_a values are from refs 193,194, 195, and 34.

group base more electron deficient, thereby accommodating the increased electron density that develops during glycosidic bond cleavage. For instance, hy-

drolysis of adenosine or guanosine shows an apparent first-order dependence on proton concentration over the entire pH range −3 to 9 reflecting the requirement for rapid equilibrium protonation of the endocyclic nitrogens of these purine bases before bond cleavage (Figure 3B). It is important to note that there is a striking absence of a pH independent hydrolysis reaction for purines (Figure 3B), indicating that the pathway involving expulsion of the protonated base predominates over the accessible pH range.[23] In contrast, the pyrimidine nucleosides deoxyuridine and thymidine both show pH independent hydrolysis reactions over the pH range ∼2 to 9 (Figure 3B).

Why is a pH-independent hydrolysis pathway observed for pyrimidine nucleosides but not purines, even though the pK_a values for the leaving glycosidic nitrogens are both about 9.5 (Figure 3A)? This result arises because the pK_a values for the exocyclic oxygen atoms of pyrimidine bases (pK_a (O2) = −2.98) [33] are about 7 log units lower than those of the endocyclic nitrogen atoms of purines (pK_a ∼ 3−4) (Figure 3A); thus the concentration of the protonated pyrimidine leaving group is insignificant at pH values greater than about 3, resulting in a change to a pH independent pathway involving expulsion of the pyrimidine anion. This change in pathway never occurs for purine bases because the protonated base is always present at a significant concentration at all accessible pH values. These considerations largely account for the observed mechanistic differences between the nonenzymatic glycosidic bond cleavage reactions of purines and pyrimidines. As will be pointed out later, these conclusions have direct bearing on the observed mechanisms for the corresponding enzymatic reactions.

The final general class of nucleobase leaving group is the cationic alkylated purine base. Nucleosides containing such bases are quite reactive, but their reactivity requires invocation of no special mechanisms. For instance, 7-Me-dG is 26 000 times more reactive than dG at neutral pH,[25] reflecting the powerful activating effect of alkylation, which simply mimics the effect of protonation, thereby lowering the pK_a of the N9 leaving group by 4.1 log units. Thus, at neutral pH values, alkylated purines react with the vigor of their protonated purine analogues. Similar activating effects of alkylation are observed for adenosine, once again reflecting the improved quality of the nitrogen leaving group that mimics the effect of protonation.[32,34] The preactivation of these purine leaving groups by alkylation suggests that enzymes that act on these unnatural cationic bases may not require mechanisms for base activation that involve strong hydrogen bonding or proton transfer. Once again, this inference based on chemical intuition appears to be substantiated by the structures and mechanisms of some enzymes that recognize alkylated purines (vide infra).

2. The Role of the Deoxynucleoside Sugar

Most mechanisms for involvement of the furanose sugar in nonenzymatic glycosidic bond cleavage invoke conformational changes in the sugar ring that promote orbital interactions that facilitate bond cleavage.[35−37] These hotly debated mechanisms involve stereoelectronic effects in the ground state or transition state of these reactions.[38] Sugar ring conformations that promote overlap of a lone pair of electrons on the furanose oxygen with the vacant σ^* orbital on C1′ produce an electronic stabilization called the "anomeric effect" (Figure 4A). The energetic magnitude of the ground state anomeric effect, which falls in the modest range of 1−2 kcal mol for deoxynucleosides, depends on the square of the overlap between the donor and acceptor orbitals and is inversely dependent on the energy difference between these orbitals. It is important to realize that the ground-state anomeric effect stabilizes the ground state (i.e., favors conformations that allow the orbital overlap). Thus, the stabilizing effect must increase in magnitude as the transition state is approached, or it will serve to increase the activation barrier for glycosidic bond cleavage.[39]

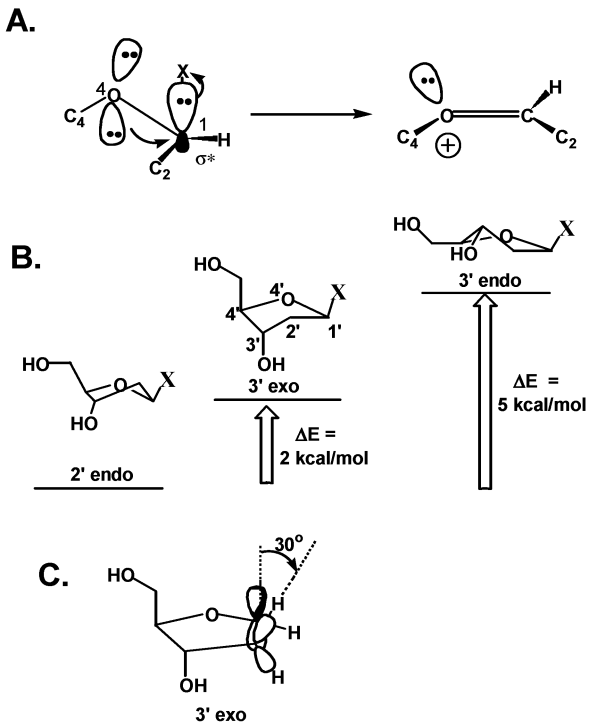
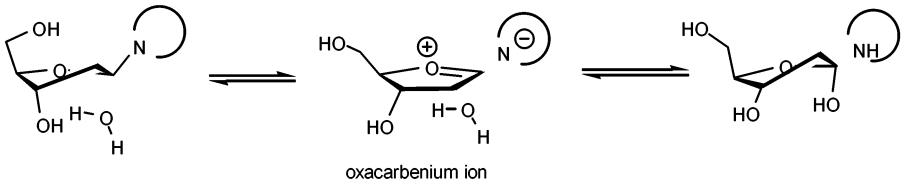


Figure 4. Stereoelectronic control and conformational aspects of glycosidic bond hydrolysis reactions of furanose sugars. (A) The anomeric effect and the antiperiplanar lone pair hypthesis (ALPH) arise from the interaction of a lone pair of electrons on the sugar oxygen with the adjacent empty antibonding σ^* orbital of the C1′ carbon. This arrangement increases the O4′−C1′ bond order and decreases the bond order to the leaving group nitrogen. Sugar conformations that allow the ground-state anomeric effect are stabilized relative to conformations that do not, but this effect is quite small (1−2 kcal/mol). (B) Potential energy differences between the most stable 2′ endo conformation of a deoxynucleoside and the two high-energy conformations that allow maximal hyperconjugation from the σ-bonded electrons of the 2′ CH bonds to C1′ in a dissociative transition state.[42] (C) Optimal orientation of the 2′ CH bonds relative to the vacant p orbital on C1′ for stabilization of a oxacarbenium ion transition state. Such an orientation is achieved in either the 3′ exo (shown) or 3′ endo sugar pucker.

Scheme 1



The kinetic anomeric effect is often termed the "antiperiplanar lone pair hypothesis" (ALPH).[36] ALPH is a stereoelectronic argument that requires orientation of the lone pair electrons on oxygen antiperiplanar to the leaving group, which is proposed to facilitate glycosidic bond cleavage by promoting favorable electron flow from the sugar onto the leaving group (Figure 4A). Sinott has pointed out that the kinetic benefit of the anomeric effect is negligible for reactions with late transition states (i.e., where the bond to the leaving group is nearly completely broken). The small benefit arises because the late transition state has lost the "memory" of whether it arose from the stereochemically correct ground state conformation with the antiperiplanar arrangement, or another conformation.[36] Therefore, the kinetic anomeric effect is expected to be small in the hydrolysis of the glycosidic bond in nucleosides because leaving group expulsion is considerably more advanced than bond formation to the nucleophile for these reactions, resulting in transition states with high glycosyl cation character, or even oxacarbenium ion intermediates that have a borderline existence in solution ($\leq$10 ps).[40,41] In addition, the kinetic anomeric effect is expected to be small for systems that have high conformational flexibility such as naturally occurring glycosides because there is only a small energetic barrier to achieving stereochemically favored conformations. Indeed, the high flexibility of the furanose ring of deoxynucleosides allows all sugar conformations to be accessed at ambient temperatures with barriers in the range 2−5 kcal/mol (Figure 4B),[42] suggesting that the catalytic advantage an enzyme could gain by preorganizing the sugar in a conformation that allows overlap would fall in this energetic range. Of course, the energetic penalty for forming this unstable conformation would be paid for by the utilization of binding energy between the substrate and enzyme,[43] and a key question is at what stage the molecular strain energy outweighs the benefits of antiperiplanar orbital overlap. Examples of ground-state preorganization are prevalent in the now extensive structural database of glycosidase-substrate complexes and will be discussed in the context of the enzyme mechanisms below.[44,45] The modest kinetic anomeric effect should not be confused with the large energetic requirement for orbital overlap in a dissociative transition state that results from the large stability difference between a resonance-stabilized oxacarbenium ion and a naked carbonium ion (Figure 4A, $\Delta\Delta G \sim 20$ kcal/mol).[46,47]

A second type of stereoelectronic effect that may be important is the hyperconjugative effect resulting from electron donation from the σ-bonded electrons at the 2′ position to the p orbital of the nascent oxacarbenium ion in the transition state (Figure 4C).[47,48] Hyperconjugative effects are most effective when orbital overlap between both 2′ C−H σ bonds and the electron deficient p orbital at C1′ are maximal.[49] For deoxyribonucleosides, this occurs when the dihedral angles between both C−H bond vectors and the vacant C1′ orbital are about 30 °,[50] which requires that the normal 2′ endo sugar pucker be transformed to a high energy 3′ exo or 3′ endo form. The energetic cost for an enzyme to induce this transformation is around 2−5 kcal/mol as estimated from ab intio QM calculations of the relative energetics of various furanose sugar puckers (Figure 4B).[42]

3. Nucleoside Hydrolysis: Nomenclature and Transition-State Analysis Methods

It is now reasonably well established that nucleophilic substitution reactions at the anomeric carbon of sugars proceed through oxacarbenium ion intermediates or transition states that resemble oxacarbenium ions (Scheme 1).[51,52] To facilitate description of these reaction mechanisms, an IUPAC-approved nomenclature has been adopted, and the two limiting mechanisms for glycosidic bond cleavage using this nomenclature are shown in Figure 5.[53,54] In this

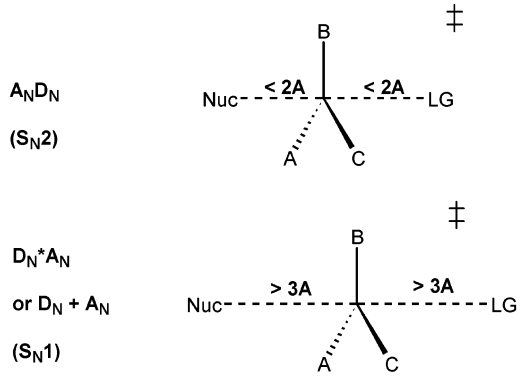


Figure 5. IUPAC approved nomenclature for designation of reaction mechanisms for glycosidic bond hydrolysis. The A_ND_N mechanism involves significant bonding to both the nucleophile and leaving groups. The D_N*A_N or $D_N + A_N$ mechanism has very little or no bonding to these groups (adapted from ref 50).

nomenclature, a bimolecular S_N2 reaction is designated A_ND_N, which represents that there is nucleophile addition (A_N) and nucleophile dissociation (D_N) in the transition state. Such a mechanism may be further described as "associative" if significant bond order to the nucleophile and leaving group exists in the transition state, or as "dissociative" if a small amount of bonding exists. Stepwise S_N1 mechanisms are termed D_N*A_N or $D_N + A_N$, reflecting the fact that leaving group dissociation (D_N) precedes addition of the nucleophile (A_N). The D_N*A_N stepwise mechanism

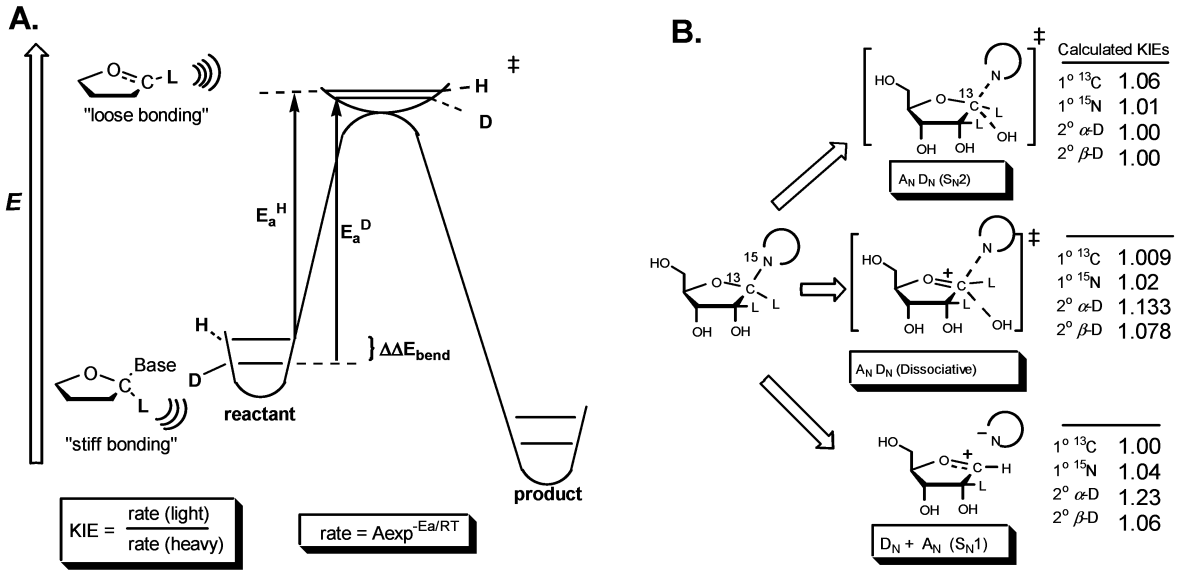


Figure 6. The energetic basis for kinetic isotope effects (KIEs) in glycosidic bond hydrolysis reactions and representative KIEs for A_ND_N and $D_N + A_N$ mechanisms. (A) KIEs are dominated by the changes in the zero point vibrational energies of a light and heavy isotope as a reactant moves from its ground state to transtition state conformation. The depiction shows that the secondary α deuterium KIE arises from looser bonding of the C−L bond in the transition state as compared to the ground state (where L = H or D). Thus, the large ΔE_{bend} for the two isotopes in the ground state is diminished as the transition state is approached. The different activation energies for the two isotopes gives rise to the KIE. (B) Calculated KIEs for three reaction mechanisms. Adapted from ref 55.

differs from the $D_N + A_N$ mechanism with respect to the lifetime of the oxacarbenium ion intermediate. For the former, the intermediate is so short-lived that it does not have time to diffuse away from the leaving group, while in the latter, a fully solvent equilibrated intermediate exists. The D_N*A_N stepwise mechanism can show rate-limiting addition of the nucleophile (D_N*$A_N^{\ddagger}$), or rate-limiting departure of the leaving group ($D_N^{\ddagger}$*A_N).

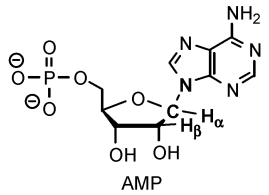
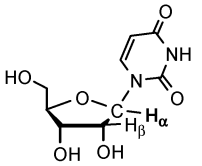
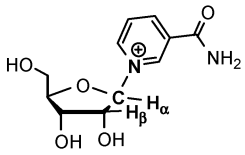
The transition state structures for nonenzymatic hydrolysis of some nucleosides have been ascertained using kinetic isotope effect measurements (KIEs), and in some cases through the use of structure−reactivity measurements in which the pK_a of the leaving group was varied. A detailed description of the KIE approach, which measures the change in reaction rate upon substitution of a heavy isotope for a light isotope in an isotopically sensitive position of a substrate (k_{light}/k_{heavy}), is beyond the scope of this review, and the reader is referred to several recent reviews on this valuable technique.[55,56] Nevertheless, normal KIEs (where $k_{light}/k_{heavy} > 1$) can be qualitatively understood through the simple appreciation that substitution of an atom with a heavier isotope results in stiffening of a bond in the reactant, which places the heavy isotope in a lower vibrational potential energy well than the light isotope (Figure 6A). If bonding at this position becomes "looser" as the transition state is approached (for instance, when the bond is being broken, or if the hybridization changes from sp^3 to sp^2), the energy differences between the light and heavy isotopes that were present in the ground state become smaller, or may even disappear if the bond is broken during the reaction. Thus, the activation energy for the heavy isotope is greater than for the light isotope largely due to the differences in the ground-state vibrational energies, leading to the normal KIE. Using the same analysis, it is possible to understand inverse KIEs (where $k_{light}/k_{heavy} < 1$, and the bond becomes tighter in the transition state). KIEs are called primary when they involve atoms that undergo bond breaking or bond making during the reaction, or secondary when the isotopes do not undergo a change in covalency. For nucleosides, it is possible to measure a family of KIEs that can provide the degrees of bond breaking and bond making in the transition state, as well as detailed geometric features of the transition state. Recently KIE methods have been successfully applied to investigate the transition state structures of DNA glycosylases (vide infra).

4. Transition State Structures for Nonenzymatic Glycosidic Bond Cleavage

The calculated KIEs for $D_N + A_N$ and A_ND_N nonenzymatic nucleoside hydrolysis mechanisms are shown in Figure 6B, and a family of experimentally measured KIEs for the acid-catalyzed hydrolysis of adenosine and uridine and the solvolysis of NAD^+ are listed in Table 3.[57−60] Although these three examples have ribose sugar rings rather than the deoxyribose sugars found in damaged DNA nucleotides and also include a nicotinamide leaving group that is not a DNA base, they represent the closest currently available nonenzymatic models for DNA glycosylase reactions. These limitations point to the need for further directed studies on the transition state structures for hydrolysis of purine and pyrmidine deoxynucleosides. For reference in the examples that follow, qualitative explanations that account for each primary and secondary KIE are provided in Table 3.

What structures do the transition states for nonenzymatic glycosidic bond hydrolysis of purine, alkylated purine, and pyrimidine nucleosides assume?

