

Organic & Biomolecular Chemistry

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Mechanisms for enzymatic cleavage of the *N*-glycosidic bond in DNA

Alexander C. Drohat and Atanu Maiti

Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine,
Baltimore, MD USA

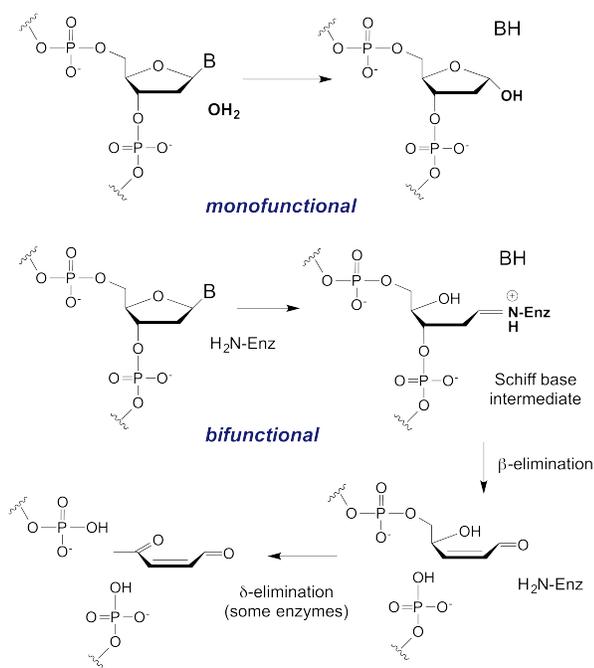
DNA glycosylases remove damaged or enzymatically modified nucleobases from DNA, thereby initiating the base excision repair (BER) pathway, which is found in all forms of life. These ubiquitous enzymes promote genomic integrity by initiating repair of mutagenic and/or cytotoxic lesions that arise continuously due to alkylation, deamination, or oxidation of the normal bases in DNA. Glycosylases also perform essential roles in epigenetic regulation of gene expression, by targeting enzymatically-modified forms of the canonical DNA bases. Monofunctional DNA glycosylases hydrolyze the *N*-glycosidic bond to liberate the target base, while bifunctional glycosylases mediate glycosyl transfer using an amine group of the enzyme, generating a Schiff base intermediate that facilitates their second activity, cleavage of the DNA backbone. Here we review recent advances in understanding the chemical mechanism of monofunctional DNA glycosylases, with an emphasis on how the reactions are influenced by properties of the nucleobase leaving-group, the moiety that varies across the vast range of substrates targeted by these enzymes.

Introduction

The nucleobases in DNA are reactive and subject to continuous modification by endogenous and exogenous agents, producing a broad spectrum of damage that can pose a severe threat to genomic integrity and contribute to aging and diseases including cancer.¹ Three major types of damage occur to the canonical bases in DNA, including alkylation by endogenous and exogenous electrophiles, hydrolytic deamination of the exocyclic amino groups (not relevant for thymine), and oxidation by a slew of reactive oxygen species (ROS). The resulting lesions are corrected by DNA base excision repair (BER), a pathway that is conserved in all forms of life and is initiated by a DNA glycosylase.² These enzymes use a nucleotide-flipping mechanism to recognize chemically modified or mismatched bases in DNA, and remove them by cleaving the carbon-nitrogen bond that

connects the target base to the anomeric carbon of the 2'-deoxyribose sugar.³

DNA glycosylases are found in two basic types, monofunctional and bifunctional.⁴ Monofunctional glycosylases catalyze hydrolysis of the *N*-glycosidic bond, yielding an apurinic/apyrimidinic (AP) or abasic site in the DNA (scheme 1). Bifunctional glycosylases cleave the *N*-glycosidic bond using an amine nucleophile of the enzyme, giving a Schiff base (imine) intermediate that facilitates a second enzymatic activity, cleavage of the phosphodiester backbone on the 3' side of the lesion (β -elimination). Some bifunctional glycosylases also cleave the DNA on the 5' side of the lesion (δ -elimination). The various reaction products of monofunctional and bifunctional glycosylases can be highly mutagenic and cytotoxic, and they are recognized and processed by follow-on BER enzymes to restore the original DNA sequence.



Scheme 1. *N*-glycosidic bond hydrolysis in DNA by monofunctional and bifunctional DNA glycosylases

Our focus here is largely on the mechanism of *N*-glycosidic bond hydrolysis by monofunctional glycosylases. We also consider the relevant non-enzymatic reactions, hydrolysis of 2'-deoxynucleosides, which inform the enzymatic mechanisms. The emphasis is on developments since the most recent reviews on this topic by Berti and McCann (2006)⁵ and Stivers and Jiang (2003),⁴ and in particular on the role of the leaving group base, the moiety which varies across the sea of substrates targeted by these enzymes. We generally do not discuss aspects of the deoxynucleoside sugar or nucleophile activation or positioning, and would direct the reader to the reviews noted above.^{4,5}