Table 3. Representative KIEs for Some Nonenzymatic Nucleoside Hydrolysis Reactions[a]

	Primary ^{13}C	Primary ^{15}N	Secondary α D	Secondary β D	Mechanism
 AMP	1.023	1.03	1.175	1.077	A_ND_N (dissociative)
 uridine			1.11		Oxacarbenium ion character
 nicotinamide riboside	1.008	1.02	1.131	1.078	A_ND_N (dissociative)
Interpretation of KIEs	Reaction coordinate motion of C in TS	Bond order to N in the transition structure	sp^2 geometry, out-of-plane vibrations in TS	Extent of 2'H hyperconjugation and electron deficiency at C1'	

[a] The results for AMP, uridine and nicotinamide were obtained from refs,[191,192,60] and.[59]

The picture for the acid catalyzed hydrolysis of adenosine is well-developed and clearly reveals a fairly dissociative A_ND_N mechanism in which departure of the adenine leaving group is farther advanced than nucleophilic addition in the transition state. The solvolysis of NAD^+ has also been extensively studied and shows an even more dissociative A_ND_N mechanism, as evidenced by the larger 2° α-D1′ KIE and the smaller 1° ^{13}C1′ KIE than for adenosine hydrolysis.[59] The slightly more dissociative transition state for NAD^+ solvolysis may be attributed to the enhanced leaving ability of the labile cationic nicotinamide. In contrast to the preceding two cases, the hydrolysis of pyrimidine nucleosides has not been extensively studied using KIE methods. Nevertheless, Santi and Prior have measured a 2° α-D1′ KIE of 1.11 for acid hydrolysis of uridine at 25 °C, which suggests considerable glycosyl cation character for this reaction.[60] Supporting this interpretation are the structure reactivity studies of Shapiro and Danzig in which the hydrolysis reactions of 5-Br-2′-deoxyuridine, 2′-deoxyuridine, and thymidine were followed at pH 6.5 and 75 °C.[24] The slope of a plot of log k_{obsd} against pK_a of the leaving group for these nucleosides was about −1, indicating the development of a full negative charge on the leaving group in the transition state, and a very dissociative transition state, perhaps even a D_N*A_N stepwise mechanism. The low activation entropies in the range +3 to +10 eu for these deoxyuridine nucleosides are also suggestive of low nucleophile participation in the transition state[24] because pure A_ND_N mechanisms would typically have negative activation entropies about 20 eu lower.[61] In addition, the reactions are severely retarded in less polar solvents than water, suggesting significant ionic character in the transition state. These results suggest that the hydrolysis reactions of deoxynucleosides may have more oxacarbenium ion character that the corresponding ribonucleosides, partially due to the destabilizing effect of the electron withdrawing 2′ hydroxyl on the electron deficient anomeric carbon. Accordingly, the model nonenzymatic reactions with ribonucleosides may set a lower limit for the dissociative character of the corresponding reactions of deoxyribonucleosides.

5. Implications of Nonenzymatic Studies for DNA Glycosylases

In general, enzymes have two mechanistic choices: to lower the activation barrier by following the same reaction coordinate and transition state as the nonezymatic reaction, or by enforcing a reaction

coordinate with a different transition state structure. The first alternative requires less catalytic power because the enzyme merely stabilizes the already lowest energy transition state or saddle point that is followed in solution. Of course, ground-state destabilization, or preorganization, may also be used to lower the activation barrier.[62] The second alternative requires more catalytic power from the enzyme because a transition state is enforced that is not already prevalent in the solution reaction, requiring greater stabilization by forces within the enzyme active site. Accordingly, the above nonenzymatic studies suggest that DNA glycosylases may follow dissociative A_ND_N or $D_N{*}A_N$ mechanisms, which would take advantage of the intrinsic tendencies of these reactions to follow dissociative-type mechanisms. We will explore below whether the electronic features of the active sites of DNA glycosylases, and the KIE measurements for several enzymatic reactions, are consistent with these ideas.

III. The Pathway for Finding a Damaged Base in Genomic DNA

A. General

It is difficult to overstate the magnitude of the "search and destroy" mission that each DNA repair glycosylase faces. Consider that in the human genome of over 3 billion base pairs, cytosine deamination occurs at the rate of several hundred events per day per cell,[18,63] which generates an estimated density of uracil residues in DNA of about 1 per 10^7 base pairs over this time period in the absence of repair. A reasonable question to ask is how do repair enzymes locate these rare sites in an enormous background of undamaged DNA? Two general solutions have been put forward to answer this question, and these solutions date back to the seminal work of Berg and von Hippel of over 20 years ago.[64,65] In this analysis, a protein may locate a specific binding site in DNA by a simple three-dimensional random search process or, alternatively, by a translocation mechanism in which it first binds to a nonspecific site on the DNA chain and then linearly diffuses along the chain to its specific site. The latter mechanism, in concert with favorable Coulombic electrostatic forces, can reduce the dimensionality of the search process, leading to bimolecular association rates that exceed the diffusion controlled limit of $\sim 10^8$ M^{-1} s^{-1}.[66] Whether DNA glycosylases use extensive translocation along the DNA chain or three-dimensional search strategies to find their specific sites is not firmly established. Nevertheless, it is useful to examine the physical limitations of these two limiting mechanisms and, in addition, the possible general role for nonspecific DNA interactions in the mechanism of damage site recognition and repair.

B. The Earliest Events in Locating a Damaged Site

1. The Kinetic Competence of a Three-Dimensional Search Mechanism

Is it possible that a three-dimensional search mechanism could be employed to remove 100 unwanted uracil residues from the vastness of human genomic DNA in a reasonable time frame? This question may be addressed in an approximate way by using the analysis shown in Figure 7A. Here we ask the simple question of how long it would take the enzyme uracil DNA glycosylase (E) to excise 100 uracil residues from DNA in the presence of 10^9 fold excess of nonspecific DNA binding sites (DNA^{NS}). The rate and equilibrium constants in this analysis are obtained from the extensive kinetic studies of *Escherichia coli* uracil DNA glycosylase (UDG),[67,68] and it is assumed that the nonspecific DNA binding sites only serve as competitive inhibitors of the enzyme with respect to binding at the specific sites that contain uracil. Thus, in this limiting model, the nonspecific binding events provide no assistance in terms of reducing the dimensionality of the search process. The calculations assume a ratio of specific to nonspecific sites of $100/10^9$ and a nuclear enzyme

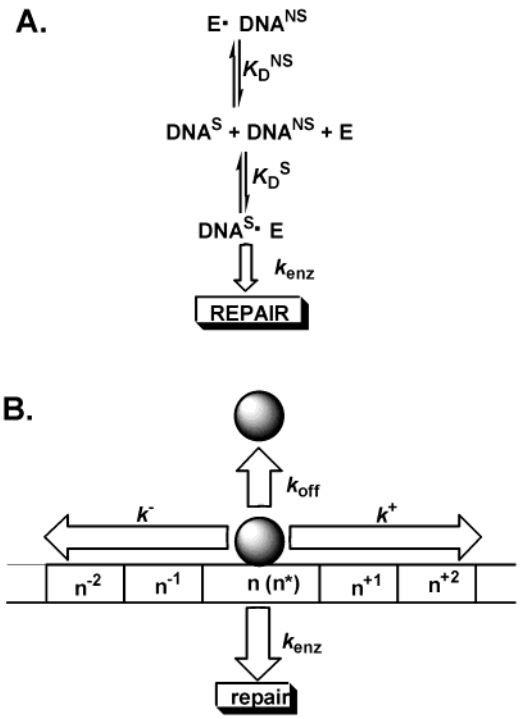


Figure 7. The effects of nonspecific DNA binding and DNA translocation on the efficiency of locating specific sites in DNA. (A) The equilibria describing competitive binding of an enzyme to specific (K_D^S) and nonspecific (K_D^{NS}) sites in DNA. If nonspecific binding does not facilitate location of the specific site by linear diffusion or other mechanisms, location of the specific site will be slowed. For the example of UDG, $K_D^S \sim 10$ nM and $K_D^{NS} \sim 5$ μM (these values are dependent on sequence and DNA length).[146,196] The first irreversible step that leads to repair (k_{enz}) is equal to $\sim$1000 s^{-1} for UDG.[76] (B) The key kinetic partitioning events that determine the effectiveness of a linear diffusion mechanism in finding a specific site in DNA (see text). In this scheme, k_{off} is the dissociation rate of the enzyme from a nonspecific (n) or specific (n*) site in DNA, $k_{trans} = k^- + k^+$ are the translocation rates along the DNA, and k_{enz} is the first irreversible step that leads to repair of the damage site. Using this nomenclature, the probability of translocation down the DNA as opposed to dissociation from a nonspecific site is $P_{trans} = k_{trans}/(k_{off} + k_{trans})$, and the probability of repair as opposed to dissociation or translocation away from a specific site is $P_{repair} = k_{enz}/(k_{off} + k_{trans} + k_{enz})$.

concentration of ~7 μM. In this analysis, the UDG concentration was estimated by the number of UDG molecules per human cell (~200 000) [69] and the nuclear volume (~0.05 pL). Using this mechanism, we find that less than 0.001% of the 100 uracils could be excised in 1 day. This calculation, which is really a worst-case scenario stacked against the kinetic competence of a simple diffusion mechanism, shows that uracils would slowly accumulate in the genome if such a scenario prevailed.

Despite the above limiting case, a diffusive search mechanism can be extremely effective if "hopping" or "jumping" within a DNA domain occurs.[64,65] Such mechanisms, which can be mistakenly assigned as linear diffusion or processivity,[70] are in fact closely related to a random diffusion process because multiple encounter and dissociation events occur before the specific site is located (see below). In fact, diffusive mechanisms offer a major advantage over translocation when the impediments of protein coated cellular DNA are considered.[71,72] Given that there is on average one UDG molecule for each 15 000 bps of human DNA,[69] hopping or jumping within this domain size would effectively reduce the dimensionality of the search process. The probability of locating the target site within this domain would be proportional to the target size (a), and inversely proportional to how far away the target site is located (r) (i.e., $P = a/r$ for the simple case of a spherical target).[64]

2. The One-Dimensional Translocation Mechanism: A Double-Edged Sword

Examining the plausibility of the alternative translocation mechanism requires explicit consideration of the kinetic partitioning of the nonspecific and specific complexes between three possible fates (Figure 7B): DNA dissociation (k_{off}), translocation down the DNA chain ($k_{trans} = k^+ + k^-$), or the first irreversible kinetic step for enzymatic recognition of the specific site (k_{enz}). The rate constants for these processes determine the efficiency of repair [$P_{repair} = k_{enz}/(k_{off} + k_{trans +} k_{enz})$], and the probability of translocation as opposed to dissociation from a nonspecific site [$P_{trans} = k_{trans}/k_{off} + k_{trans}$] as shown in Figure 7B.

It is instructive to use the experimentally measured dissociation rate constant (k_{off}) for UDG to estimate the magnitude of k_{trans} that would be required to provide significant translocation along the DNA chain for this enzyme. For interaction at a nonspecific site (n), UDG binds weakly with a K_D = 2−5 μM and an estimated dissociation rate approaching 1000 s^{-1}.[67] Thus, k_{trans} would have to be much larger than 1000 s^{-1} to achieve efficient translocation as compared to dissociation. Even assuming k_{trans} = 100 000 base pairs per second, only 100 base pairs would be covered in a random walk process before 50% of the enzyme molecules initially bound at the nonspecific site would have dissociated from the DNA (50% of the enzyme molecules will have dissociated when $1/(\Delta n\ k_{trans}) = 1/k_{off}$, where Δn is the average number of base pairs translocated after the initial binding event at site n). Such a large value for k_{trans}, yielding such a modest amount of translocation before dissociation, suggests that the translocation mechanism has severe limitations when applied to UDG. One study with the restriction enzyme EcoRV suggested that enzyme translocation along the DNA chain could indeed be as fast as 10^6 base pairs per second,[73,74] but this large value has been subsequently cast in doubt due to limitations in the experimental design for the initial study.[70] In fact, subsequent work concluded that EcoRV located its specific site by a predominantly three-dimensional search process, with an average processivity of much less than 400 base pairs before a dissociation event.[70]

The requirement for an exceedingly rapid translocation rate in order to allow UDG to find its specific site by a processive search mechanism has a profound effect on the probability of repair once a specific site is located [$P_{repair} = k_{enz}/(k_{off} + k_{trans +} k_{enz})$]. Consider that for UDG, k_{off} from a specific damaged site and k_{enz} are both about 1000 s^{-1}.[67,75] In the case of UDG, k_{enz} is the rate of extruding the uracil base from the base stack into the enzyme active site (see below and refs 67 and 76). Thus, in the absence of translocation ($k_{trans} = 0$), the probability of repair is 0.5 ($k_{enz}/k_{off} + k_{enz}$). The introduction of a significant contribution from k_{trans} of 100 000 s^{-1} would reduce P_{repair} to 10^{-2}, requiring multiple encounter events before repair would be affected. Thus, the requirement for rapid translocation is a double-edged sword that in principle may facilitate the search process but, in addition, has ramifications on the efficiency of repair once a specific site is located. In general, several studies examining the processivity of UDG have not found evidence for a large amount of translocation along the DNA chain (<50 bp at 0 salt concentration), and when such effects were investigated at physiological concentrations of salt, no evidence for translocation was observed.[77−79] This amount of processivity is much less than would be expected from a pure scanning mechanism in vivo, where there is an average density of one UDG molecule per 15 000 base pairs of human DNA (see above). Mirroring the findings with UDG, processivity studies with the bifunctional pyrimidine dimer glycosylase, Endonuclease V, MutY, and FPG have also been performed, and no evidence for significant processivity at physiological salt concentrations was obtained.[80,81] It should be noted that these processivity studies have resolution limitations because the experiments used substrates that possessed specific sites that were in quite close proximity along the DNA chain (25−170 bps), requiring very little translocation (or short range hopping) between encounter events at specific sites. It would be very desirable to revisit these measurements using the elegant approach recently used by Stanford et al. on the restriction enzyme *Eco*RV which can distinguish between translocation and hopping mechanisms.[70] As pointed out by these authors, linear diffusion and hopping mechanisms can only be distinguished rigorously by varying the number of base pairs between specific sites. An additional complexity in understanding the damage site location mechanism in a cellular environment is that DNA repair can be localized to transcribed regions of DNA [82] or replication foci.[83]

IV. The Mechanism of DNA Base Flipping

A. General

The nonenzymatic mechanisms of glycosidic bond hydrolysis involving leaving group activation strongly suggest that DNA glycosylases must gain access to sites on the damaged base and sugar that are normally buried in the duplex structure of DNA. For DNA glycosylases, the enzymatic solution to this problem was first uncovered in the crystal structure of human UDG bound to its cognate DNA[84] (although a similar solution had been found previously for cytosine-5-methyltransferase).[85] In a remarkable feat of molecular gymnastics, it was observed that UDG had rotated the damaged deoxyuridine nucleotide 180 ° from its normal location in the DNA base stack into a position firmly ensconced in the enzyme active site (Figure 8). Subsequent crystallographic studies have

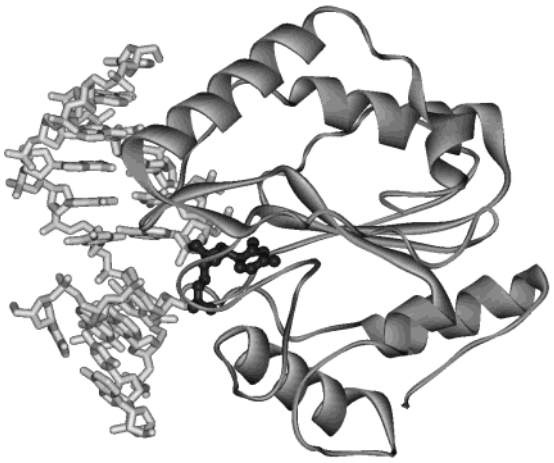


Figure 8. Crystal structure of human uracil DNA glycosylase bound to DNA containing the substrate analogue, pseudodeoxyuridine (ΨdU) (see Figure 17 for chemical structure of ΨdU). The flipped-out ΨdU nucleotide is highlighted in dark gray (PDB code 1EMH).

revealed that "base flipping" is a ubiquitous mechanistic feature of DNA glycosylases, reflecting the evolutionary pressure to recognize damaged bases by their unique electronic and structural features and, in some cases, to form strong interactions that render the base and glycosidic bond electron deficient. To date, enzymatic base flipping has been found to take on three general forms: (i) flipping of the damaged base itself, (ii) flipping of the base located on the opposite DNA strand to the damaged base, and (iii) flipping of both the damaged base and the opposing base. Although crystallography has provided a solid description of the final extrahelical state in the base flipping process, the reaction pathway that is followed is still largely unknown. Before presenting examples of enzymatic base flipping, we first examine the dynamic properties of DNA base pairs that may play a seeding role in initiating base flipping. We then consider the possible role of enzyme induced DNA strain in promoting flight of a base from its position in the DNA base stack and into the enzyme active site.

B. Spontaneous Base Pair Opening in DNA

Enzymatic base flipping introduces a very localized structural perturbation into the DNA duplex that does not extend much beyond the cognate nucleotide. Although DNA base pairs are often viewed as static entities, it is well established that internal base pairs in DNA can also undergo spontaneous and localized opening events at ambient temperature, suggesting that a process related to base flipping can occur in the absence of an enzyme. The pathway for spontaneous opening may be quite different from the enzymatic process, but it is still important to understand whether these spontaneous events are kinetically competent to account for the observed rates of enzymatic base flipping, and what physical barriers exist that must be overcome by enzymes during base flipping.

The dynamic process of spontaneous base pair opening has been studied extensively by following base pair imino proton exchange using NMR spectroscopy.[86] In general, these studies have revealed that AT base pairs have shorter lifetimes ($1/k_{open}$ = 1−5 ms) than GC base pairs ($1/k_{open}$ = ∼10−50 ms), as might be anticipated on the basis of the additional hydrogen bond in the GC pair. The thermodynamic stability of the GC pair usually results in a smaller equilibrium constant for base pair opening ($K_{open} \sim 10^{-7}$) as compared to AT base pairs ($K_{open} \sim 10^{-5}$).[87] Consistent with this simple correlation with hydrogen bond stability, the base pair lifetime for GT mismatches is much shorter than matched base pairs ($1/k_{open} \leq 1$ ms), and the opening equilibrium is much larger ($K_{open} \sim 10^{-4}-10^{-3}$).[88] However, this simple correlation is not universal, as AT base pairs located in AT tracts have exceptionally long lifetimes (≥100 ms)[89] and GC base pairs located in GC tracts can have much shorter lifetimes than those in other sequence contexts (≤5 ms).[90] It is likely that the differential stacking interactions of the various bases, sequence-specific steric conflicts, and differences in the widths and hydration of the DNA grooves in different DNA sequences also play a significant role in the thermodynamics and kinetics of spontaneous base flipping.[90] Thus, in the absence of deformations to the DNA induced by enzyme binding, sequence-dependent structural features of the native DNA duplex can have a strong influence on the dynamics of base pair opening, and possibly enzymatic recognition of damaged sites in DNA.

What does the pathway for nonenzymatic base flipping look like? The only approach to address this question at the current time is computational chemistry. One recent study used umbrella sampling and a novel center-of-mass pseudodihedral reaction coordinate to calculate the potential of mean force for flipping a C or G base from a Watson−Crick base pair.[91] The complete pathways for base flipping from the major and minor grooves were investigated from which it was concluded that both pathways have nearly equal energetic barriers. This was a surprising result, given that flipping through the minor groove was previously considered to have a large steric barrier due to clashes between the departing base and its base-paired partner.[92] However, the compu-

tational results suggest that the DNA backbone flexes to remove this steric obstacle, thereby allowing efficient departure via the minor groove. The calculated barrier heights were in the range 15−20 kcal/mol for C and G flipping, respectively, which are consistent with the rates of imino proton exchange measured by NMR. The trajectories of the bases during the flipping process were quite unexpected, with the base normal oriented nearly perpendicular to the helical axis, with the target base tracking along the exit groove (Figure 9). Both the minor and major

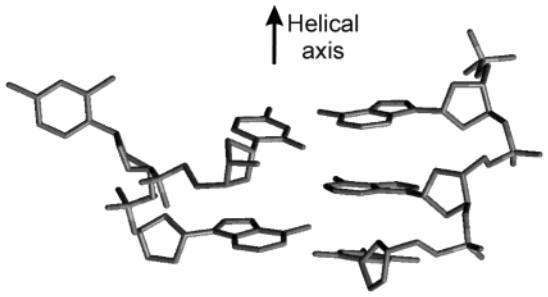


Figure 9. Structural model of a fully extrahelical deoxyuridine base. The model is based on the major groove spontaneous base-flipping pathway determined in potential-of-mean-force calculations.[91] Note that the normal of the base is perpendicular to the helical axis (see text).

groove pathways involved similar phosphodiester backbone dihedral distortions, suggesting that such distortions are a requirement for base flipping and may be relevant in the enzymatic process.

C. How Do Enzymes Initiate Base Flipping?

As will become apparent from the upcoming discussion of enzymatic base flipping, all enzymes that flip bases from DNA also bend DNA at the target site. This observation raises the questions of when such distortions to the DNA occur (i.e., prior to, simultaneous with, or after base flipping) and whether DNA glycosylases bend undamaged DNA. Consideration of the role of DNA distortion and strain in base flipping has been driven by the computational proposal that base pair opening required less energy in bent or underwound DNA.[92] More recent computational studies have suggested a linkage between the flexibility of a damaged site in DNA and the energy required to bend the DNA.[93,94] This result suggests that damaged DNA may be intrinsically more flexible, which may facilitate enzyme induced DNA bending and base flipping.

These computational proposals have found some validation in experiment. It has been found that introduction of torsional stress in a DNA duplex, through the cross-linking of two guanine residues across the minor groove with a chemical tether, promoted partial rupture of a base pair and partial extrusion of a cytosine residue resulting in DNA bending.[95] Thus, base flipping and DNA bending apparently relieved the torsional strain induced by constraining the minor groove width. It was estimated that the energetic cost of inducing base pair rupture by torsional stress was only 3 kcal/mol, suggesting that enzymes could easily employ similar strategies.

V. Examples of Enzymatic Base Flipping

A. Damaged Base Flippers

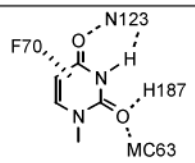
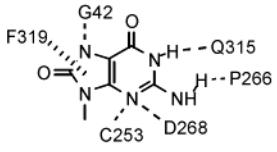
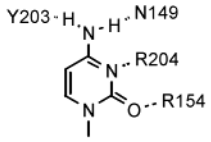
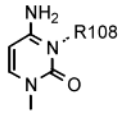
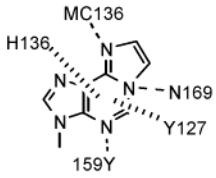
The most common type of base flipping that has been observed to date is where the damaged base itself is yanked from the DNA base stack into the active site pocket of the enzyme (Figure 9). Five examples of enzymes that use this type of strategy are listed in Table 4, along with a summary of the contact interfaces for each enzyme.[96−100] From this summary, it is clear that these enzymes share many similar interactions, the conservation of which would indicate a shared functional role in base flipping worthy of understanding.

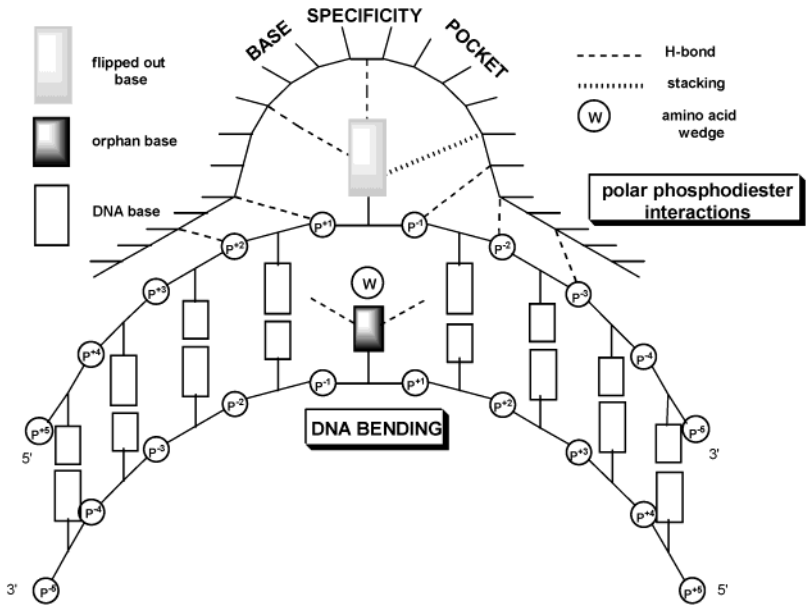
1. DNA Bending

One key commonality between these structurally diverse enzymes is the high density of interactions with the DNA phosphodiester groups on the damaged strand that extends 2 or 3 nucleotides on either side of the damaged base. These polar interactions can assume various forms, but invariably, a main chain or side chain hydrogen bond donor interacts with a nonbridging oxygen acceptor of the phosphodiester linkage. In contrast, much fewer or no interactions are observed with the phosphodiesters of the undamaged strand. The most likely functional role for these interactions is to distort the DNA backbone chain leading to its bending away from the surface of the enzyme. Such induced bending may serve to destabilize the base pairing and stacking of the damaged base and introduce a torsional stress that serves to promote flight of the base from the DNA stack. In this viewpoint, the enzymes sacrifice binding energy to distort the DNA into a conformation that is predisposed for base flipping. Recent computational studies have suggested that damaged sites are more flexible than undamaged sites, which also decreases the kinetic barrier for DNA bending and base flipping at damaged sites (see above).

Recently, single-molecule atomic force microscopy has been used to address the key question of whether hOGG1 bends nonspecific DNA.[101] In this work, the interaction of hOGG1 with linear DNA containing a single oxoG:C specific site or, alternatively, only nonspecific binding sites was visualized. For both types of DNA, significant bending was observed, with an average bend angle of 72° that was indistinguishable from the angle observed in the crystal structure of hOGG1 bound to specific DNA (Table 4).[98] Interestingly, for the nonspecific binding events a bimodal distribution of bend angles was observed, with a second Gaussian distribution centered on a bend angle of 0°. Assuming that interactions with the surface did not perturb the equilibrium for DNA bending, it was calculated that K_{bend} = [bent DNA]/[unbent DNA] = 2. It was further argued, on the basis of structural observations, that bending of nonspecific DNA required base flipping, and therefore, hOGG1 must flip undamaged bases as well. The authors proposed a damage search model where hOGG1 scans down DNA, bending and flipping bases into its active site until the damage site is encountered. Of course, AFM experiments

Table 4. Summary of the DNA Contact Interactions for Selected DNA Glycosylases that Flip Damaged Bases

Enzyme	Damaged strand phosphate Interactions	Undamaged strand interaction	Wedge group (W)	Damaged base interactions	Orphan base interactions	DNA bend angle
UDG	+1, -1, -2		L191			45°
hOGG1	+1, -1, -2, -3		N149			70°
FPG	+1, -1, -2	-1, -2	R108			75°
AlkA	+2, +1, -1, -2, -3	-1	L125			60°
AAG	+1, -1, -2, -3	+1, -3, -4, -5	Y162			22°



provide no information on the processivity of hOGG1, and the extremely weak binding affinity of hOGG1 for nonspecific DNA ($K_D \sim 45\ \mu M$) would not be expected to allow large amounts of DNA translocation, as discussed above for UDG. Regardless of the finer interpretations of these results, these observations clearly suggest that nonspecific bending does occur, at least with hOGG1, which likely promotes flipping of the 8-oxoguanine base when it is encountered.

2. Pushing and Plugging: The Role of the Amino Acid Wedge

A second highly conserved interaction shared by these enzymes is the side chain wedge group that occupies the hole in the DNA base stack left behind by the departed cognate base (Table 4). The identity of the wedge group is less important than its steric bulk, which serves to plug the hole and prevent reinsertion of the damaged base, and may also serve

to push the damaged base from the duplex as the enzyme bends and embraces the DNA. The strongest evidence for these proposed roles are from studies with UDG where it was found that removal of the Leu191 wedge caused an 60-fold weakening of site-specific binding.[102,103] This effect was attributed to the inability of the L191G mutant to undergo a conformational change that docks the uracil base into the active site. Since fluorescence measurements with 2-aminopurine labeled DNA established that L191G was able to partially flip the uracil from the duplex (see part 4 below), the combined results indicated that the process was arrested at a metastable flipped state, in which the uracil was out of the duplex but not stabilized by the specific hydrogen bonding groups and stacking interactions that are located deep in the active site pocket (Table 4). Thus, the bulk from Leu191 appears to be required for stable docking of the base. Another conclusion that emerges from these experiments is that the metastable flipped state can be formed in the absence of Leu191, which suggests that the DNA backbone distortions arising from phosphodiester interactions can propel the uracil into this partially flipped state. Structural characterization of this metastable intermediate on the base flipping pathway would perhaps validate recent computed trajectories for base flipping on enzymes.[104]

The role of Leu191 has also been elucidated using a chemical rescue approach.[102,105] In this strategy, the wedge function of Leu191 is replaced by a bulky pyrene wedge nucleotide that is located on the opposite DNA strand directly facing the uracil in the DNA substrate (Figure 10). The initial hypothesis

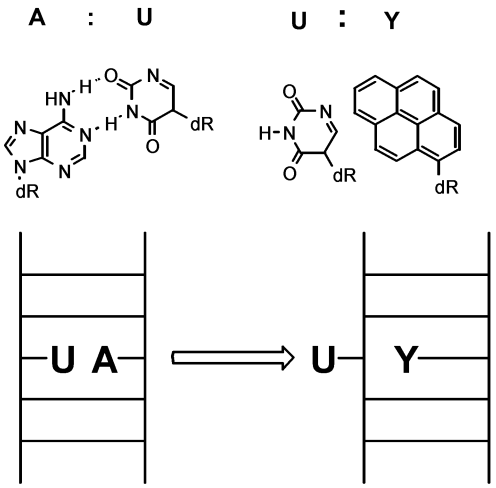


Figure 10. Utilization of an artificial pyrene nucleotide wedge to force a uracil base into an extrahelical position and facilitate damaged site recognition (see text).

that drove these investigations was that the large planar pyrene molecule would occupy the entire base stack and push the uracil into an extrahelical conformation in the free DNA (Figure 10) and, therefore, rescue the damaging effects of the L191A and L191G mutations. In fact, the strategy worked exceedingly well, and it was found that pyrene fully restored wild-type activity to the L191A mutant. Most importantly, pyrene had little or no palliative effect on mutations that exerted their effects in the chemical step of the reaction, or other base flipping mutations that exerted their effects in other ways than Leu191. Thus, the proposed wedge function of Leu191 has been directly confirmed by this novel chemical rescue approach.

3. The Base Specificity Pocket

The final interactions with the extrahelical base are found deep in the active site pockets of these glycosylases. For most, the active site is designed both to be sterically compatible with the shape of the damaged base and to allow formation of hydrogen bonds between donor and acceptor groups on the damaged base and enzyme (Table 4). These two effects—steric exclusion of alternative bases and specific hydrogen bonding—largely explain the specificity of these enzymes for binding damaged DNA.[76,106−108] For others, such as AlkA and AAG, the active site pockets are devoid of hydrogen bonding groups, and recognition seems to involve π-cation interactions between aromatic groups in the active site and the cationic damaged base.[109] This more promiscuous recognition strategy allows these enzymes to act on unmodified purine bases, albeit at a much reduced rate.[110]

How is base flipping by DNA glycosylases used to obtain their tremendous catalytic specificity? The available evidence indicates that the largest component in determining the catalytic specificity of DNA glycosylases is not found at the binding step, but instead arises at the chemistry step. For instance, hOGG1 has a specificity for binding to damaged sites as opposed to undamaged sites of about 400,[101] and for UDG, the binding discrimination for specific and nonspecific sites is only around 100.[67] These binding specificities cannot account for the enormous catalytic specificities of these enzymes, which can approach 10^6-fold.[67] It is likely that the specific interactions with the damaged base are used to drive a conformational change in the enzyme and/or DNA that is required for catalysis. Such conformational events, which cannot be induced by nonspecific DNA binding, provide a gateway to the catalytic machinery that can only be unlocked by the damaged base. Induced-fit specificity is well documented for UDG, where free uracil, or dU-DNA, can induce an open to closed conformational change in the enzyme that is not observed upon nonspecific DNA binding.[67,111,112]

4. A Detailed Kinetic Description of the Base Flipping Pathway for UDG

What are the temporal events in the pathway for forming the final flipped-out state? For the enzyme UDG, extensive kinetic, mutagenesis and chemical rescue studies have uncovered the stepwise nature of the enzymatic flipping pathway, and established the temporal sequence for formation of the phosphodiester, wedge, and specific base interactions described above.[67,75,76,102,105] These kinetic studies were facilitated by the incorporation of a fluorescent reporter nucleotide, 2-aminopurine (2-AP), adjacent to a nonhydrolyzable deoxyuridine analogue (2′-fluoro-2′-deoxyuridine) in a DNA context. 2-Ami-

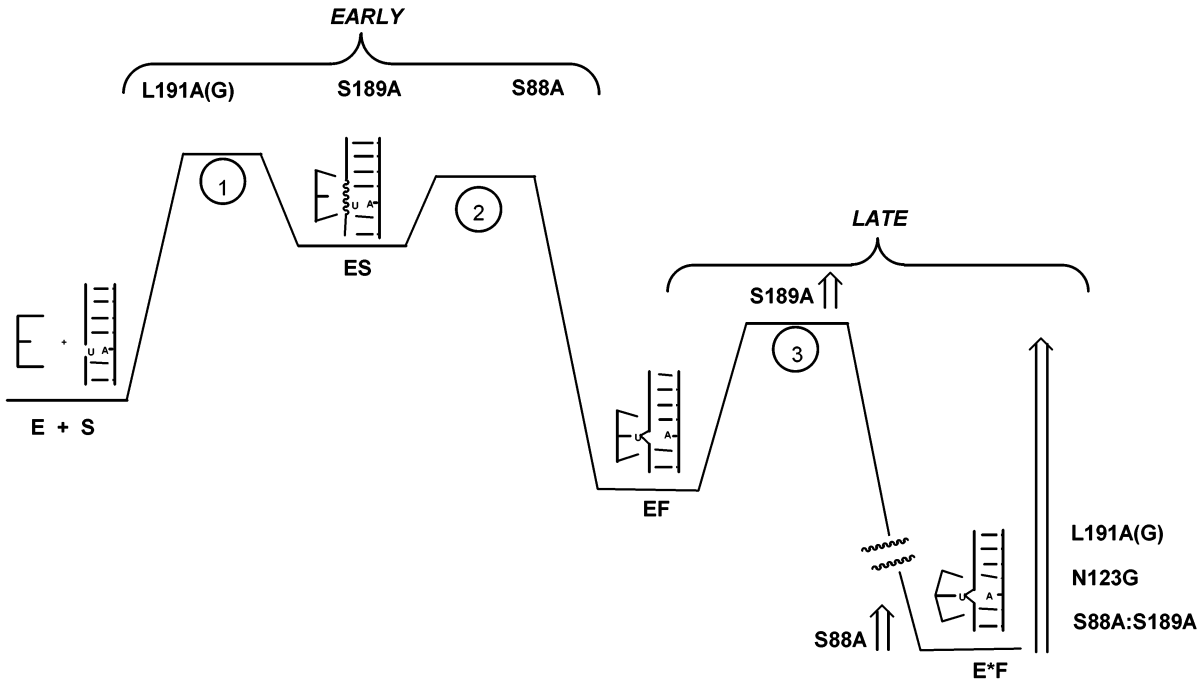


Figure 11. Temporal pathway for base flipping by UDG. The pathway was explored using mutagenesis and stopped-flow fluorescence measurements.[76,105] The perturbing effects of removing the enzyme groups that are involved in pushing and pulling the uracil into the active site are indicated by the arrows (see text).

nopurine has the favorable properties that its fluorescence is quenched when it is stacked in the DNA duplex, but increases 2−10-fold when stacking interactions are removed.[67,113,114−118] Thus, flipping of the adjacent uracil results in a large increase in 2-AP fluorescence that allows real-time monitoring of the base flipping process. In addition, it was found that base flipping results in a 2-fold decrease in the tryptophan fluorescence of UDG, which provided an independent signal to monitor whether the induced-fit conformational change is concerted with extrusion of the uracil from the duplex, or occurred in a distinct step. Finally, these two spectroscopic tools allowed microscopic dissection of the kinetic behavior of site-directed UDG mutants, allowing mapping of the steps that were defective for each mutant. The results from these comprehensive studies are summarized in the free energy reaction coordinate diagram in Figure 11.