Shown in Fig. 1 are the four canonical nucleobases found in DNA, four modified forms of cytosine that are enzymatically generated in the DNA of vertebrates, and four examples from the vast range of damaged bases that are recognized and excised by DNA glycosylases. While some of these enzymes act predominantly on a single type of lesion, others excise a variety of modified bases,

in keeping with findings that the number of different glycosylases in any given organism is typically far smaller than the number of different lesions that must be excised to initiate repair. Some glycosylases excise one of the four canonical bases, if that base resides in a mismatched pair. For example, MutY excises adenine that is mispaired with 8-oxoguanine,⁶ and other glycosylases (TDG, MBD4) selectively excise thymine bases that are mismatched with guanine.⁷

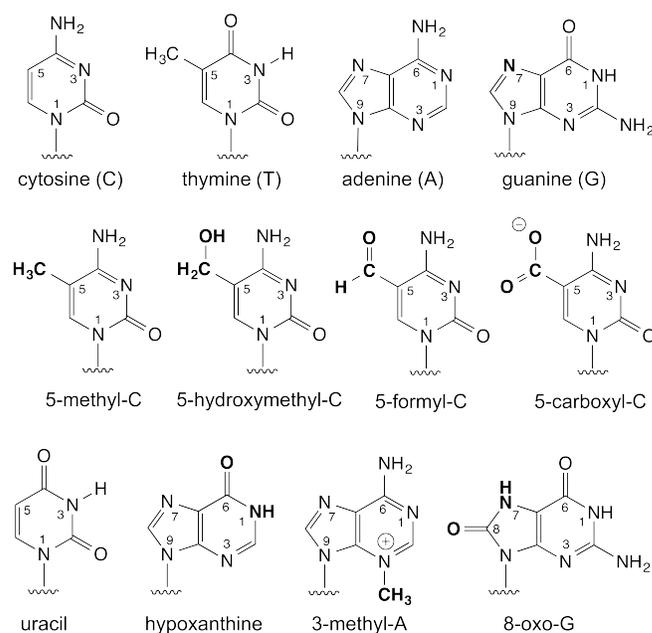


Fig. 1. The four canonical bases in DNA (top), four enzymatically-generated forms of cytosine found in vertebrates (middle), and four examples of damaged bases (bottom) from among the many dozens that arise in DNA.

In addition to removing damaged bases, some glycosylases excise modified bases that are enzymatically derived from one of the four normal DNA bases. For example, active cytosine deaminases convert cytosine to uracil (Fig. 1), which is then excised by uracil DNA glycosylase (UDG), a process that is important for maturation of antibody affinity and the innate defense against retroviruses and retroelements.^{8,9} In addition, methyltransferases generate bases including 5-methylcytosine (mC), a modification that is important for maintaining genomic stability and

transcriptional regulation.¹⁰ In mammals, TET (ten-eleven translocation) enzymes catalyze oxidation of mC, in a stepwise manner, to give 5-hydroxymethyl-C (hmC),^{11, 12} 5-formyl-C (fC), and 5-carboxyl-C (caC).¹³⁻¹⁵ Glycosylase-mediated excision of mC (in plants)¹⁶⁻¹⁸ or of fC or caC (vertebrates)^{14, 19} is an essential step in active DNA demethylation, a process that converts mC back to C and plays a central role in epigenetic regulation of gene expression.¹⁰

Transition-state analysis of reaction mechanisms

The detailed mechanism of many *N*-glycoside hydrolysis reactions has been studied by a variety of approaches, including structure-activity relationships and transition-state analysis. The latter involves the determination of kinetic isotope effects (KIEs) for reactants harboring an isotopic label at multiple positions and computational analysis to identify the TS structure(s) that best match the observed KIEs.^{5, 20} Transition-state analysis has been completed for many enzymatic and non-enzymatic ribonucleotide *N*-glycoside hydrolysis reactions, but only for a limited number of deoxyribonucleoside reactions.

The two limiting mechanisms for hydrolysis of the *N*-glycosidic bond are shown in Figure 2.^{4, 5, 21} In a stepwise $D_N^*A_N$ (or S_N1) mechanism, departure of the nucleobase leaving group (D_N) gives rise to a discrete but highly unstable oxocarbenium ion intermediate (*) that exists for a finite (albeit short) lifetime before addition of the nucleophile (A_N). Alternatively, in a concerted biomolecular A_ND_N (or S_N2) mechanism, nucleophile addition (A_N) begins prior to departure of the leaving group (D_N), such that bond order to both groups exists in a single transition state (Fig. 2). For *N*-glycoside hydrolyses, the A_ND_N reactions are typically highly dissociative, with a single oxocarbenium-ion-like transition state in which leaving group departure is greatly advanced over nucleophile approach.⁵

The stepwise reactions can be further resolved, where the rate limiting step, indicated by the double dagger (\ddagger), is either departure of the

leaving-group ($D_N^*A_N$) or addition of the nucleophile ($D_N^*A_N^\ddagger$).²² Notably, in the latter mechanism, cleavage of the glycosidic bond is reversible. The asterisk in $D_N^*A_N$ indicates that the intermediate is too short-lived for the leaving group to diffuse into solution prior to nucleophile attack. By contrast, for a $D_N + A_N$ mechanism, the intermediate lifetime allows for diffusional equilibration of the leaving group with solvent.

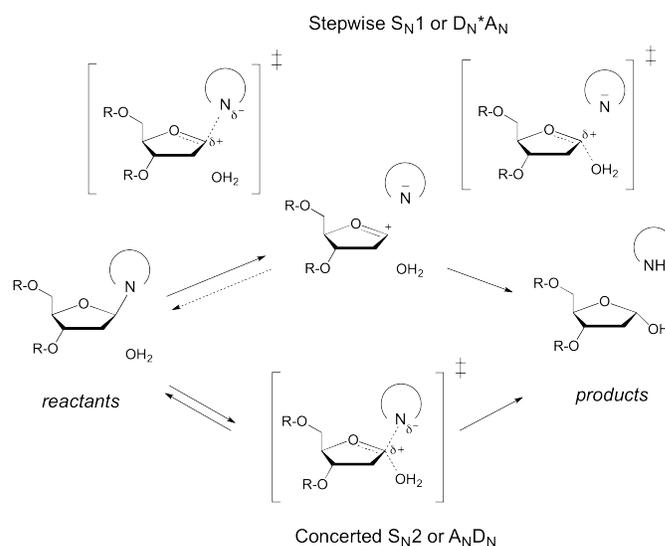


Fig. 2. Two limiting mechanisms for *N*-glycosidic bond hydrolysis in (deoxy)ribonucleosides. The stepwise reaction (top) has two transition states, for departure of the leaving group (D_N) and nucleophile addition (A_N).

While most ribonucleotides are hydrolyzed through a highly dissociative A_ND_N mechanism with some rare examples of $D_N^*A_N$, all studies to date find that hydrolysis of deoxyribonucleotides proceeds through a $D_N^*A_N$ mechanism.²³⁻²⁶ As indicated in Fig. 2, these reactions are nearly always irreversible, though one DNA glycosylase can catalyze the synthesis of an *N*-glycosidic bond, using abasic DNA and one particular nucleobase.²⁷

Non-enzymatic N-glycoside hydrolysis of dAMP
Transition-state analysis of non-enzymatic *N*-glycoside hydrolysis has been reported for only one deoxynucleoside, that being 2'-deoxyadenosine monophosphate (dAMP).^{26, 28} The acid-catalyzed

reaction, collected in 0.1 M HCl, proceeds through a stepwise mechanism in which the glycosidic bond breaks and reforms repeatedly, giving an equilibrium between the reactant and the intermediate, an adenine-oxocarbenium contact ion pair complex (CIPC).²⁶ Expulsion of the adenine leaving-group is catalyzed by protonation at two endocyclic nitrogens, N1 and N7, such that it departs as a monocation. The first irreversible step is either nucleophile (water) attack on the oxocarbenium in the CIPC, 80% of the time, or dissociation of the CIPC followed by nucleophile addition, 20% of the time. These conclusions were based in part on methanolysis reactions, collected in 50% MeOH/0.1 M HCl, which also indicated that the oxocarbenium-ion intermediate has a lifetime of 10^{-11} - 10^{-10} s. Thus, the mechanism is 80% $D_N^*A_N^\ddagger$ and 20% $D_N + A_N$ (more specifically, $D_N^*P^\ddagger + A_N$, where "P" indicates diffusional separation). The results suggest that dAMP hydrolysis at neutral pH follows a $D_N^*A_N^\ddagger$ mechanism.

Differing substantially from dAMP, acid-catalyzed hydrolysis of AMP exhibits a concerted $A_N D_N$ transition-state, with simultaneous approach of the nucleophile and departure of the leaving group.²⁹ The vastly different mechanisms for hydrolysis of AMP and dAMP may reflect the ~ 4 kcal/mol lower stability of a ribocation relative to a 2'-deoxyribocation.³⁰

Ricin-catalyzed N-glycoside hydrolysis of dA in DNA

Ricin, a ribosome-inactivating enzyme that normally excises adenine from RNA, was employed to study the mechanism of enzymatic *N*-glycoside hydrolysis of dA nucleotides in DNA.²⁴ The KIEs indicated a stepwise mechanism but did not distinguish clearly between a $D_N^\ddagger A_N$ or a $D_N^*A_N^\ddagger$ mechanism. Subsequent computational studies suggest that the KIEs support a $D_N^\ddagger A_N$ mechanism, with rate-limiting and irreversible rupture of the *N*-glycosidic bond.²⁶ An important component of the ricin-catalyzed reaction is activation (protonation) of the adenine leaving group, as discussed below.