The base flipping process for UDG was found to consist of at least three discrete steps: the rapid formation of a weak encounter complex in which the uracil base is still intrahelical ($k_1 = 220\ \mu M^{-1}\ s^{-1}$, $k_{-1} = 600\ s^{-1}$), formation of a metastable complex in which the uracil base is partially extrahelical but not yet fully docked in the active site pocket ($k_2 = 700\ s^{-1}$, $k_{-2} = 180\ s^{-1}$), and a conformational docking step in which the enzyme clamps around the fully extrahelical base ($k_3 = 350\ s^{-1}$, $k_{-3} = 100\ s^{-1}$). Thus, formation of the final Michaelis complex requires passage through three gates, and it is likely that nonspecific DNA only makes it through the first gate.[67] Interestingly, UDG interacts with uracil-containing single-stranded DNA with no greater facility than with duplex DNA, which is surprising given that no obvious barriers exist for flipping a base from single stranded DNA.[67] This intriguing result suggests that single-stranded DNA may be entropically disadvantaged for binding and flipping into the UDG active site, whereas the duplex structure restricts the number of backbone DNA conformations and provides for enhanced binding. To facilitate a qualitative description of the mutational effects in Figure 11, we have classified the effects as "early or "late". Early effects include all ground states and transition states leading to formation of the metastable extrahelical state (EF), and late effects include the transition state and ground state for formation of the final closed conformation (E*F).

Independent removal of the two serine groups (Ser88 and Ser189) that interact with the +1 and −1 phosphodiester groups of the substrate (Figure 11) was found to destabilize specific DNA binding by less than 1.8 kcal/mol, reflecting modest perturbations at both early and late steps of the base flipping pathway.[76] However, removal of both serine groups severely destabilized the closed complex, essentially arresting the reaction at the metastable EF intermediate. These results indicate that formation of the metastable state does not require these serine phosphodiester interactions, but they become quite important in stabilizing the final closed state that precedes the chemical step. Thus, if these serine groups are involved in DNA bending,[119] it is possible that profound bending does not take place until the final E*F state is reached.

Removal of the Leu191 wedge group of UDG affected the kinetics for formation of the ES and/or EF complexes in Figure 12, indicating that Leu191 interacts with the DNA substrate before the final docking step to form E*F. The most profound effect of removing Leu191 was the extreme destabilization of the final E*F closed state: no UDG tryptophan fluorescence decrease was observed for either the

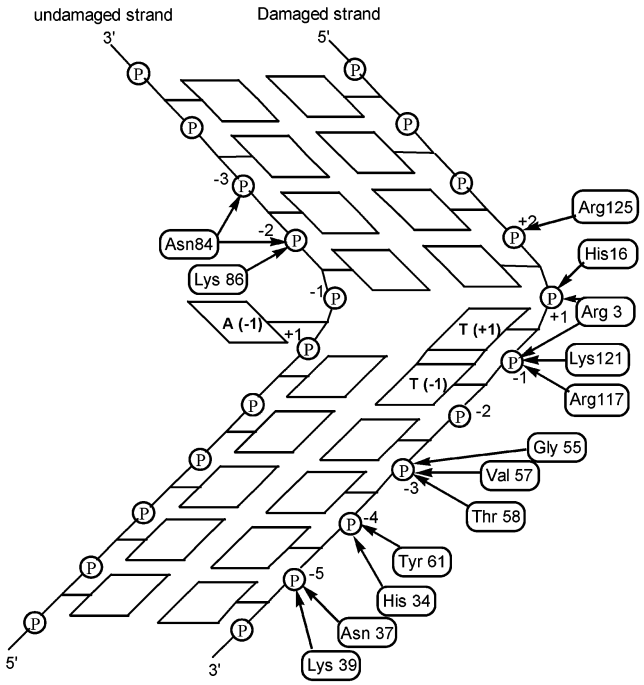


Figure 12. Interaction map for T4 PDG with its cognate pyrimidine dimer containing DNA. Unlike the enzymes in Table 4, PDG flips the base that is on the opposite strand to the damaged site.

L191A or L191G mutant, and the specific DNA binding affinity of these mutants was decreased by 10−60-fold.[76,102] Thus, the pushing function of Leu191 may play a role in the early effects, and the plugging function of Leu191 plays a late role in stabilizing the E*F complex. Remarkably, the pyrene wedge substrate (see Figure 10) fully rescued every kinetic and binding defect arising from removing Leu191.[102,105]

What effect on base flipping does removal of a group in the uracil specificity pocket produce? We asked this question by deletion of the side chain of Asn123, which forms hydrogen bonds to uracil O4 and N3 (Figure 11 and Table 4). The N123G mutant was damaged in specific DNA binding by 1.8 kcal/mol and, like L191G, was unable to proceed through the conformational change, requiring a late effect (i.e., destabilization of E*F). Surprisingly, the 2-AP fluorescence increase for N123G was 7 times greater than that for wild-type UDG or any other mutant that was investigated.[76] This intriguing result suggests that alterations of the specificity pocket that change its shape or hydrogen bonding properties may allow flipping of other bases—even 2-AP itself—as suggested by the large fluorescence increase observed here. Previous work supports this conclusion, as UDG has been converted into a cytosine DNA glycosylase by changing Asn123 to Asp and, additionally, into a thymine DNA glycosylase by removing a bulky tyrosine residue (Tyr66) that would otherwise sterically clash with the 5-methyl group of thymine.[110]

B. An Opposing Base Flipper: Pyrimidine Dimer DNA Glycosylase (PDG)

Pyrimidine dimer glycosylases are bifunctional enzymes that act to remove cis−syn cyclobutane pyrimidine dimers resulting from ultraviolet light damage to DNA (Table 1, see below). An atomic model of the enzyme from T4 phage (T4 PDG) complexed with a duplex DNA substrate containing a thymine dimer has been solved, and reveals that this enzyme has apparently taken a different strategy for accessing the glycosidic bond of the 5′ thymime base: flipping of the opposing adenine base from the complementary strand (Figure 12).[97]

The structure reveals several aspects of damage site recognition that differ considerably from the damaged base flippers discussed above. First, the enzyme bends the DNA in different way by grabbing onto both strands on either side of the damaged site, whereas the previous examples showed a distinct bias for forming DNA backbone interactions on the damaged strand only (Table 4). However, much like the damaged base flippers, there is a conspicuous high density of enzyme interactions with the phosphodiester groups that immediately flank the damaged thymines, suggesting a similar conserved function for these interactions in DNA bending. It is likely that the unique interactions on both strands explain the 60° kink in the DNA, which in turn promotes the unusual flipping of the adenine base opposite to the damage site. Once the opposing adenine is extrahelical, it does not dock into a highly specific binding pocket as seen with the damaged base flippers. Instead, the adenine is sandwiched between a group of protein residues that apparently hold it in position by weak van der Waals forces, and accordingly, the enzyme has little preference for the base that opposes the 5′ thymine of the dimer.[120] In addition, there is no wedge group apparent that might be involved in stabilizing the extrahelical adenine as observed for damaged base flippers. The requirement for displacing the opposing adenine apparently lies in the rigid nature of the pyrimidine dimer, which precludes direct flipping of the damaged base. Opposing base flipping provides an alternate route for catalytic residues to access the damaged thymine, much like the function of damaged base flipping.

Flipping of the opposing base by T4 endo V has also been studied using the fluorescent adenine analogue 2-aminopurine as discussed above for UDG, although detailed kinetic measurements have not yet been performed in this system.[121] Evidence for base flipping in solution was provided by a selective 4-fold increase in 2-AP fluorescence when 2-AP was placed opposite to the excised 5′ thymine, but not when it was placed opposite to the adjacent 3′ thymine of the dimer. The extrahelical state was further confirmed by the high sensitivity of the 2-AP fluorescence to quenching by added acrylamide, indicating significant solvent exposure of the 2-AP base opposite to the 5′ thymine. The extrahelical base seems to persist after cleavage of the glycosidic bond because DNA containing a product-like abasic site analogue opposite to the 2-AP base also showed a large 2-AP fluorescence increase upon PDG binding. Thus, flipping does not require the presence of the specific damage, but it is likely facilitated by the flexibility of the site, or the constraints on the DNA duplex imposed by the thymine dimer.

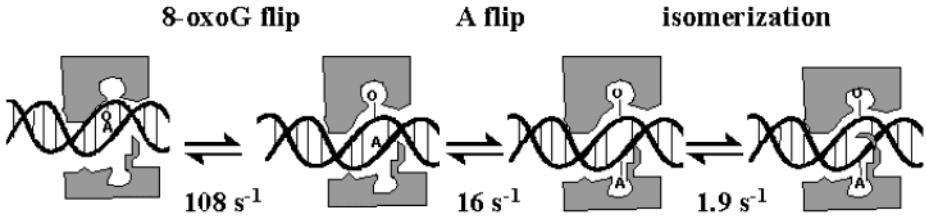


Figure 13. Double base flipping mechanism of MutY adenine DNA glycosylase. Adapted from ref.[129]

C. A Double Flipper: MutY Glycosylase

MutY glycosylase acts to remove normal adenine bases that have been misincorporated by DNA polymerase opposite to a damaged 8-oxoguanine base.[122,123] This occurs due to the ability of 8-oxoguanine to form a Hoogsteen base pair with adenine, as well as a normal Watson−Crick base pair with cytosine.[124,125] Thus paradoxically, MutY must remove a natural base that is paired with a damaged base in order to prevent G/C → A/T transversions in the genome. MutY is a monofunctional glycosylase and a member of the helix−hairpin−helix superfamily of DNA glycosylases.[126] It consists of a 26 kDa catalytic core and a 13 kDa C-terminal domain that was, until recently, of unknown function. A crystal structure of the catalytic domain bound to the target base adenine established that MutY flipped adenine from the duplex into an active site pocket, reminiscent of the damage base flippers discussed above.[126] However, the interesting question of how this enzyme recognized adenine only in the context of A/8-oxoG and A/G mismatches was unresolved until Noll and Clark made the observation that the C-terminal domain of MutY showed significant homology to the MutT enzyme.[27] This was a seminal insight because MutT is a pyrophosphohydrolase that cleaves 8-oxo-dGTP to 8-oxo-dGMP and pyrophosphate, thereby ridding the nucleotide pool of a mutagenic nucleotide triphosphate.[127] Thus, the intriguing suggestion was that MutY possessed an 8-oxoguanine binding pocket similar to MutT. This proposal was later confirmed by a limited NMR structural study of MutY, which established a similar secondary structure between MutY and MutT.[128] Putting the crystallographic and informatics studies together strongly suggested that MutY had an adenine binding pocket in its catalytic domain, and an 8-oxoguanine binding pocket in its MutT-like C-terminal domain.

With this information in hand, Wong and colleagues set out to test whether MutY flipped both adenine and 8-oxoguanine (i.e., a double-flip).[129] In experiments using duplex DNA with an 8-bromoguanine mispaired with adenine, they observed UV-cross-linking between MutY and the DNA. Although it was not established which residues of MutY were cross-linked, these results strongly supported the double-flip mechanism. Borrowing from the 2-AP fluorescence methods originally developed for UDG and other enzymes (see above), these workers studied the rapid kinetics of base flipping by MutY and proposed a multistep pathway as shown in Figure 13. In this model, 8-oxoguanine flipping occurs first with a rapid rate constant of 108 s^{-1}, which is then followed by much slower adenine flipping (16 s^{-1}). A final very slow enzyme isomerization step was proposed to account for a slow but large amplitude 2-AP fluorescence change (1.9 s^{-1}). This work suggests that there will be considerable diversity in the mechanisms and rates for base flipping and that Nature has taken great care to evolve a mechanism that ensures that MutY only removes adenine opposite to 8-oxoG, and not from normal A/T base pairs. In addition, a very recent study on the interaction of MutY with specific DNA has concluded, quite unexpectedly, that the active form of the enzyme is a dimer bound to one DNA duplex ($MutY_2$·DNA), and that binding to a second DNA duplex completely inhibits the glycosylase activity of the enzyme ($MutY_2$·DNA_2).[130] It is not clear how these complexities affect the mechanism of base flipping shown in Figure 13.

VI. Enzymatic Strategies for Cleaving the Glycosidic Bond

Since 1995, there have been more than 50 crystal structures and one NMR structure of DNA glycosylases deposited in the protein data bank, and many of these are cocrystal structures in the presence of damaged DNA analogues. It is difficult to overstate the importance of these structures in the pursuit to understand the chemical mechanisms of these enzymes. However, as will become apparent below, the emergence rate of new structures has far eclipsed the rate of detailed biophysical studies. Thus, many proposed glycosylase mechanisms are based solely on observations from static crystal structures or limited mutagenesis studies and, therefore, have not been rigorously tested using the tools of enzymology, chemistry and spectroscopy. Therefore, each of the proposed mechanisms discussed below must be embraced cautiously, keeping closely in mind the sometimes low density of supporting data.

A. Monofunctional Pyrimidine Specific Glycosylases: UDG, MUG, and TDG

1. General

It is of interest to understand whether enzymes that catalyze similar reactions have evolved similar transition states and chemical mechanisms. Although transition state and spectroscopic information is lacking on all DNA glycosylases except UDG, it is still possible to ask the question of whether the active site interactions of these three pyrimidine specific enzymes are consistent with promoting formation of a similar transition state as observed for UDG (see below). Each of the enzymes considered here has been crystallized by itself or with damaged DNA, and the

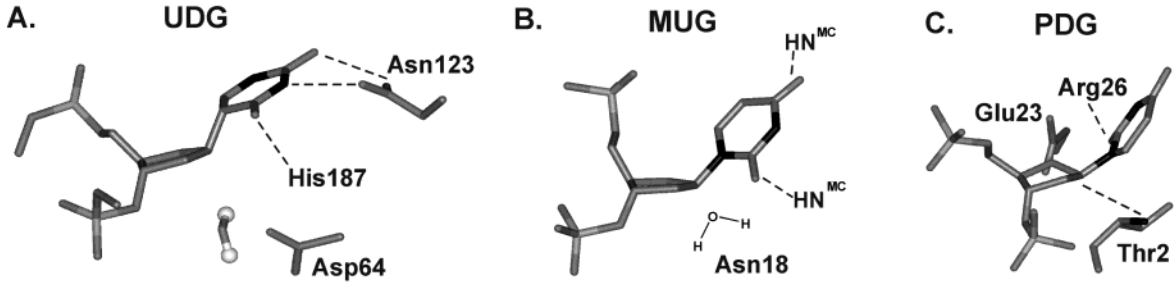


Figure 14. Selected active site interactions of UDG, MUG, and T4 PDG with their cognate substrates. These depictions were obtained from crystal structures with the following PDB accession codes: UDG (1EMH), MUG (1MWJ), and PDG (1VAS). (A) The complex of UDG was obtained with a stable C-glycoside analogue of deoxyuridine (ψdU); see text and Figure 17. (B) The complex of MUG was obtained with a stable 2′-fluoro-2′-deoxyuridine substrate analogue. (C) The stable complex of T4 PDG with pyrimidine dimer DNA was obtained by removal of the catalytic glutamate (E23Q). For clarity, only the 5′ pyrimidine of the dimer is shown.

key active site interactions with their respective substrates are shown in Figure 14 for reference in the following discussion. Our tactical approach will be to summarize the observations and conclusions that have been obtained from detailed mechanistic studies of UDG, and then examine whether the active sites of the other enzymes suggest similar mechanisms. The substrates for these glycosylases are summarized in Table 2.

2. The Nature of the Transition State

Only one DNA repair glycosylase, UDG, has been subjected to transition state analysis using the KIE approach.[131] This work required development of a total enzymatic synthesis method for incorporation of isotopically labeled dUTP in DNA.[131,132] For UDG, a family of four KIE measurements was made as indicated in Figure 15. These experimental KIEs

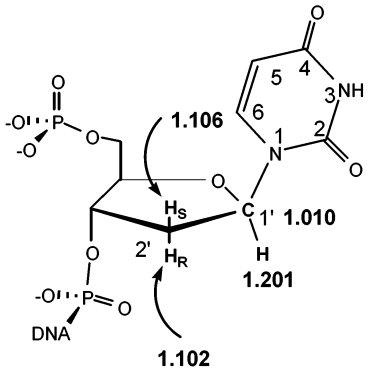


Figure 15. A family of four KIE measurements performed on the UDG enzyme that indicated a $D_N{*}A_N$ mechanism.

unambiguously indicated an extremely dissociative transition state for glycosidic bond cleavage, with extensive oxacarbenium ion character, and strongly suggested, but were not sufficient to prove, that a $D_N{*}A_N$ reaction was followed. Indeed, both stereospecific 2′ β deuterium KIEs (~1.11) were near the theoretical values expected for $D_N{*}A_N{}^{\ddagger}$ mechanism (1.1−1.13).[133] The β deuterium KIEs indicated that the stabilizing hyperconjugative effects resulting from electron donation from the 2′ CH σ bonds to the electron deficient anomeric carbon were maximal. The geometric dependence of the 2′ β deuterium secondary KIEs required that the sugar assume a unusual 3′-exo or endo pucker in the transition state, and they were subsequently validated when it was observed that the sugar pucker in the strained Michaelis complex with the pseudodeoxyuridine substrate analogue was 3′-exo (Figure 14A). The 1′ α deuterium secondary KIE of 1.22 was also near the theoretical maximum of 1.24 for a $D_N{*}A_N{}^{\ddagger}$ mechanism, indicating complete rehybridization of this sp^3 center to sp^2 geometry in the transition state, implying extensive oxacarbenium ion character. The primary ^{13}C1′ KIE was only 1.01 ± 0.009, which overlaps with the theoretical value expected for a $D_N{*}A_N{}^{\ddagger}$ reaction mechanism (1.005). These measurements established that the UDG active site must be assembled to stabilize a highly ionic transition state and intermediate. Thus, all mechanistic proposals for UDG must be viewed in the light of this key observation. Similar transition state structures might be expected for the other enzymes if analogous active site interactions are observed in their structures, although subtle changes in transition state structure cannot not be reliably predicted from ground-state structures alone.

3. Pyrimidine Leaving Group Activation I: Base Interactions

Because of the absence of viable proton acceptor groups on pyrimidine bases, it seems difficult to envision how pyrimidine specific glycosylases could activate their substrates by base protonation. Nevertheless, the structure of the UDG active site shows that every electron lone pair and hydrogen bond donor on the uracil base is involved with an interaction with a conserved group in the enzyme active site, and a potential proton donor, His187, is directed at the uracil O2 atom (Figure 14A). However, extensive heteronuclear NMR measurements have unambiguously determined that His187 is neutral, negating the possibility that an extremely low pK_a imidazolium group donates a proton to uracil O2.[112,134,135] Thus, the key questions are the following: (i) which of the observed interactions with the uracil base are important in lowering the activation barrier, and (ii) does full proton transfer to the uracil base occur? With respect to these questions we consider the relative roles of His187 and Asn123, the two conserved groups that interact with the O2 and O4 electronegative atoms on the uracil ring.