MutY-catalyzed N-glycoside hydrolysis of dA in DNA

MutY is a DNA repair enzyme that excises adenine bases that are mispaired with 8-oxo-G. Transition-state analysis using multiple KIEs shows that MutY-catalyzed *N*-glycosidic bond hydrolysis follows a stepwise $D_N^*A_N^\ddagger$ mechanism, where the C-N bond breaks and reforms repeatedly prior to irreversible addition of the water nucleophile (A_N).²⁵ The results also demonstrate that the adenine base is protonated (at N7) prior to nucleophile addition, probably before C-N bond cleavage. MutY imparts ground-state distortion to favor a sugar ring conformation that resembles the oxocarbenium ion intermediate. Findings that the commitment to catalysis is negligible and that C-N bond cleavage is reversible lead to the remarkable conclusion that once the oxocarbenium ion intermediate arises in the active site, the most likely outcome is that the glycosidic bond will be reformed and the intact DNA substrate will dissociate from the enzyme. The results also indicate that, following bond cleavage, the adenine leaving group remains in a position that allows for reformation of the C-N bond, at least until the A_N transition state. Thus, the MutY reaction and acid-catalyzed dAMP hydrolysis both follow a $D_N^*A_N^\ddagger$ mechanism, while ricin-catalyzed dA hydrolysis involves a $D_N^\ddagger A_N$ mechanism.

UNG-catalyzed N-glycoside hydrolysis of dU in DNA

Uracil DNA glycosylase (UNG) catalyzes hydrolysis of 2'-deoxyuridine (dU) residues in DNA, and transition-state analysis revealed a stepwise $D_N^*A_N$ mechanism, with an oxocarbenium ion uracil anion intermediate (Fig. 3A).²³ This conclusion was supported by computational studies of the enzymatic reaction using hybrid quantum-mechanical/molecular-mechanical (QM/MM) methods.³¹ The oxocarbenium ion intermediate adopts a slightly 3'-*exo* sugar pucker that differs from the 2'-*endo* conformation of B-type DNA but resembles that observed in the enzyme-substrate complex.³² This suggests that UNG imparts ground-state strain to impose a sugar conformation that

favors a highly dissociative oxocarbenium-ion-like transition-state on path to the oxocarbenium ion intermediate.²³ The KIEs and other findings indicate that the first step, C-N bond cleavage, is rate-limiting, that is, a $D_N^\ddagger A_N$ mechanism.^{23, 26, 31}

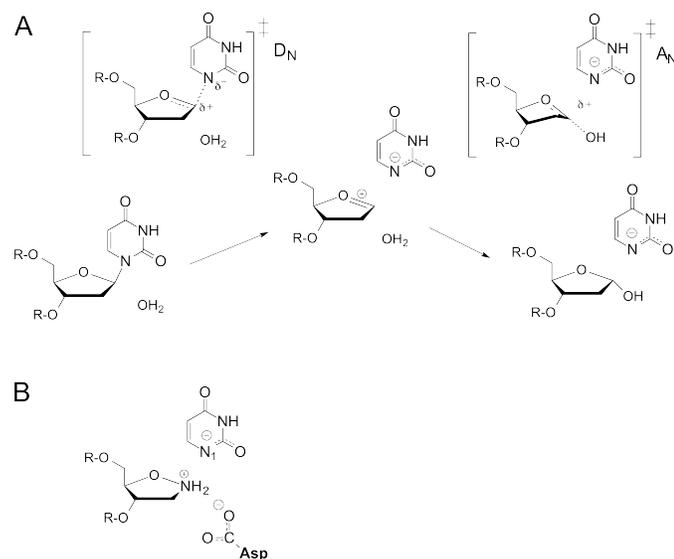


Fig. 3. UNG-catalyzed excision of uracil from DNA. (A) Transition-state analysis and computational studies indicate that UNG follows a stepwise mechanism with irreversible C-N bond cleavage. NMR studies show that the uracil anion is stabilized in the ternary product complex. (B) The uracil anion and a conserved Asp side chain interact with the 1-aza-2'-deoxyribose cation, suggesting similar stabilization of the oxocarbenium ion in the chemical transition state(s).

Observation of a rate-limiting D_N step typically suggests that the leaving group is either (i) activated via protonation, (ii) electrostatically stabilized after departure, or (iii) sequestered in order to suppress reformation of the C-N bond. Consistent with the second mechanism, NMR studies show that UNG dramatically stabilizes the uracil anion in the ternary (UNG-DNA-uracil) product complex at neutral pH (Fig. 3A).³³⁻³⁵ Uracil is also anionic in a ternary complex with UNG and DNA containing a cationic mimic of the oxocarbenium ion intermediate (Fig. 3B).³⁶ These findings, which were not predicted by crystal structures, suggest a similar mechanism for stabilization of the uracil anion in the chemical

transition state(s), as discussed below. Together, the uracil anion and a conserved Asp side chain stabilize the cationic oxocarbenium-ion via an “electrostatic sandwich” (Fig. 3B).^{34, 36}

A crystal structure of UDG bound to DNA containing a cationic 1-aza-2'-deoxyribose and the uracil anion provides a mimic of the second (A_N) transition state.³⁷ Remarkably, the structure indicates an electrophile migration mechanism, whereby electrostatic forces pull the electrophile (oxocarbenium ion) towards the bound water nucleophile (Fig. 3A).

While transition state analysis of non-enzymatic glycosidic bond hydrolysis has not been reported for dU or any other pyrimidine deoxynucleotide, other studies are informative. Non-enzymatic hydrolysis of dU, dT, and 5-halogen-substituted dU proceeds with a strong dependence on leaving group quality (see below) and a low and positive entropy of activation ($T\Delta S$), suggesting a highly dissociative $A_N D_N$ mechanism or perhaps a stepwise ($D_N A_N$) mechanism.^{38, 39} Computational studies of dU hydrolysis using QM/MM methods indicate the former.³¹

Phosphorylase-catalyzed hydrolysis of dT in DNA

Thymidine phosphorylase (TP) catalyzes reversible phosphorolysis of the *N*-glycosidic bond in thymidine (dT), and in the absence of phosphate it can use water as the nucleophile to hydrolyze dT. Transition state analysis shows that the hydrolytic reaction proceeds through a stepwise $D_N A_N^\ddagger$ mechanism, with reversible C-N bond cleavage followed by irreversible and rate-limiting capture of the water nucleophile by the deoxyribocation intermediate.⁴⁰ Thus, the intermediate is in equilibrium with the reactant (dT) prior to nucleophile capture, as observed for acid-catalyzed dAMP hydrolysis and the MutY reaction. The equilibrium isotope effect (EIE) between free dT and N1-protonated thymine indicates that a departing thymine anion is protonated at N1 prior to nucleophile capture of the deoxyribocation. The near-unity ¹⁵N KIE indicates little or no

rehybridization of N1, hence no participation of the leaving group, at the rate-determining transition state for nucleophile addition. The steep reaction barrier for A_N is thought to reflect improper alignment of water or its weak nucleophilic nature relative to the preferred phosphorylytic reaction of TP. The results suggest general base activation of the water nucleophile, potentially by a His side chain. In remarkable contrast, the phosphorylytic reaction follows a nearly synchronous $A_N D_N$ transition state with no cationic character.⁴¹

Role of the Leaving Group

The various nucleotides in DNA, be they canonical, enzymatically modified, or damaged, differ in the chemical nature of the nucleobase. As such, it is of interest to consider how *N*-glycosidic bond hydrolysis may be influenced by the base, given its key role as the leaving group in these reactions. Findings that a stepwise ($D_N^* A_N$) pathway predominates in glycosidic bond hydrolysis for 2'-deoxynucleosides indicate that properties of the leaving-group can have a large impact on the reaction mechanism and rate.

The leaving-group quality of a nucleobase depends on its capacity to accommodate the increased electron density generated by heterolytic cleavage of the glycosidic bond (Fig. 4A). In the absence of catalysis, a neutral base departs as an anion, which is usually unstable in solution and a relatively good nucleophile, and typically a poor leaving group. By contrast, a cationic base departs as a neutral species, which is generally a better leaving group than the anion. As such, leaving-group quality of a neutral base can be enhanced by protonation (acid catalysis), typically prior to C-N bond cleavage, and/or electrostatic stabilization of the departing anion. Alkylation, if it imparts a positive charge, improves leaving group quality because the base departs as a neutral species.

A key indicator of leaving-group quality is the acidity of the glycosidic nitrogen (N1 for pyrimidines, N9 for purines). Greater acidity, or a lower pK_a , reflects a more stable N-deprotonated

species (conjugate base) and a weaker nucleophile, hence a better leaving group for deoxynucleotide hydrolysis reactions (Fig. 4B). For instance, N1 is more acidic for 5FU relative to U because the N1 anion is more stable for 5FU (F substituent stabilizes the anion), and the leaving-group quality is better for 5FU relative to U. The lower N9 pK_a for 3-methyl-A relative to A reflects greater stability of the N9-deprotonated species for the former, which is a better leaving group (Fig. 4B).

Below, we consider mechanisms for catalyzing leaving-group departure in glycosidic bond hydrolysis reactions, first for purines and then for pyrimidines.

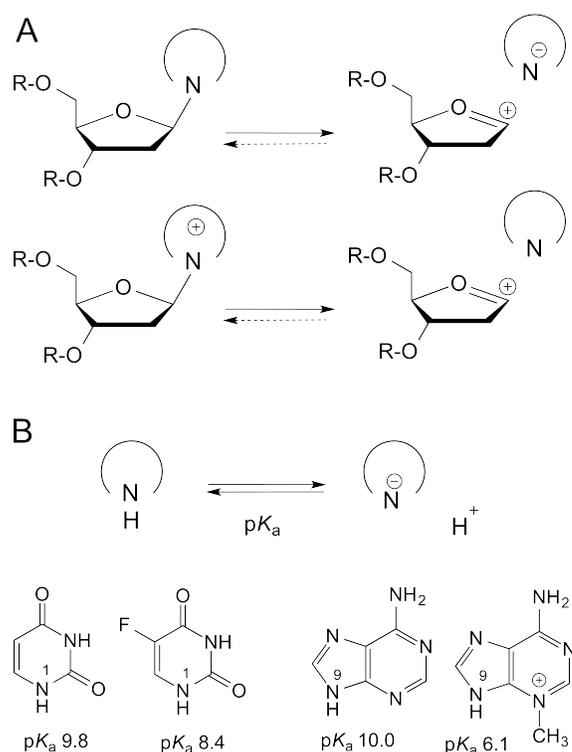


Fig. 4. Leaving group quality of the nucleobase depends on the acidity (pK_a) of its glycosidic nitrogen. (A) In the stepwise reactions for glycosidic bond hydrolysis, a neutral base departs as an anion and a cationic base departs as a neutral species. (B) Acidity of the glycosidic nitrogen for a given base depends on the stability of the N-deprotonated species (conjugate base). The glycosidic nitrogen is N1 for pyrimidines, N9 for purines. Acidity (pK_a) is shown for U, 5FU, A, and 3-methyl-A.