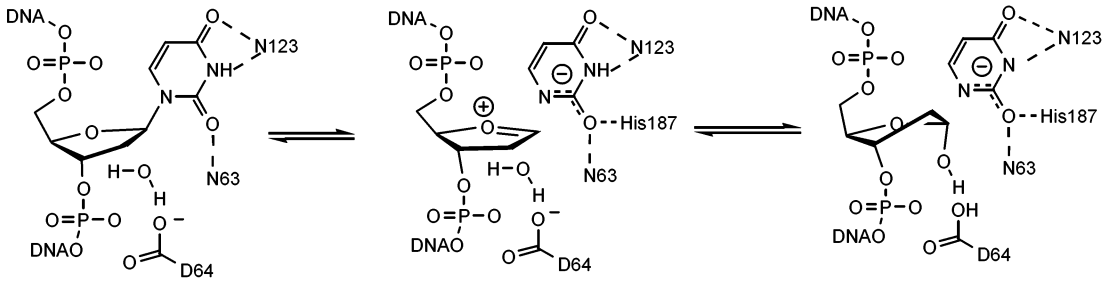


Figure 16. Electrostatic sandwich mechanism for stabilization of the oxacarbenium ion in the UDG active site. Stabilization of this otherwise unstable species is brought about by the combined electrostatic environment provided by the uracil anion, Asp64, and the phosphodiester groups of the DNA. Adapted from ref 139.

One approach these enzymes could take to facilitate charge migration onto the pyrmidine base, without having to surmount the large thermodynamic barrier for full protonation of the exocyclic carbonyl oxygen atoms of the substrate, is to form hydrogen bonds that are weak in the ground state and become increasingly stronger as the transition state is approached. As pointed out by Herschlag, such mechanisms are catalytic when there is a large pK_a mismatch between donor and acceptor groups in the ground state (weak hydrogen bonding), which diminishes or disappears as the anionic transition state is approached, and the pK_a values become more closely matched (stronger hydrogen bonding).[136,137] Such effects can be energetically amplified in an enzyme active site of low dielectric as compared to water because of the stronger Coulombic attractive forces between acceptor and donor groups with matched pK_a values. For instance, uracil O2 accepts a hydrogen bond from the NH$^\epsilon$ donor of His187 in the UDG ground state. Since the estimated pK_a values for O2 and His187 are -3 and $\sim$14,[33,138] a huge pK_a mismatch of 17 units is indicated, leading to weak hydrogen bonding. However, in the transition state, where a significant amount of charge has resonated onto the O2 atom of the leaving group, the O2 pK_a increases by 10 pK_a units (p$K_a^{\text{O2 imidol}} \sim 7$),[135] reducing the p$K_a$ mismatch between the O2 acceptor and the His187 donor by perhaps 10 units (this assumes no pK_a change in the transition state for His187). Consistent with this model, UDG binds the neutral uracil base 500 times more weakly than the N1$-$O2 imidate form, and removal of His187 by mutagenesis destabilizes the transition state of the reaction by 5 kcal/mol without affecting ground-state substrate binding.[68]

How does the interaction between His187 and uracil O2 facilitate leaving group departure? Detailed heteronuclear NMR studies have established that His187 forms a strong hydrogen bond between its N$^\epsilon$H$^\epsilon$ and the uracil O2 anion.[112] The role of the His187-uracil O2 hydrogen bond, which results in an extremely desheilded chemical shift for H$^\epsilon$ (15.75 ppm),[112] is to lower the pK_a of the departing N1 atom of the uracil leaving group by at least 3.4 units (5 kcal/mol), which corresponds nicely with the energetic effect of removing His187.[68] The similarity between the kinetic effect of removing His187 and its role in lowering the pK_a of the leaving group is reminiscent of the nonenzymatic studies described earlier, which showed a direct linear correlation between the hydrolysis rate of deoxyuridine derivatives in water, and the pK_a of the uracil leaving group. These results suggest that the character of the nonenzymatic and enzymatic transition states may not differ by large extremes.

Quite surprisingly, heteronuclear NMR studies have also shown that the uracil base remains anionic in the enzyme active site (p$K_a \leq 6.4$) until it is released to solution after the departure of the abasic DNA.[112,139] Motivated by these findings, we have suggested a novel role for "product assisted" catalysis in which the stable uracil anion serves to stabilize the positive charge on the oxacarbenium ion intermediate (Figure 16). Together with the negative charge provided by the conserved aspartate located on the opposite face of the sugar, an "electrostatic sandwich" is formed that cradles the unstable oxacarbenium ion.[140] Thus, evolution has licked the problem of pyrimidine leaving group activation by dramatically lowering its pK_a in the transition state through strong hydrogen bonding to the O2 position, and then avoiding full proton transfer, which would negate the electrostatic stabilization of the cationic intermediate provided by the anionic leaving base (Figure 16). We have estimated a 6 kcal/mol contribution of the uracil anion to the stability of the oxacarbenium ion intermediate based on binding studies using a cationic oxacarbenium ion mimic.[140]

The second group that interacts with the uracil base is Asn123 (Figure 14A). In contrast with the exclusive interaction of His187 in the transtion state, removal of Asn123 has a $+2.5$ kcal/mol effect on ground state binding, indicating that its bidentate hydrogen bonds with uracil O4 and N3 are important in ground state stabilization. In fact the dissociation constant of the N123G enzyme for deoxyuridine containing DNA is indistinguishable from the affinity of wild-type UDG for nonspecific DNA (see above), resulting from the inability of N123G to stabilize the closed conformation of the Michaelis complex (E*F in Figure 11). Thus, the low activity of N123G may be partially attributed to nonproductive binding. In addition, Asn123 may play a lesser role in lowering the activation barrier as compared to His187 because, as the C1′$-$N1 bond is cleaved, the dominant flow of electron density is to uracil O2 rather than O4 (Figure 14A).

Do the other pyrimidine specific enzymes show evidence for a similar mechanism of leaving group activation and sugar stabilization? Both MUG and TDG are 10^2-10^4-fold less powerful enzymes than

UDG, as determined from single turnover kinetic measurements (Figure 2). Single-turnover measurements are important here because steady-state catalysis by all of these enzymes is severely limited by slow release of the abasic DNA product.[28,141] Although these enzymes have nearly identical folds as UDG, structural studies have revealed that both lack the conserved histidine of UDG that strongly hydrogen bonds with uracil O2, and also, the conserve aspartate that is proposed to electrostatically stabilize the glycosyl cation intermediate (Figure 14B).[142] It is difficult to envision how these enzymes could follow such a dissociative mechanism as found with UDG, given that the histidine and aspartate groups suggested to be very important for establishment of the $D_N{*}A_N$ mechanism are missing. We speculate that these enzymes could still follow a very dissociative A_ND_N type of mechanism, but the likelihood of forming a discrete oxacarbenium ion intermediate is greatly diminished. KIE studies of mutant UDG enzymes that lack the conserved histidine and aspartate are being performed and will establish whether a stepwise mechanism is possible in the absence of these groups.

4. Pyrimidine Leaving Group Activation II: Coupled Stereoelectronic Effects and Strain

Potentially one of the most remarkable features of the UDG mechanism is observed in the structure of UDG cocrystallized with DNA containing a stable C-glycoside analogue of deoxyuridine, pseudodeoxyuridine (ΨdU, Figure 17).[96] In this structure, the C1

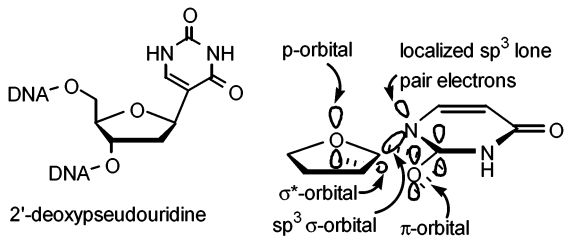


Figure 17. Structure of the UDG substrate analogue ΨdU (see also Figure 14A) and depiction of coupled p-σ^*/σ-π orbital interactions. Adapted from refs 143 and 111.

carbon of ΨdU is distorted from its usual trigonal planar geometry toward a tetrahedral geometry (C1 of ΨdU is equivalent to N1 of dU). The crystallographic group that solved this structure proposed that this unprecedented ground-state destabilization promoted coupled stereoelectronic effects between the sugar and base. In this proposal, weakening of the glycosidic bond of dU in the Michaelis complex results from two stereoelectronic effects: the interaction of the empty p-orbital of O4′ with the σ^* orbital of C1′ (the anomeric effect), and the interaction between the C1′-N1 σ bond with the π electron system of the uracil base (Figure 17). The enzymatic distortion of the glycosidic bond by rotating the base 90° from its normal anti configuration, and bending the plane of the uracil base by 40° from its normal 180° orientation relative to the vector described by the glycosidic bond, is proposed to enhance the latter $\sigma-\pi$ interaction (Figure 17). However, this effect is a double-edged sword because the induced interaction of the C1′-N1 σ bond with the π electron system of the uracil disrupts the normal aromatic delocalization of the lone pair electrons on uracil N1 with the π system. Thus, the coupled stereoelectronic hypothesis "borrows from Peter to pay Paul".

Berti has performed quantum mechanical calculations to address whether such a distortion in the normal deoxyuridine nucleoside leads to the expected changes in bond lengths resulting from coupled anomeric and $\sigma-\pi$ stereoelectronic effects.[143] Although the computations supported the expected O4′−C1′ bond order increase resulting from the anomeric effect, the changes in the uracil ring were inconsistent with the expectations from the coupled anomeric/$\sigma-\pi$ stereoelectronic hypothesis.[143] Consistent with these QM calculations, off-resonance Raman studies of UDG complexed with 2′-fluorodeoxyuridine substrate analogue DNA showed a 34 cm^{-1} decrease in the uracil C2−O2 carbonyl stretch frequency resulting from the strong hydrogen bond to His187 (see above), indicating increased C2−O2 single bond character. However, this very sensitive method provided no evidence for loss of aromaticity of the uracil ring, which is inconsistent with the presence of a highly populated bent state in solution. These computational and experimental results are most simply explained by ground-state destabilization involving the anomeric effect, and conventional resonance stabilization of the developing negative charge on N1 through conjugation of its lone pair electrons to the π system of uracil O2. We conclude that if the distorted conformation is present in solution, it is a high-energy state that is not observable by NMR or Raman spectroscopy. The discrepancies between the solution and crystallographic measurements are not yet resolved.

Although there is no crystal of TDG bound to DNA, the crystal structure of MUG bound to DNA containing the substrate analogue, 2′-fluoro-2′-deoxyuridine reveals some interesting similarities to the UDG complex with ΨdU−DNA.[99] First, the sugar pucker is in a nearly identical 3′ exo conformation, suggesting a similar sugar preorganization strategy as UDG. It is likely that the similar DNA backbone distortions in UDG and MUG are used to drive the sugar into this conformation, thereby coupling the base flipping process to catalysis. In the MUG structure, there is also a rotation of the uracil ring around the glycosidic bond, but this induced change is only about 45° as compared to the 90° rotation in the UDG−ΨdU structure (Figure 14B). Thus, the orientation of the uracil base with respect to the sugar is not optimal for coupled anomeric/$\sigma-\pi$ interactions as proposed for UDG. Given these observations, it would seem that the activation barrier decrease provided by the MUG active site involves ground-state destabilization of the sugar, conformational facilitation of the anomeric effect, stabilization of electron density on the base through the formation of hydrogen bonds between the exocyclic carbonyl groups of the base and main chain amide groups of the enzyme, and, possibly, favorable electrostatic interactions of a dissociative transition state with the DNA phosphodiester backbone (see below).

5. Substrate Autocatalysis: The Contribution from Phosphodiester Electrostatics

A recent QM/MM computational study that calculated the complete reaction surface for deoxyuridine hydrolysis in the active site of UDG concluded that electrostatic interactions between the cationic sugar and the phosphodiester backbone of the DNA substrate provided a stabilizing effect of 17.5−21.9 kcal/mol, and provided essentially all of the energetic basis for UDG catalysis [144] (Figure 14A). Although the importance of these interactions was qualitatively supported by extensive experimental studies, the maximal combined electrostatic contribution from these groups was estimated to be no greater than 6.3−8.3 kcal/mol.[145,146] The experimental estimates were obtained from investigations of UDG's reactivity on a simple deoxyuridine substrate that contained no phosphodiester groups at all, and also using substrates that contained single stereospecific methylphosphonate (MeP) substitutions at the +2, +1, −1, and −2 positions (see Table 4 for nomenclature), thereby removing the negative charges that were implicated to be essential in the computational work. Evidence that the phosphodiester electrostatic interactions were especially important in stabilizing the cationic sugar in the transition state and intermediate was obtained from a plot of log k_{cat}/K_m for the series of MeP substituted substrates against log K_i for binding of an identical DNA that contained a cationic 1-azadeoxyribose transition-state analogue instead of dU (Figure 18). A linear correlation of unit

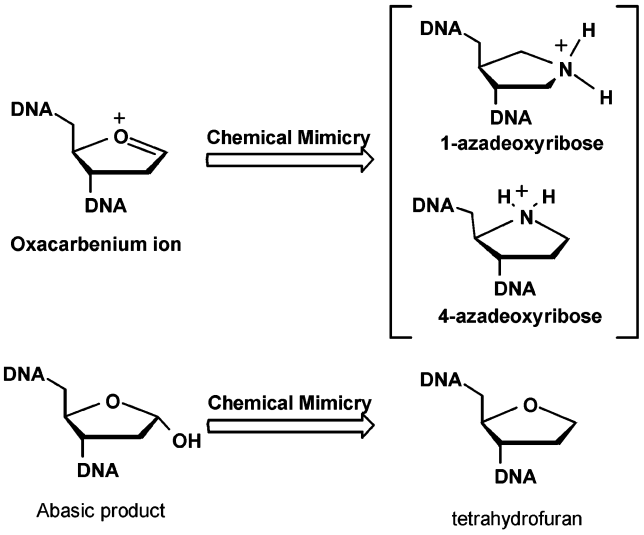


Figure 18. Stable chemical mimics of the glycosyl cation transition state and the abasic product of DNA glycosylase reactions.

slope was obtained, confirming that the electronic features of the transition-state resembled that of the cationic 1-azadeoxyribose inhibitor and that the anionic backbone of DNA was used in transition-state stabilization. In contrast, the correlation between log k_{cat}/K_m and log K_m for the substrate showed a slope of only 0.5, indicating that phosphate interactions had a much lesser effect in ground-state binding. These differential effects in the ground state and transition state are required for an interaction to lower the activation barrier and provide credible evidence that the charged DNA backbone can play a major role in such reactions. It is likely that other glycohydrolases that act on DNA or RNA substrates will use similar catalytic strategies that take advantage of the electronic features of the substrate. Indeed, the 3′ and 5′ DNA phosphodiesters of the cognate nucleotide of MUG are poised even closer to anomeric carbon than the same groups in the UDG active site, suggesting that such a mechanism could easily be operative.

B. A Bifunctional Pyrimidine Specific Glycosylases: PDG

Endonuclease V from T4 phage is the most thoroughly studied bifunctional thymine dimer glycosylase/lyase. Such enzymes have also been isolated from the single-celled algae *Chorella*,[147] yeast, and other eubacteria.[148] The catalytically inactive E23Q enzyme was cocrystallized with thymine dimer-containing DNA in 1995, which has provided the largest reservoir of data for mechanistic interpretations of this enzyme's activity (Figure 14C). This structure and biochemical studies[149−151] have established that the amino terminal threonine serves as the amine nucleophile that attacks the anomeric carbon of the 5′ thymine of the dimer. However, the specific roles of three other active site groups that are observed in the crystal structure remain ambiguous (Arg22, Arg26, and Glu23).[152] Part of the problem in inferring mechanism from the crystal structure arises from the rather large distance between the amine nucleophile and the anomeric center (3.9 Å), and its unusual angle of attack with respect to the departing leaving group. This distance and geometry suggests that an early conformation on the reaction coordinate has been trapped and that further structural rearrangements are required to facilitate the reaction chemistry. In addition, the catalytic Glu23 (which was mutated to glutamine to obtain the structure) is located almost equidistance from thymine O2 and O4′ of the sugar (~2.7 Å), giving rise to multiple proposals for its catalytic role. One proposal is that it serves as a proton donor to thymine O2 in the ground state, which is problematic given the extremely low pK_a of O2,[33] but certainly becomes a viable mechanism in a concerted transition state, where the pK_a of O2 increases significantly (see discussion above for UDG). A second proposal is that Glu23 hydrogen bonds to O4′, thereby increasing the electrophilicity of C1′, which would promote attack of the amino group of Thr2 at C1′ in a associative A_ND_N mechanism.[152] This precarious activation mechanism requires that the enzyme balance the leaving group potential of the N1 nitrogen of thymine and O4′ of the sugar. Both proposals require a high pK_a Glu23, and consistent with this, computational studies predict a large 6 unit increase in the Glu23 pK_a upon DNA binding due to electrostatic effects.[152] A favorable electrostatic role for Arg22 and Arg26 in stabilizing the negatively charged leaving base has also been proposed on the basis of this computational work. It should be pointed out that the thymine

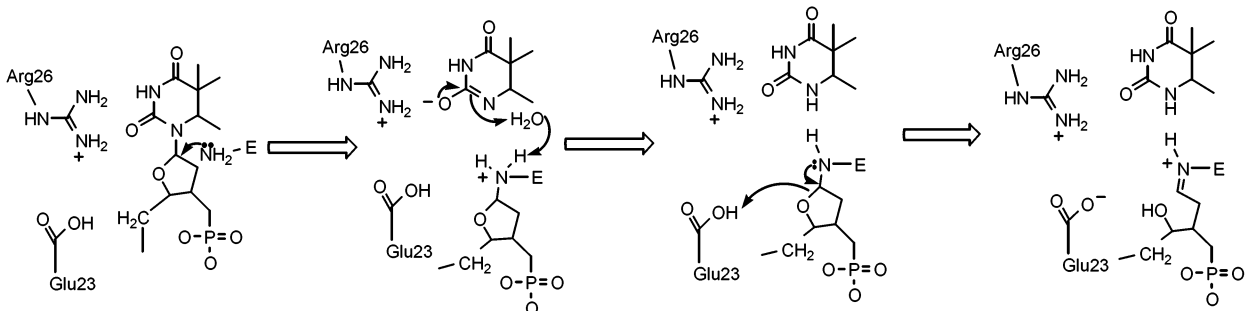


Figure 19. Possible reaction mechanism for the bifunctional glycosylase T4 PDG. This mechanism is adapted from ref 152.

dimer does not possess the aromaticity of the uracil ring due to the reduced C5−C6 double bond, and therefore, electron release into the pyrimidine must be localized on N1 or O2. Electrostatic interactions of these arginines with this localized negative charge may fulfill a similar catalytic role as the strong hydrogen bond between the active site histidine of UDG and uracil O2. One reasonable, but by no means the xclusive mechanism for this enzyme, is depicted in Figure 19.