Catalysis of purine departure – Hydrolysis of purine deoxynucleosides (dA, dG) is strongly acid catalyzed via protonation of nitrogen(s) in the nucleobase, which can dramatically improve leaving-group quality, as noted above.^{5, 42, 43} The mono-protonated base departs as a neutral species, a much better leaving group than the anion, and the di-protonated base departs as a monocation, an even better leaving group.

The pH profiles for hydrolysis of purine nucleosides (dA, dG) and for DNA depurination show acid-catalysis at pH 7.4, even as the pK_a values for the most basic ring nitrogens are well below 7.4. The depurination rates are very slow at physiological pH, reflecting the very low but significant population of the protonated species. For hydrolysis of dA at pH 7.4, acid catalysis involves N1 protonation (pK_a^{N1} 3.7) and gives an estimated rate enhancement of $10^{3.6}$ -fold.^{5, 42} An additional rate enhancement of about $10^{5.3}$ -fold is realized by protonation at a second site, N7 (pK_a^{N7} -1.3) to give diprotonated dA. Notably, transition-state analysis gives direct evidence for protonation at the second site (N7) in acid-catalyzed hydrolysis of dAMP (performed at pH 1).²⁶ Similarly, pH profiles indicate that acid catalyzed dG hydrolysis at pH 7.4 involves N7 protonation (pK_a^{N7} 2.4) and gives an estimated rate enhancement of 10^5 -fold.^{5, 42} An additional rate enhancement of roughly $10^{4.7}$ -fold is realized by ionization at a second site, N3 (pK_a^{N3} -2.4) to give di-protonated dG.^{5, 42}

As noted previously,⁵ the pK_a of -1.3 for N7 of dA is a *macroscopic* value, reflecting the interconversion of mono- and di-protonated species. However, protonation of *neutral* dA at N7 occurs with a *microscopic* pK_a of about 2.4, thus acidity is not dramatically greater for N7 relative to N1. Similar reasoning likely applies to N3, the most acidic purine nitrogen. This is relevant to acid catalysis by enzymes, which do not always protonate the most basic nitrogen of the purine leaving-group.

Enzymatic catalysis of adenine departure – Ricin-catalyzed hydrolysis of dA in DNA has been

proposed to involve mono- or di-protonation of adenine to activate it for departure, but the site(s) of nitrogen protonation are not yet established. Transition state analysis suggests that N7 is protonated in ricin-catalyzed hydrolysis of adenosine in RNA.³⁰ It was proposed that N7 and possibly N1 are protonated in ricin-mediated hydrolysis of dA in DNA.²⁴ The enzyme has no side chain that can directly protonate N7 via general acid catalysis, but it does provide backbone oxygen contacts that could favor N7 protonation, with protons derived from solvent.⁴⁴ Structural and mutagenesis studies suggest that a conserved Arg side chain could protonate adenine at N3.⁴⁴⁻⁴⁶ Thus, evidence to date favors N3 and N7 as sites of protonation for ricin-catalyzed dA hydrolysis.

Transition-state analysis demonstrates that MutY-mediated hydrolysis of dA in DNA involves protonation of the adenine base at N7, prior to glycosidic bond cleavage.²⁵ N7-protonation increases the calculated acidity of adenine (N9-H) by over 100 kcal/mol in the gas-phase.⁴⁷ A conserved Asp side chain, which is essential for base excision, contacts adenine N7, suggesting a role in general acid catalysis of adenine expulsion.⁴⁸⁻⁵⁰ MutY activity is reduced at least $10^{4.4}$ -fold when N7 of adenine is replaced by carbon; it cannot excise 7-deaza-A.^{47, 51} Moreover, activity is reduced 40-fold for 1-deaza-A, 100-fold for 3-deaza-A,^{47, 52} and 6000-fold for an adenine analogue lacking both N1 and N3,^{47, 53} suggesting that electrostatic contacts to N1 and N3 stabilize the departing adenine in the transition-state. An Arg side chain contacts N1, and an ordered water molecule contacts N3 in the enzyme-substrate (ES) complex.⁴⁹ While N1 is more basic than N3 and N7, it is not the site of protonation and appears to be the least important endocyclic nitrogen in the enzymatic reaction.

Enzymatic excision of damaged purines – Mammalian alkyladenine DNA glycosylase (AAG) protects against DNA damage caused by alkylation and oxidative deamination of purine bases in DNA. Acid catalysis is observed for AAG excision of two

damaged forms of purine, hypoxanthine (Hx, Fig. 2) and 1,*N*⁶-ethenoadenine (ϵ A, Fig. 5).⁵⁴ The pH profiles suggest an essential protonated group in the ES complex, with pK_a values of 6.4 and 7.3 for Hx and ϵ A, respectively. Because mutational analysis revealed no general acid for AAG,⁵⁴ the data could potentially reflect ionization of the nucleobase in the ES complex, with a proton derived from solvent. This could be facilitated by an ordered water molecule, and one is proximal to N7 of ϵ A in the ES complex.^{5, 55} Evidence that acid catalysis for Hx involves N7-protonation is indicated by findings that activity is reduced by over 10^7 -fold when N7 is replaced by carbon; AAG cannot excise 7-deaza-Hx (Fig. 5).⁵⁴ This idea is supported by the absence of acid catalysis for 7-methylguanine (7meG, Fig. 5).⁵⁴ 7meG is cationic (pre-activated for departure), and N7-methylation mimics the effect of N7-protonation. Indeed, the glycosidic nitrogen (N9) is much more acidic for 7meG relative to Hx and ϵ A, indicating 7meG is an inherently good leaving group.⁵⁶ Catalysis for ϵ A could involve protonation at N10, pK_a 4.05,⁵⁷ which has greater proton affinity than N7.⁵⁸ Notably, it was also proposed that AAG excises some lesions without acid catalysis, such that the base departs as an anion; the glycosidic nitrogen (N9) is more acidic for Hx and ϵ A relative to A and G in the gas phase, and this trend may hold in a hydrophobic active site, which could make Hx and ϵ A more susceptible to enzymatic excision.^{56, 58, 59}

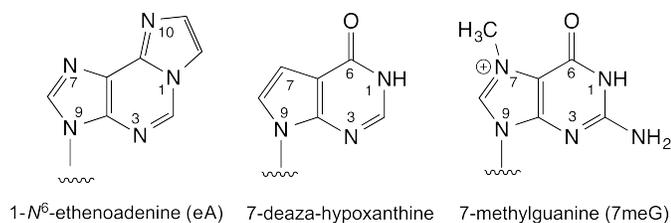


Fig. 5. Damaged forms of adenine and guanine. AAG can remove Hx, ϵ A, and 7meG but not 7-deaza-Hx.

The *E. coli* 3-methyladenine DNA glycosylase I (TAG) excises 3-methyladenine (3meA) and 3-methylguanine (3meG). These cationic bases are expelled as neutral species after C-N bond

cleavage. The effect of N3-methylation on leaving group quality is reflected by the much greater acidity of the glycosidic nitrogen (N9) for 3meA and 3meG relative to A.⁵⁶ Structures of TAG and structure-activity relationships for binding 3meA analogues indicate that the hydrophobic enzyme active site stabilizes the neutral form of 3meA, i.e., the leaving group, relative to the cationic ground-state.⁶⁰⁻⁶² The results suggest TAG imposes ground-state electrostatic strain on the cationic nucleotide substrate and selectively stabilizes the departing neutral base in the transition state. Specificity against neutral substrates may involve tight ground-state binding and poor stabilization of a departing anion in the transition state.

In contrast to AAG and TAG, the *E. coli* 3-methyladenine DNA glycosylase II (AlkA) does not distinguish between methylated and undamaged bases.^{63, 64} Rather, it excises a broad variety of alkylated purines and pyrimidines and has some activity for undamaged DNA bases. The remarkably indiscriminant active site of AlkA exhibits a relatively uniform rate enhancement of $\sim 10^5$ -fold for excision of alkylated and normal bases. As such, specificity depends largely on the reactivity of the *N*-glycosidic bond.⁶⁴

The Fpg-MutM enzymes target oxidized purines such as formamidopyrimidines and 8-oxoguanine (oxoG, Fig. 1).⁶⁵ The reaction mechanism is unknown but could be stepwise ($D_N^*A_N$), where an oxocarbenium intermediate captures a nucleophile provided by the enzyme, the α -nitrogen of an *N*-terminal Pro residue.⁶⁶ Structures suggest these enzymes could catalyze departure of the oxoG anion via electrostatic stabilization.⁶⁶ The glycosidic nitrogen (N9) is more acidic for oxoG relative to G,⁶⁷ due perhaps to resonance stabilization of the oxoG anion (Fig. 6), such that leaving group quality is likely better for oxoG relative to G. Remarkably, Fpg provides four hydrogen bonds (backbone N-H) to the O6 oxygen of oxoG, which could stabilize an oxoG anion.⁶⁶ Findings that acid-catalyzed C-N hydrolysis is slower for 8-oxo-dG versus dG⁶⁸ likely reflect the fact that acid catalysis involves N7-

protonation for dG (see above), which cannot occur for 8-oxo-dG. Rather, acid-catalyzed (nonenzymatic) hydrolysis of 8-oxo-dG could involve N3-protonation; N3 has a calculated pK_a of 0.2.⁶⁷ While acid catalysis at N3 cannot be excluded for the Fpg-MutM reaction, the site is rather acidic and the enzymes have no obvious group to perform such a role.