In summary, if the proposed mechanisms for PDG are correct, then there would appear to be a large change in transition state structure for this enzyme as compared to UDG. It is possible that the greater nucleophilicity of the amine nucleophile as compared to water, and its rigid positioning in the enzyme active site by covalent bonding, may play a role in pushing the transition state to an earlier point on the reaction coordinate than observed with the monofunctional enzymes. These conjectures are far from being established.

C. Purine Specific Glycosylases: MutY, HOGG1, and FPG

1. General

As discussed in the introductory sections on nonenzymatic hydrolysis reactions, enzymatic excision of purine bases is a distinct chemical problem from the analogous pyrimidine reactions due to the different acid−base properties of purine bases. As paradigms for these reactions, we will look at the adenine specific glycosylase, MutY, and two bifunctional glycosylase/lyase enzymes, hOGG1 and FPG, that recognize 8-oxoguanine opposite to cytosine in DNA. Together, these enzymes represent the cellular armament to combat the deleterious effects of the oxidized base, 8-oxoguanine (8-oxoG) in DNA.[122] The biological functions of the human enzyme, hOGG1, and its bacterial counterpart, FPG, are to excise 8-oxoG before the DNA is replicated, as DNA polymerases will readily insert adenine opposite to 8-oxoG using the Hoogsteen hydrogen bonds of syn 8-oxo-dG.[153] The role of MutY is to excise any adenine that has been mistakenly incorporated opposite to 8-oxo-dG during DNA replication, thus providing a second line defense to combat accumulation of this mutagenic lesion. The bacterial MutY and hOGG1 are both helix−hairpin−helix superfamily members,[98,126] but FPG belongs to its own structural family and has recently been cocrystallized with DNA by a number of groups.[100,154,155] As with most DNA repair enzymes, the discussion of mechanism for these examples relies heavily on structural interpretations and mutagenesis experiments.

2. Leaving Group Activation

A crystal structure of MutY bound to DNA has not yet been reported, but a structure of the enzyme bound to the product base adenine has been described.[126] Like many purine nucleoside hydrolases and phosphorylases,[157] MutY shows an acidic group (Glu37) directed at the N7 nitrogen of the adenine ring, the removal of which abolishes enzyme activity.[157] It is interesting that N7 is the site of leaving group protonation by these enzymes, as the N1 position is the easiest to protonate in solution. Instead, MutY forms specificity hydrogen bonds to N6 and N1 of the adenine base utilizing a conserved glutamine residue (Gln182). Thus, the inference from this arrangement is that MutY may protonate the leaving base in a preequilibrium step before nucleophilic attack by water, or before leaving group departure in a D_NA_N mechanism. This type of mechanism would be quite distinct from the pyrimidine enzymes, where full proton transfer does not occur (see above).

There is no structural information of FPG bound to DNA containing an 8-oxoG base, but the structure of a hOGG1 substrate complex has been solved which provides insights into how both of these enzymes may facilitate cleavage of the glycosidic bond of 8-oxoG (Figure 19A).[98] A priori, it might be expected that enzymatic activation of the 8-oxoG base would involve strong hydrogen bonding to the 8-carbonyl group, which would make the base electron deficient and destabilize the glycosidic bond. Surprisingly, the crystal structure shows that the enzyme is completely devoid of interactions with this position, which is the most conspicuous handle to distinguish 8-oxoG from guanine. Instead, the only apparent discriminating interaction with the base is the hydrogen bond between N7−H and the carbonyl group of Gly42. It seems improbable that the exquisite specificity of the enzyme for 8-oxoG can be explained by this single hydrogen bond, and the absence of interactions that would facilitate leaving group departure is an unsatisfying outcome from this structure. As suggested below, it may be that hOGG1 uses hydrogen bonding with the 4′ OH of the ring opened sugar to facilitate leaving group departure.

3. Nucleophile Activation

MutY is a prototypic monofunctional HhH glycosylase[158] and, therefore, possesses a highly conserved aspartate residue (Asp138) that has been implicated in activating the water nucleophile or in stabilization of the cationic sugar in a dissociative transition state.[126,159] The latter role is not supported because the binding affinity of the enzyme for a cationic pyrrolidine abasic site analogue (65 pM)[160] is similar to an uncharged tetrahydrofuran abasic site analogue (45 pM)[161] and also falls in a similar range as a series of stable substrate analogues (50−280 pM).[162,163] The tightest binding inhibitor of MutY yet found is the pyrrolidine analogue in which the adenine base is attached through a methylene linkage (K_D < 1 pM).[160] However, it is not clear whether this compound is a glycosyl cation mimic because the 4′-imino group is expected to have a low pK_a of about 6.5 based on the closely related phenyliminoribitol,[164] and therefore may not be cationic in the enzyme active site.

In contrast with MutY, both hOGG1 and FPG use amine nucleophiles to attack the anomeric carbon. For hOGG1, a lysine is employed, as found in another HhH family member, EndoIII.[165] FPG uses its N-terminal proline residue,[166,167] as found in another closely related enzyme, Endonuclease VIII, which removes oxidized pyrimidines.[168] It is not known whether these amine groups are in the protonated or deprotonated states in the ES complexes, but the neutral forms clearly have to serve as the nucleophiles in these reactions. Although low pK_a lysines are unusual,[169] it is possible that the active site environment of hOGG1 promotes formation of such a species. Moreover, it is highly likely that the N-terminal proline of FPG is in the required neutral protonation state because N-terminal amines have pK_a values near neutrality, and a low pK_a N-terminal proline general base (pK = 6.4) has been previously established for the enzyme 4-oxalocrotonate tautomerase.[170] Supporting the notion of low pK_a values for the nucleophilic groups of hOGG1 and FPG, there are no groups nearby that could reasonably serve as a general bases to deprotonate these amine nucleophiles, although a conserved aspartate (Asp268) of hOGG1 could conceivably fulfill this role if this group were rotated 180° around its χ1 axis (Figure 20A). If so, then Asp268 would be fulfilling the same general base role as proposed for the conserved aspartate groups in other HhH family members such as MutY (see above) and AlkA (see below). As described below, an alternative role for this group as a general acid is also compatible with the current structural and biochemical data.

4. Possible Mechanisms for MutY, HOGG1, and FPG

Given the limited mechanistic and structural data, it is difficult to confidently conclude where on the reaction coordinate the MutY transition state lies. The prevailing evidence, base primarily on structural insights and chemical mimicry of the transition state (see above), would seem to indicate a more associative mechanism involving leaving group activation by protonation at N7 and water activation by a conserved aspartate. Without additional information on the interactions of the enzyme with the sugar and, preferably, KIE measurements, the detailed mechanism will remain ambiguous. It should be pointed out that KIE studies have been performed on the RNA/DNA adenine glycosylase, ricin A chain, which have indicated a highly dissociative mechanism for this functionally related enzyme.[133,171]

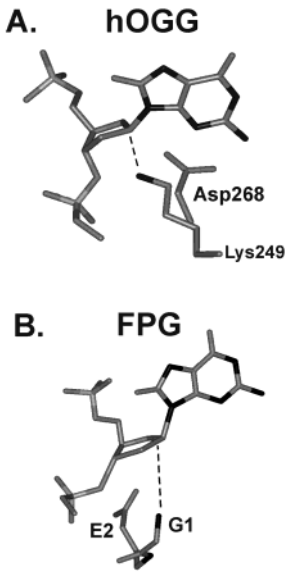


Figure 20. Selected active site interactions of hOGG1 and FPG (PDB accession codes 1EBM and 1KFV, respectively). (A) The residue lys249 of hHOGG1 was modeled into the crystal structure of the K249Q mutant. (B) The sugar and 8-oxoG base were modeled onto the structure of FPG bound to a hydroxypropyl abasic site analogue.

A unified catalytic mechanism for the glycosylase reactions of hOGG1 and FPG can be proposed on the basis of the crystal structures of these enzymes bound to 8-oxo-dG containing DNA and a propyl abasic site analogue, respectively.[98,155] For the crystallographic studies with hOGG1, the catalytic lysine was mutagenized to glutamine to prevent catalytic turnover, and we have modeled the native lysine residue into the crystal structure in the depiction shown in Figure 19A. For FPG, we have reconstructed the full sugar and 8-oxoguanine base from the propyl abasic site analogue used in the crystallization, using the base orientation observed in the hOGG1 structure (Figure 19B). As can be seen in these models, both enzymes have a carboxylic acid residue positioned that would allow protonation of O4′ of the sugar as the proline or lysine nucleophile attacks the anomeric carbon to form an aminal intermediate. In both structures, the observed angle of attack of the amine nucleophile, is highly unfavorable for C−N bond cleavage, suggesting that the first step for both of these enzymes could involve expulsion of the O4′ leaving group facilitated by general acid catalysis, rather than cleavage of the C−N glycosidic bond, as shown in Figure 20A for hOGG1. Thus, expulsion of the 8-oxoG leaving group could occur in the subsequent step of forming the protonated Schiff base intermediate. The catalytic carboxylate could now stabilize the cationic Schiff base intermediate through a favorable electrostatic interaction, providing the thermodynamic driving force for expulsion of the base. An attractive feature

of this mechanism is that cleaving the sugar first allows the proton that is attached to O4′ to stabilize the developing negative charge on the 8-carbonyl group during the glycosidic bond cleavage step, thereby solving the structural enigma of no hOGG1 interactions with the 8-carbonyl group (Figure 19A). This interaction would provide a highly specific mechanism for cleavage of the glycosidic bond of 8-oxoG but not dG. For hOGG1, the Verdine group has proposed two distinct mechanisms involving C−N bond cleavage before C−O bond cleavage (Figure 20B,C),[29,98] and as of yet, the relative merits of these two mechanisms have not been tested.

D. Alkylated Purine Specific Glycosylases: AlkA, AAG, and TAG

1. General

Alkyl purine specific DNA glycosylases represent the most catalytically feeble of the DNA glycosylase enzymes, due to the large k_{non} values for the nonenzymatic reactions (Figure 2). Thus, the enzymatic mechanisms for recognition and removal of lesions with such unusual reactivity and electronic structure would be expected to differ considerably from the previous examples with neutral bases. The three alkyl purine specific enzymes discussed here represent the currently known spectrum of enzymatic solutions to these problems. AlkA and TAG both belong to the helix−hairpin−helix superfamily of DNA glycosylases, and thus share a similar fold, but as detailed below, distinct specificities and catalytic mechanisms. The human counterpart, AAG, has an entirely different fold but shares the broad substrate specificity and mechanism of AlkA (Table 2).

2. Leaving Group Interactions: The Aromatic π-Cation Hypothesis

As first pointed out by Lindahl, enzymatic removal of alkylated bases need not involve profound mechanisms for leaving group activation, as these bases are electron-deficient and prone to spontaneous hydrolysis at comparatively rapid rates.[172] For AlkA and AAG, this prediction appears to be verified, as the active site pockets of these enzymes possess no obvious polar groups that could form strong hydrogen bonds to the electronegative acceptor groups on the base. They also possess no specialized binding pockets for the alkyl modification, presumably because such a diverse array of damaged base are removed by these enzymes.[110] In fact, the active sites are lined with conserved tryptophan and tyrosine residues that form stacking and edgewise interactions with the damaged cationic base. On the basis of the aromatic character of the active sites, it has been proposed that these enzymes use aromatic π-cation interactions to attract the cationic damaged base into the active site, and thus discriminate cationic damaged purines from neutral undamaged purines.[109] Such interactions can be quite strong, and computational studies investigating similar Arg−Trp interactions in many proteins have found an average strength of about −3 kcal/mol,[173] which if true for alkyl base recognition, might lead to significant specific binding energy for a cationic base.[174] However, it should be kept in mind that there exists no structure of a cationic nucleotide bound to any of these enzymes, partly because of the difficulty in synthesis of stable analogues for these substrates. Thus, the structural information is not strong and relies on chemical intuition and modeling of alkylated bases into a crystallographic model. Nevertheless, modeling 3-methyladenine into the crystal structure of AlkA bound to DNA containing the transition state mimic, 1-azadeoxyribose, suggested that a tryptophan residue stacked with the cationic base (Figure 21A).[159] Similarly, the structure of a catalytically neutered mutant of AAG with DNA containing the neutral substrate, 1,N^6-ethenoadenine, showed a stacking interaction between a tyrosine and histidine side chain of the enzyme and the base (Figure 21B).[175] It should be kept in mind that true π−cation interactions require partial charge transfer from the HOMO of the aromatic ring to the LUMO of the base. Such interactions will be most energetically favorable when there is a large overlap area between the donor and acceptor, close interplanar spacing (<3.4 Å), and an almost perfect parallel stacking arrangement.[176] Since there are no structures with cationic bases bound in the active sites of these enzymes, it is impossible to know if these requirements are met.

Can the π−cation interaction provide a large amount of binding specificity? If we define specificity as the ratio of the binding constant for the cationic alkylated base as compared to the normal neutral base and we assume that there is no special recognition of the alkyl modification as suggested by the available structures, then specificity would in part depend on the relative energetics of the π−cation versus the $\pi-\pi$ stacking interactions for these bases. In fact, all DNA glycosylases that recognize neutral cognate bases show an aromatic stacking interaction between an active site group and the base, which suggests a favorable $\pi-\pi$ stacking energy in such cases. Although the π−cation interaction with an alkylated base may be stronger than the $\pi-\pi$ interaction with a neutral base of identical structure, this interaction would not appear to lend itself to exquisite discrimination between neutral and cationic bases. Consistent with this, AAG binds tightly to DNA containing the neutral bases, ϵA and hypoxanthine, as well as cationic bases such as 7-methylguanine and 3-methyladenine.[30,106]

An alternative mechanistic view that has not been considered to our knowledge is that the nonpolar active sites of these alkylated base specific enzymes provide an environment that promotes ground-state electrostatic strain, thereby selectively lowering the activation barrier for glycosidic bond cleavage of cationic nucleobases. In this mechanism, substrate binding energy is used to drive the cationic damaged base into an active site of low dielectric, where it is then destabilized. As the glycosidic bond is broken and the electrons are released into the aromatic system of the base, the electrostatic strain decreases and the departing base, which is now neutral, binds more tightly as the transition state is approached. Thus, the nonpolar environment lowers the activation

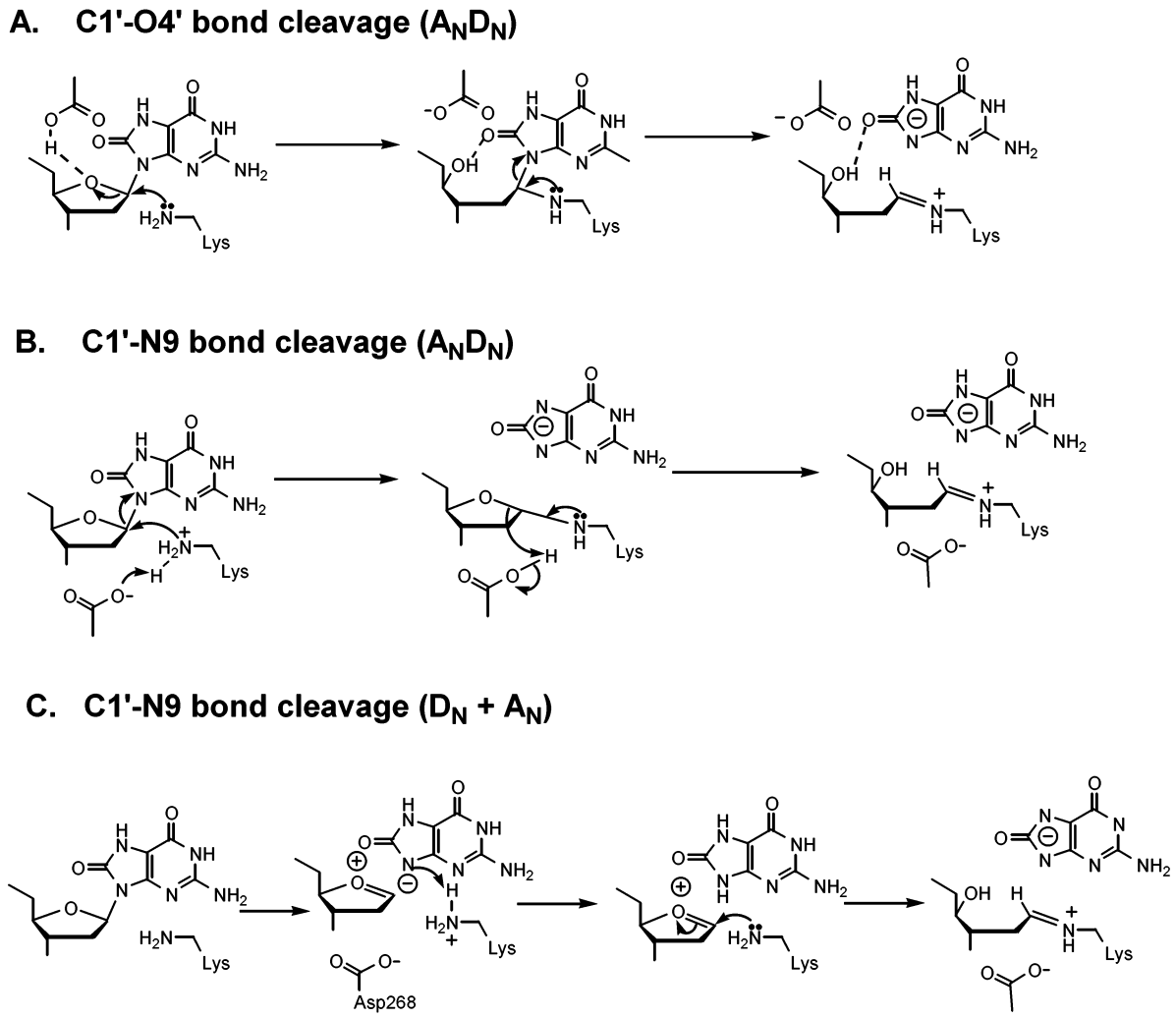


Figure 21. Three possible reaction mechanisms for hOGG1 involving (A) C1′−O4′ bond cleavage before glycosidic bond cleavage utilizing Asp268 as a general acid, (B) C1′−N9 bond cleavage before sugar ring opening utilizing Asp268 as a general base, and (C) C1′−N9 bond cleavage by a $D_N + A_N$ mechanism utilizing Asp268 as an electrostatic catalyst.

barrier by straining the charged ground state and binding tightly to the neutral base in a dissociative transition state. This would provide a plausible mechanism to promote specificity at the catalytic step because neutral bases would bind tightly in the ground state, and more weakly as the anionic transition state is approached, which is anticatalytic. More mechanistic studies need to be performed before these alternative proposals can be confirmed or dismissed, but given the small catalytic powers of these enzymes, such a ground state strain mechanism could easily account for the observed rates and specificities.

A new example of alkylated base recognition and cleavage has been revealed from the solution structure and 3-methyladenine binding studies of TAG, the second enzyme in *E. coli* that removes 3-methyladenine from DNA.[177,178] Although it only shows 10% overall sequence identity with AlkA, the structure revealed that TAG shared the helix−hairpin−helix structural motif that is the hallmark of this DNA repair superfamily, and contained a novel structural zinc binding site that tethered the amino and carboxyl termini together.[178] (At almost the same time, a computational study also proposed that TAG was a member of this superfamily.).[179] From monitoring backbone amide chemical shift changes upon 3-methyladenine binding, the damaged base binding pocket was assigned to an extremely aromatic rich pocket, reminiscent of AAG and AlkA.[177] However, some interesting differences were observed. First, TAG has a conserved glutamate residue that is also found in the adenine binding pocket of MutY. In the crystal structure of MutY bound to adenine,[126] this group forms hydrogen bonds with the N6 and N7 positions of the base, suggesting that TAG may using a similar recognition strategy for 3-methyl adenine. [However, this group is also close enough to the sugar binding site that it may serve as a general base to activate the attacking water (see below).] Assuming the same role as in MutY, this group could lower the activation barrier by donating a proton or hydrogen bond to N7. Second, it was found that the neutral 3-methyladenine base bound specifically and tightly to TAG by an induced fit mechanism (K_D = 60 μM), whereas binding of the normal base adenine could not be detected.[177] This result reflects on the π−cation hypothesis because it directly demonstrates that the TAG active site can specifically recognize a 3-methyladenine base even when it has no positive charge.

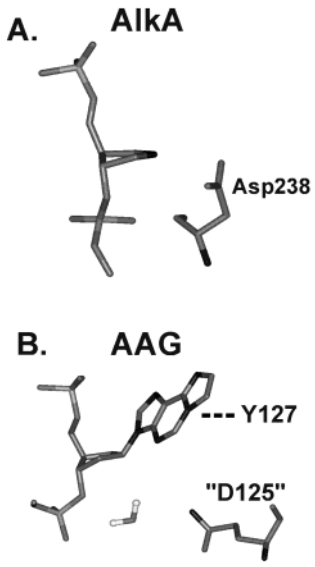


Figure 22. Selected active site interactions of the alkylated purine specific enzymes, AlkA and AAG (PDB accession codes 1DIZ and 1EWN, respectively). (A) The complex of AlkA with 1-azadeoxyribose containing duplex DNA. (B) The complex of AAG with ϵA containing DNA.

Clearer illumination of the recognition mode for this enzyme awaits further structural and mechanistic studies.

3. Water Activation and Stabilization of the Cationic Sugar

The crystal structures of the AlkA and AAG complexes with DNA suggest different transition state structures for these two enzymes. Potential mechanisms for stabilization of an oxacarbenium-like transition state are indicated for AlkA, while for AAG, a mechanism for activating a water nucleophile is indicated. For AlkA, the 1-azadeoxyribose transition state analogue is found in a 3′-exo sugar pucker, with a conserved aspartate side chain poised beneath the sugar (Figure 21A).[159] This arrangement is remarkably similar to the UDG complex with ψdU (see above) and suggests that AlkA may also use substrate binding energy to strain the sugar into a conformation that is favorable for oxacarbenium ion formation. In addition, the 3′ phosphodiester of the 1-azadeoxyribose is located 3.8 Å from the anomeric position which could provide electrostatic stabilization of a glycosyl cation. Diverging from the UDG mechanism, there is no water molecule visible in the electron density, and indeed, the catalytic aspartate is positioned only 3.2 Å from the anomeric center, which precludes positioning of an intervening water molecule. The arrangement of the AlkA active site seems architecturally consistent with a D_N*A_N type of mechanism, although structures of high-energy transition states cannot be reliably predicted from stable ground-state complexes. The AAG structure with ethenoadenine (ϵA) containing DNA shows a near identical sugar pucker at the target base as AlkA (Figure 21B),[175] but strong electrostatic stabilization mechanisms do not appear to be in place: the catalytic glutamate of AAG and both the 3′ and 5′ phosphodiester groups of ϵA are located greater than 4.5 Å away from the anomeric center. Finally, there is strong electron density for a water nucleophile wedged between the catalytic glutamate and the anomeric carbon, which may suggest a more associative A_ND_N type mechanism for the AAG catalyzed cleavage of ϵA.

The NMR structure of TAG presented a mechanistic surprise with respect to water activation and cationic sugar stabilization by this enzyme. Previous to this structure, it was believed that all HhH superfamily glycosylases possessed a conserved aspartate residue that served to deprotonate water or stabilize the positive charge that develops on the sugar.[158] On the contrary, TAG has no aspartate group in the vicinity of the conserved aspartate observed in the structures of other HhH members.[177] This is an intriguing observation that has not been explained to date.

4. Envisioning the Transition State for Alkyl Purine Glycosylases

A hint at the transition state structure for AlkA has been provided through the use of chemical mimics of the oxacarbenium ion.[161,181] These mimics contain cationic nitrogen atoms at the C1′ or O4′ positions of the sugar which mimic the delocalized positive charge on C1′−O4′ in a dissociative transition state. Such analogues have been widely employed to inhibit various glycosidases that act on pyranose or furanose sugars,[181,182] and the Verdine group extended this design concept to DNA glycosylases by incorporation of the 4-azadeoxyribose pyrrolidine analogue into DNA using standard phosphoramidite chemistry (Figure 18).[160] Ichikawa later synthesized the 1-azadeoxyribose analogue (Figure 18),[181] which was also incorporated into DNA and tested as a transition-state mimic of AlkA (as well as UDG).[159,183] Both oxacarbenium analogues were found to bind to AlkA with picomolar affinities, while a uncharged tetrahydrofuran abasic analogue bound 2800 times more weakly (Figure 18).[161] These findings are consistent with the hypothesis that AlkA follows a reaction course that includes a significant amount of charge development on the sugar ring.

VII. Concluding Remarks

The importance of DNA glycosylases in biology and medicine continues to expand as our understanding of their functions increases. New discoveries are constantly uncovering new and surprising biological roles for these enzymes beyond their classical function of combating genetic mutations. Some recent examples are the central role of UDG in generating somatic hypermutation during development of the antibody response to antigen,[184] its role in the cell killing mechanism of antifolate drugs,[185] and the emerging realization that UDG plays an essential role in the life cycles of several viruses that affect human health.[186] There is also a growing appreciation that progression of cancer involves mutations in DNA repair genes,[187,188] thereby increasing the genetic plasticity of tumor cells, which may play a role in allowing these cells to rapidly develop resistance to

chemotherapeutic agents. There is considerable interest in using DNA repair proteins in gene therapy applications to enhance the effectiveness of cancer chemotherapeutic regimes.[189,190] We anticipate that further diverse roles and applications for these enzymes will emerge in the years that follow.

What mechanistic questions remain for DNA glycosylases? The answer to this question is straightforward—many. However, the age of structural biology is winding down for these enzymes, and the remaining important questions will require methods that are capable of linking the structural observations to binding energy—the basic currency that all biological molecules use to perform useful work. New methods should be employed to understand the site searching mode for these enzymes in protein-coated cellular DNA, and the tools of spectroscopy and kinetic isotope effects should be applied to other representative glycosylases so that the chemical mechanisms for enzymatic cleavage of the complete spectrum of damaged bases can be ascertained. With such information in hand, we predict small molecules that modulate the activities of these enzymes will emerge. Such molecules could help unravel the roles of DNA repair pathways in apoptosis, cancer progression, and anticancer drug efficacy.

VIII. Acknowledgments

We thank Daniel J. Krosky for his careful reading of this manuscript and for helpful suggestions.

IX. Abbreviations

AAG	human alkyl purine DNA glycosylase
AlkA	*E. coli* alkyl purine DNA glycosylase II
2-AP	2-aminopurine
KIE	kinetic isotope effect
MUG	*E. coli* mismatch uracil DNA glycosylase
8-oxoG	8-oxoguanine
PDG	pyrimidine dimer DNA glycosylase
TAG	*E. coli* 3-methyladenine DNA glycosylase I
T4 PDG	pyrimidine dimer DNA glycosylase from T4 phage
TDG	human thymine DNA glycosylase
UDG	uracil DNA glycosylase

X. References

(1) Seeberg, E.; Eide, L.; Bjoras, M. *Trends Biochem. Sci.* **1995**, *20*, 391.
(2) Mol, C. D.; Parikh, S. S.; Putnam, C. D.; Lo, T. P.; Tainer, J. A. *Annu. Rev. Biophys. Biomol. Struct.* 1999, *28*, 101.
(3) Lindahl, T.; Andersson, A. *Biochemistry* **1972**, *11*, 3618.
(4) Cuniasse, P.; Fazakerley, G. V.; Guschlbauer, W.; Kaplan, B. E.; Sowers, L. C. *J. Mol. Biol.* **1990**, *213*, 303.
(5) Lindahl, T. *Mutat. Res.* **1990**, *238*, 305.
(6) Barzilay, G.; Mol, C. D.; Robson, C. N.; Walker, L. J.; Cunningham, R. P.; Tainer, J. A.; Hickson, I. D. *Nat. Struct. Biol.* **1995**, *2*, 561.
(7) Mol, C. D.; Izumi, T.; Mitra, S.; Tainer, J. A. *Nature* **2000**, *403*, 451.
(8) Kubota, Y.; Nash, R. A.; Klungland, A.; Schar, P.; Barnes, D. E.; Lindahl, T. *EMBO J.* **1996**, *15*, 6662.
(9) Mullen, G. P.; Wilson, S. H. *Biochemistry* **1997**, *36*, 4713.
(10) Sun, B.; Latham, K. A.; Dodson, M. L.; Lloyd, R. S. *J. Biol. Chem.* **1995**, *270*, 19501.
(11) Piersen, C. E.; McCullough, A. K.; Lloyd, R. S. *Mutat. Res.* **2000**, *459*, 43.
(12) Klungland, A.; Lindahl, T. *EMBO J.* **1997**, *16*, 3341.
(13) Pascucci, B.; Stucki, M.; Jonsson, Z. O.; Dogliotti, E.; Hubscher, U. *J. Biol. Chem.* **1999**, *274*, 33696.
(14) Parikh, S. S.; Mol, C. D.; Hosfield, D. J.; Tainer, J. A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 37.
(15) Lloyd, R. S. *Prog. Nucleic Acid Res. Mol. Biol.* **1999**, *62*, 155.
(16) David, S. S.; Williams, S. D. *Chem. Rev.* **1998**, *98*, 1221.
(17) Wyatt, M. D.; Allan, J. M.; Lau, A. Y.; Ellenberger, T. E.; Samson, L. D. *Bioessays* **1999**, *21*, 668.
(18) Lindahl, T.; Wood, R. D. *Science* **1999**, *286*, 1897.
(19) Gogos, A.; Jantz, D.; Senturker, S.; Richardson, D.; Dizdaroglu, M.; Clarke, N. D. *Proteins* **2000**, *40*, 98.
(20) Aravind, L.; Walker, D. R.; Koonin, E. V. *Nucleic Acids Res.* **1999**, *27*, 1223.
(21) Oivanen, M.; Hovinen, J.; Lehikoinen, P.; Lonnberg, H. *Trends Org. Chem.* **1993**, 4, 397.
(22) Wolfenden, R.; Snider, M. J. *Acc. Chem. Res.* **2001**, *34*, 938.
(23) Shapiro, R.; Danzig, M. *Biochemistry* **1972**, *11*, 23.
(24) Shapiro, R.; Kang, S. *Biochemistry* **1969**, *8*, 1806−10.
(25) Zoltewicz, J. A.; Clark, F. D.; Sharpless, T. W.; Grahe, G. *J. Am. Chem. Soc.* **1970**, *92*, 1741.
(26) Chernikov, A. V.; Usacheva, A. M.; Bruskov, V. I. *Biochemistry (Moscow)* **1996**, *61*, 35.
(27) Noll, D. M.; Gogos, A.; Granek, J. A.; Clarke, N. D. *Biochemistry* **1999**, *38*, 6374.
(28) Waters, T. R.; Swann, P. F. *J. Biol. Chem.* **1998**, *273*, 20007.
(29) Norman, D. P. G.; Chung, S. J.; Verdine, G. L. *Biochemistry* **2003**, *42*, 1564.
(30) Asaeda, A.; Ide, H.; Asagoshi, K.; Matsuyama, S.; Tano, K.; Murakami, A.; Takamori, Y.; Kubo, K. *Biochemistry* **2000**, *39*, 1959.
(31) Riazuddin, S.; Lindahl, T. *Biochemistry* **1978**, *17*, 2110.
(32) Fujii, T.; Saito, T.; Nakasaka, T. *Chem Pharm. Bull.* **1989**, *37*, 2601.
(33) Poulter, C. D.; Frederick, G. D. *Tetrahedron Lett* **1975**, *26*, 2171.
(34) Fujii, T.; Itaya, T. *Heterocycles* **1998**, *48*, 1673.
(35) Briggs, A. J.; Evans, C. M.; Glenn, R.; Kirby, A. J. *J. Chem. Soc., Perkin Trans. 2* **1983**, *11*, 1637.
(36) Sinnott, M. L. *Adv. Phys. Org. Chem.* **1988**, *24*, 113.
(37) Gorenstein, D. G. *Chem. Rev.* **1987**, *87*, 1047.
(38) In *The Anomeric Effect and Associated Stereoelctronic Effects*; Thatcher, G. R. J., Ed.; American Chemical Society: Washington, DC, 1993; Vol. 539.
(39) Sinnott, M. L. In *The Anomeric Effect and Associated Stereoelectronic Effects*; Thatcher, G. R. J., Ed.; American Chemical Society: Washington, DC, 1993; Vol. 539.
(40) Zhu, J., Bennet, A. J. *J. Am. Chem. Soc.* **1998**, *120*, 3887.
(41) Banait, N. S.; Jencks, W. P. *J. Am. Chem. Soc.* **1991**, *113*, 7951.
(42) Foloppe, N.; MacKerell, A. D. J. *J. Phys. Chem. B* **1998**, *102*, 6669.
(43) Jencks, W. P. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1975**, *43*, 219.
(44) Vasella, A.; Davies, G. J.; Bohm, M. *Curr. Opin. Struct. Biol.* **2002**, *6*, 619.
(45) Fedorov, A.; Shi, W.; Kicska, G.; Fedorov, E.; Tyler, P. C.; Furneaux, R. H.; Hanson, J. C.; Gainsford, G. J.; Larese, J. Z.; Schramm, V. L.; Almo, S. C. *Biochemistry* **2001**, *40*, 853.
(46) Forcasiu, D.; Horsley, J. *J. Am. Chem. Soc.* **1980**, *102*, 4906.
(47) Hehre, W. J. *Acc. Chem. Res.* **1975**, *8*, 369.
(48) Sunko, D. E.; Szele, I.; Hehre, W. J. *J. Am. Chem. Soc.* **1977**, *99*, 5000.
(49) Ashwell, M., Guo, X., Sinnott, M. L. *J. Am. Chem. Soc.* **1992**, *114*, 10158.
(50) Schramm, V. L. *Curr. Opin. Struct. Biol.* **2001**, *5*, 556.
(51) Banait, N. S.; Jencks, W. P. *J. Am. Chem. Soc.*, **1991**; *113*, 7951.
(52) Huang, X. C., Surry, C., Hiebert, T., Bennet, A. J. *J. Am. Chem. Soc.* **1995**, *117*, 10614.
(53) Guthrie, R. D.; Jencks, W. P. *Accs. Chem. Res.* **1989**, *22*, 343.
(54) Schramm, V. L.; Shi, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 657.
(55) Schramm, V. L. *Methods Enzymol.* **1999**, *308*, 301.
(56) Berti, P. J. *Methods Enzymol.* **1999**, *308*, 355.
(57) Parkin, D. W.; Leung, H. B.; Schramm, V. L. *J. Biol. Chem.* **1984**, *259*, 9411.
(58) Mentch, F.; Parkin, D. W.; Schramm, V. L. *Biochemistry* **1987**, 26, 921.
(59) Berti, P. J.; Schramm, V. L. *J. Am. Chem. Soc.* **1997**, *119*, 12069.
(60) Prior, J. J.; Santi, D. V. *J. Biol. Chem.* **1984**, *259*, 2429.
(61) Schaleger, L. L.; Long, F. A. *Adv. Phys. Org. Chem.* **1963**, *1*, 1.
(62) Bruice, T. C.; Benkovic, S. J. *Biochemistry* **2000**, *39*, 6267.
(63) Lindahl, T.; Nyberg, B. *Biochemistry* **1974**, *13*, 3405.
(64) Berg, O. G.; Winter, R. B.; von Hippel, P. H. *Biochemistry* **1981**, *20*, 6929.
(65) von Hippel, P. H.; Berg, O. G. *J. Biol. Chem.* **1989**, *264*, 675.
(66) Selzer, T.; Albeck, S.; Schreiber, G. *Nat. Struct. Biol.* **2000**, *7*, 537.
(67) Stivers, J. T.; Pankiewicz, K. W.; Watanabe, K. A. *Biochemistry* **1999**, *38*, 952.
(68) Drohat, A. C.; Jagadeesh, J.; Ferguson, E.; Stivers, J. T. *Biochemistry* **1999**, *38*, 11866.
(69) Cappelli, E.; Hazra, T.; Hill, J. W.; Slupphaug, G.; Bogliolo, M.; Frosina, G. *Carcinogenesis* **2001**, *22*, 387.

(70) Stanford, N. P.; Szczelkun, M. D.; Marko, J. F.; Halford, S. E. *EMBO J.* **2000**, *19*, 6546.
(71) Ishiwata, K.; Oikawa, A. *Biochim. Biophys. Acta* **1982**, *698*, 15.
(72) Hildebrandt, E. R.; Cozzarelli, N. R. *Cell* **1995**, *81*, 331.
(73) Jeltsch, A.; Wenz, C.; Stahl, F.; Pingoud, A. *EMBO J.* **1996**, *15*, 5104.
(74) Jeltsch, A.; Pingoud, A. *Biochemistry* **1998**, *37*, 2160.
(75) Wong, I.; Lundquist, A. J.; Bernards, A. S.; Mosbaugh, D. W. *J. Biol. Chem.* **2002**, *20*, 20.
(76) Jiang, Y. L.; Stivers, J. T. *Biochemistry* **2002**, *41*, 11236.
(77) Bennett, S. E.; Sanderson, R. J.; Mosbaugh, D. W. *Biochemistry* **1995**, *34*, 6109.
(78) Purmal, A. A.; Lampman, G. W.; Pourmal, E. I.; Melamede, R. J.; Wallace, S. S.; Kow, Y. W. *J. Biol. Chem.* **1994**, *269*, 22046.
(79) Higley, M.; Lloyd, R. S. *Mutat. Res.* **1993**, *294*, 109.
(80) Lloyd, R. S. *Methods Mol. Biol.* **2001**, *160*, 3.
(81) Francis, A. W.; David, S. S. *Biochemistry* **2003**, *42*, 801.
(82) Friedberg, E. C. *DNA Repair*; W. H. Freeman: New York, 1985.
(83) Nilsen, H.; Rosewell, I.; Robins, P.; Skjelbred, C. F.; Andersen, S.; Slupphaug, G.; Daly, G.; Krokan, H. E.; Lindahl, T.; Barnes, D. E. *Mol. Cell.* **2000**, *5*, 1059.
(84) Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A. *Nature* **1996**, *384*, 87.
(85) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357.
(86) Gueron, M.; Leroy, J.-L. *Methods Enzymol.* **1995**, *261*, 383.
(87) Gueron, M.; Kochoyan, M.; Leroy, J. L. *Nature* **1987**, *328*, 89.
(88) Moe, J. G.; Russu, I. M. *Biochemistry* **1992**, *31*, 8421.
(89) Moe, J. G.; Russu, I. M. *Nucleic Acids Res.* **1990**, *18*, 821.
(90) Dornberger, U.; Leijon, M.; Fritzsche, H. *J. Biol. Chem.* **1999**, *274*, 6957.
(91) Banavali, N. K.; MacKerell, A. D., Jr. *J. Mol. Biol.* **2002**, *319*, 141.
(92) Ramstein, J.; Lavery, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7231.
(93) Fuxreiter, M.; Luo, N.; Jedlovszky, P.; Simon, I.; Osman, R. *J. Mol. Biol.* **2002**, *323*, 823.
(94) Seibert, E.; Ross, J. B.; Osman, R. *Biochemistry* **2002**, *41*, 10976−84.
(95) van Aalten, D. M.; Erlanson, D. A.; Verdine, G. L.; Joshua-Tor, L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11809.
(96) Parikh, S. S.; Walcher, G.; Jones, G. D.; Slupphaug, G.; Krokan, H. E.; Blackburn, G. M.; Tainer, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5083.
(97) Vassylyev, D. G.; Kashiwagi, T.; Mikami, Y.; Ariyoshi, M.; Iwai, S.; Ohtsuka, E.; Morikawa, K. *Cell* **1995**, *83*, 773.
(98) Bruner, S. D.; Norman, D. P.; Verdine, G. L. *Nature* **2000**, *403*, 859.
(99) Barrett, T. E.; Scharer, O. D.; Savva, R.; Brown, T.; Jiricny, J.; Verdine, G. L.; Pearl, L. H. *EMBO J.* **1999**, *18*, 6599.
(100) Fromme, J. C.; Verdine, G. L. *Nat. Struct. Biol.* **2002**, *9*, 544.
(101) Chen, L.; Haushalter, K. A.; Lieber, C. M.; Verdine, G. L. *Chem. Biol.* **2002**, *9*, 345.
(102) Jiang, Y. L.; Kwon, K.; Stivers, J. T. *J. Biol. Chem.* **2001**, *276*, 42347.
(103) Handa, P.; Roy, S.; Varshney, U. *J. Biol. Chem.* **2001**, *276*, 17324.
(104) Huang, N.; Banavali, N. K.; MacKerell, A. D., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 68.
(105) Jiang, Y. L.; Song, F.; Stivers, J. T. *Biochemistry* **2002**, *41*, 11248.
(106) Abner, C. W.; Lau, A. Y.; Ellenberger, T.; Bloom, L. B. *J. Biol. Chem.* **2001**, *276*, 13379.
(107) Connor, E. E.; Wyatt, M. D. *Chem. Biol.* **2002**, *9*, 1033.
(108) Kavli, B.; Slupphaug, G.; Mol, C. D.; Arvai, A. S.; Peterson, S. B.; Tainer, J. A.; Krokan, H. E. *EMBO J.* **1996**, *15*, 3442.
(109) Labahn, J.; Scharer, O. D.; Long, A.; Ezaz-Nikpay, K.; Verdine, G. L.; Ellenberger, T. E. *Cell* **1996**, *86*, 321.
(110) Berdal, K. G.; Johansen, R. F.; Seeberg, E. *EMBO J.* **1998**, *17*, 363.
(111) Werner, R. M.; Jiang, Y. L.; Gordley, R. G.; Jagadeesh, G. J.; Ladner, J. E.; Xiao, G.; Tordova, M.; Gilliland, G. L.; Stivers, J. T. *Biochemistry* **2000**, *39*, 12585.
(112) Drohat, A. C.; Stivers, J. T. *Biochemistry* **2000**, *39*, 11865.
(113) Stivers, J. T. *Nucleic Acids Res.* **1998**, *26*, 3837.
(114) Allan, B. W.; Beechem, J. M.; Lindstrom, W. M.; Reich, N. O. *J. Biol. Chem.* **1998**, *273*, 2368.
(115) Allan, B. W.; Reich, N. O. *Biochemistry* **1996**, *35*, 14757.
(116) Nordlund, T. M.; Andersson, S.; Nilsson, L.; Rigler, R.; Graslund, A.; McLaughlin, L. W. *Biochemistry* **1989**, *28*, 9095.
(117) Nordlund, T. M.; Xu, D.; Evans, K. O. *Biochemistry* **1993**, *32*, 12090.
(118) Xu, D.; Evans, K. O.; Nordlund, T. M. *Biochemistry* **1994**, *33*, 9592.
(119) Parikh, S. S.; Mol, C. D.; Slupphaug, G.; Bharati, S.; Krokan, H. E.; Tainer, J. A. *EMBO. J.* **1998**, *17*, 5214.
(120) McCullough, A. K.; Scharer, O.; Verdine, G. L.; Lloyd, R. S. *J. Biol. Chem.* **1996**, *271*, 32147.
(121) McCullough, A. K.; Dodson, M. L.; Scharer, O. D.; Lloyd, R. S. *J. Biol. Chem.* **1997**, *272*, 27210.
(122) Michaels, M. L.; Miller, J. H. *J Bacteriol* **1992**, *174*, 6321.
(123) Michaels, M. L.; Cruz, C.; Grollman, A. P.; Miller, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7022.
(124) Michaels, M. L.; Tchou, J.; Grollman, A. P.; Miller, J. H. *Biochemistry* **1992**, *31*, 10964.
(125) Grollman, A. P. *Struct. Funct., Proc. Conversation Discip. Biomol. Stereodyn., 7th* **1992**, *1*, 165.
(126) Guan, Y.; Manuel, R. C.; Arvai, A. S.; Parikh, S. S.; Mol, C. D.; Miller, J. H.; Lloyd, S.; Tainer, J. A. *Nat. Struct. Biol.* **1998**, *5*, 1058.
(127) Abeygunawardana, C.; Weber, D. J.; Gittis, A. G.; Frick, D. N.; Lin, J.; Miller, A. F.; Bessman, M. J.; Mildvan, A. S. *Biochemistry* **1995**, *34*, 14997.
(128) Volk, D. E.; House, P. G.; Thiviyanathan, V.; Luxon, B. A.; Zhang, S.; Lloyd, R. S.; Gorenstein, D. G. *Biochemistry* **2000**, *39*, 7331.
(129) Bernards, A. S.; Miller, J. K.; Bao, K. K.; Wong, I. *J. Biol. Chem.* **2002**, *8*, 6.
(130) Wong, I.; Bernards, A. S.; Miller, J. K.; Wirz, J. A. *J. Biol. Chem.* **2003**, *278*, 2411.
(131) Werner, R. M.; Stivers, J. T. *Biochemistry* **2000**, *39*, 14054.
(132) Gilles, A.-M.; Cristea, I.; Palibroda, N.; Hilden, I.; Jensen, K. F.; Sarfati, R. S.; Namane, A.; Ughetto-Monfrin, J.; Barzu, O. *Anal. Biochemistry* **1995**, *34*, 197.
(133) Chen, X.-Y.; Berti, P. J.; Schramm, V. L. *J. Am. Chem. Soc.* **2000**, *122*, 6527.
(134) Drohat, A. C.; Xiao, G.; Tordova, M.; Jagadeesh, J.; Pankiewicz, K. W.; Watanabe, K. A.; Gilliland, G. L.; Stivers, J. T. *Biochemistry* **1999**, *38*, 11876.
(135) Drohat, A. C.; Stivers, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 1840.
(136) Shan, S. O.; Loh, S.; Herschlag, D. *Science* **1996**, *272*, 97.
(137) Shan, S. O.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14474.
(138) Lodi, P. J.; Knowles, J. R. *Biochemistry* **1991**, *30*, 6948.
(139) Jiang, Y. L.; Drohat, A. C.; Ichikawa, Y.; Stivers, J. T. *Biochemistry* **2002**, *41*, 7116.
(140) Jiang, Y. L.; Drohat, A. C.; Ichikawa, Y.; Stivers, J. T. *J. Biol. Chem.* **2002**, *277*, 15385.
(141) Waters, T. R.; Gallinari, P.; Jiricny, J.; Swann, P. F. *J. Biol. Chem.* **1999**, *274*, 67.
(142) Sartori, A. A.; Fitz-Gibbon, S.; Yang, H.; Miller, J. H.; Jiricny, J. *EMBO J.* **2002**, *21*, 3182.
(143) Berti, P. J.; Tanaka, K. S. E. *Adv. Phys. Org. Chem.* **2002**, *36*, 239.
(144) Dinner, A. R.; Blackburn, G. M.; Karplus, M. *Nature* **2001**, *413*, 752.
(145) Jiang, Y. L.; Ichikawa, Y.; Song, F.; Stivers, J. T. *Biochemistry* **2003**, *42*, 1922.
(146) Jiang, Y. L.; Stivers, J. T. *Biochemistry* **2001**, *40*, 7710.
(147) Furuta, M.; Schrader, J. O.; Schrader, H. S.; Kokjohn, T. A.; Nyaga, S.; McCullough, A. K.; Lloyd, R. S.; Burbank, D. E.; Landstein, D.; Lane, L.; Van Etten, J. L. *Appl. Environ. Microbiol.* **1997**, *63*, 1551.
(148) Lloyd, R. S. *Mutat. Res.* **1998**, *408*, 159.
(149) Dodson, M. L.; Schrock, R. D., 3rd; Lloyd, R. S. *Biochemistry* **1993**, *32*, 8284.
(150) Schrock, R. D., 3rd; Lloyd, R. S. *J. Biol. Chem.* **1993**, *268*, 880.
(151) Mazumder, A., Gerlt, J. A., Rabow, L., Absalon, M. J., Stubbe, J.; Bolton, P. H *J. Am. Chem. Soc.* **1989**, *111*, 8029.
(152) Fuxreiter, M.; Warshel, A.; Osman, R. *Biochemistry* **1999**, *38*, 9577.
(153) Grollman, A. P. *Kankyo Hen'igen Kenkyu*; 1995; Vol. 16.
(154) Gilboa, R.; Zharkov, D. O.; Golan, G.; Fernandes, A. S.; Gerchman, S. E.; Matz, E.; Kycia, J. H.; Grollman, A. P.; Shoham, G. *J. Biol. Chem.* **2002**, *277*, 19811.
(155) Serre, L.; Pereira de Jesus, K.; Boiteux, S.; Zelwer, C.; Castaing, B. *EMBO J.* **2002**, *21*, 2854.
(156) Mazzella, L. J., Parkin, D. W., Tyler, P. C., Furneaux, R. H.; Schramm, V. L *J. Am. Chem. Soc.* **1996**, *118*, 2111.
(157) Wright, P. M.; Yu, J.; Cillo, J.; Lu, A. L. *J. Biol. Chem.* **1999**, *274*, 29011.
(158) Nash, H. M.; Bruner, S. D.; Scharer, O. D.; Kawate, T.; Addona, T. A.; Spooner, E.; Lane, W. S.; Verdine, G. L. *Curr. Biol.* **1996**, *6*, 968.
(159) Hollis, T.; Ichikawa, Y.; Ellenberger, T. *EMBO J.* **2000**, *19*, 758.
(160) Deng, L.; Schaerer, O. D.; Verdine, G. L. *J. Am. Chem. Soc.* **1997**, *119*, 7865.
(161) Scharer, O. D.; Nash, H. M.; Jiricny, J.; Laval, J.; Verdine, G. L. *J. Biol. Chem.* **1998**, *273*, 8592.
(162) Chepanoske, C. L.; Porello, S. L.; Fujiwara, T.; Sugiyama, H.; David, S. S. *Nucleic Acids Res.* **1999**, *27*, 3197.
(163) Chmiel, N. H.; Golinelli, M. P.; Francis, A. W.; David, S. S. *Nucleic Acids Res.* **2001**, *29*, 553.
(164) Horenstein, B. A.; Schramm, V. L. *Biochemistry* **1993**, *32*, 9917−9925.
(165) Cunningham, R. P.; Ahern, H.; Xing, D.; Thayer, M. M.; Tainer, J. A. *Ann. N.Y. Acad. Sci.* **1994**, *726*, 215.
(166) Zharkov, D. O.; Rieger, R. A.; Iden, C. R.; Grollman, A. P. *J. Biol. Chem.* **1997**, *272*, 5335.
(167) Sidorkina, O. M.; Laval, J. *J. Biol. Chem.* **2000**, *275*, 9924.

(168) Zharkov, D. O.; Golan, G.; Gilboa, R.; Fernandes, A. S.; Gerchman, S. E.; Kycia, J. H.; Rieger, R. A.; Grollman, A. P.; Shoham, G. *EMBO J.* **2002**, *21*, 789.
(169) Harris, T. K.; Turner, G. J. *IUBMB Life* **2002**, *53*, 85.
(170) Stivers, J. T.; Abeygunawardana, C.; Mildvan, A. S.; Hajipour, G.; Whitman, C. P. *Biochemistry* **1996**, *35*, 814.
(171) Chen, X.-Y.; Berti, P. J.; Schramm, V. L. *J. Am. Chem. Soc.* **2000**, *122*, 6527.
(172) Lindahl, T. *Annu. Rev. Biochem.* **1982**, *51*, 61.
(173) Gallivan, J. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9459.
(174) Biot, C.; Buisine, E.; Kwasigroch, J. M.; Wintjens, R.; Rooman, M. *J. Biol. Chem.* **2002**, *277*, 40816.
(175) Lau, A. Y.; Wyatt, M. D.; Glassner, B. J.; Samson, L. D.; Ellenberger, T. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13473.
(176) Ishida, T.; Doi, M.; Ueda, H.; Inoue, M.; Scheldrick, G. M. *J. Am. Chem. Soc.* **1988**, *110*, 2286.
(177) Drohat, A. C.; Kwon, K.; Krosky, D. J.; Stivers, J. T. *Nat. Struct. Biol.* **2002**, *9*, 659.
(178) Cao, C.; Kwon, K.; Jiang, Y. L.; Stivers, J. T. *J. Biol. Chem.* **2003**, *278*, 19442.
(179) Bujnicki, J. M.; Rychlewski, L. *DNA Repair (Amsterdam)* **2002**, *1*, 391.
(180) Makino, K.; Ichikawa, Y. *Tetrahedron Lett.* **1998**, *39*, 8245.
(181) Zechel, D. L.; Withers, S. G. *Acc. Chem. Res.* **2000**, *33*, 11.
(182) Horenstein, B. A.; Zabinski, R. F.; Schramm, V. L. *Tetrahedron Lett.* **1993**, *34*, 7213.
(183) Jiang, Y. L.; Ichikawa, Y.; Stivers, J. T. *Biochemistry* **2002**, *41*, 7116.
(184) Petersen-Mahrt, S. K.; Harris, R. S.; Neuberger, M. S. *Nature* **2002**, *418*, 99.
(185) Tinkelenberg, B. A.; Hansbury, M. J.; Ladner, R. D. *Cancer Res.* **2002**, *62*, 4909.
(186) Chen, R.; Wang, H.; Mansky, L. M. *J. Gen. Virol.* **2002**, *83*, 2339.
(187) Frosina, G. *Eur. J. Biochem.* **2000**, *267*, 2135.
(188) Loeb, L. A.; Christians, F. C. *Mutat. Res.* **1996**, *350*, 279.
(189) Smith, M. L.; Seo, Y. R. *Mutagenesis* **2002**, *17*, 149.
(190) Limp-Foster, M.; Kelley, M. R. *Environ. Mol .Mutagen* **2000**, *35*, 71.
(191) Parkin, D. W.; Leung, H. B.; Schramm, V. L. *J. Biol. Chem.* **1984**, *259*, 9411.
(192) Mentch, F.; Parkin, D. W.; Schramm, V. L. *Biochemistry* **1987**, *26*, 921.
(193) Kampf, G.; Kapinos, L. E.; Griesser, R.; Lippert, B.; Sigel, H. *J. Chem. Soc., Perkin Trans.* **2002**, *2*, 1320.
(194) Jang, Y. H.; Goddard, W. A., III.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. *Chem. Res. Toxicol.* **2002**, *15*, 1023.
(195) Jaworski, S.; Schoellhorn, H.; Eisenmann, P.; Thewalt, U.; Lippert, B. *Inorg. Chim. Acta* **1988**, *153*, 31.
(196) Bellamy, S. R.; Baldwin, G. S. *Nucleic Acids Res.* **2001**, *29*, 3857.

CR010219B