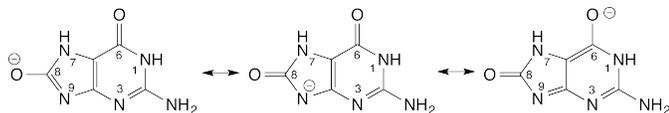


Fig. 6. Resonance structures of the 8-oxoguanine anion.

OGG1 (8-oxoguanine DNA glycosylase) enzymes also excise oxoG,⁶⁹ and the nucleophile can be either an enzyme group (Lys in this case) or a water molecule.⁷⁰ OGG1 provides no contacts to O8 and few contacts to other regions of oxoG.⁷¹ Contacts to O6 involve only two water molecules for OGG1, compared to four backbone N-H contacts for Fpg. Despite the absence of contacts to O8, structure-activity relationships with oxoG analogues indicate that a carbonyl oxygen at C8 is important for OGG1 activity; 8-methoxy-G is not a substrate.⁷² While this finding was interpreted as evidence that OGG1 protonates O8 to activate oxoG,⁷² it is also consistent with departure of an oxoG anion, stabilized by charge delocalization to O8 (Fig. 6). Interestingly, it was proposed that an anionic oxoG leaving group participates in the lyase activity of OGG1, by deprotonating the Lys nucleophile to generate the Schiff base intermediate (scheme 1).⁷³

Some lesions resulting from purine deamination are excised by enzymes in the UDG superfamily, including TDG-MUG, SMUG1, UDGa/b, and HDG.⁷⁴⁻⁷⁸ Yeast TDG (thymine DNA glycosylase) excises Hx and xanthine (X, Fig. 7).⁷⁵ Remarkably, *E. coli* MUG (mismatch uracil glycosylase) has greater activity for X than it does for U.⁷⁷ As discussed below, enzymes of the UDG superfamily do not use acid catalysis to activate uracil departure, and this might also be true for their excision of purine lesions. The glycosidic nitrogen of Hx is more acidic

(pK_a^{N9} 8.9) than that of U (pK_a^{N1} 9.8), suggesting that the Hx anion could be a good leaving group.^{39, 56}

The glycosidic nitrogen of X is even more acidic (pK_a^{N9} 7.3, calculated), thus the X monoanion is predicted to be an excellent leaving group (Fig. 7).^{56, 79} However, N3-H is also relatively acidic, with a pK_a^{N3} of 5.7 for the xanthosine nucleoside,⁸⁰ such that 2'-deoxyxanthosine (dX) is predominantly anionic at pH 7.4 (Fig. 7). Non-enzymatic hydrolysis of dX in ssDNA is pH-independent for pH >7 and acid-catalyzed for pH 3-7, with an apparent pK_a of ~5 that likely reflects N3-protonation of the dX anion (though the results were suggested to reflect N7-protonation of neutral dX).⁸¹ The predominance of the dX anion in DNA at neutral pH likely explains its resistance to depurination,⁸² which was not initially expected.¹ Because the dianion of X is likely to be a very poor leaving group, excision of the X monoanion seems unlikely. Thus, enzymatic hydrolysis of dX in DNA could potentially be acid-catalyzed via N3-protonation, such that X departs as a monoanion, which is predicted to be a good leaving group (Fig. 7).

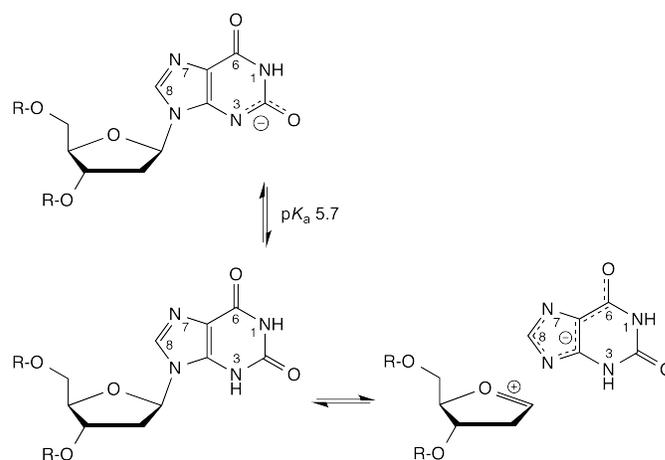


Fig. 7. 2'-deoxyxanthosine (dX) ionizes at N3 with a pK_a of about 5.7 (based on xanthosine) and is predominantly anionic at pH 7.4. Hydrolysis of neutral dX results in departure of the xanthine (X) monoanion, which should be a good leaving group, given that N9 is acidic for X relative to other purines, due perhaps to charge delocalization (N7, O6).

Non-enzymatic hydrolysis of dU and its analogs – While glycosidic bond hydrolysis is acid catalyzed at

neutral pH for dA, dG, and dC (below), hydrolysis of dU (and dT) is pH-independent for pH 2-8 and acid catalyzed for pH <2.^{5, 39} For these pyrimidines, N3 is protonated at neutral pH ($pK_a^{N3} > 9$) and the two carbonyl oxygens (O2, O4) are highly acidic, with $pK_a < -3$. Accordingly, for non-enzymatic hydrolysis of dU and dT at neutral pH, the base is not protonated prior to C-N bond cleavage and it departs as an anion (Fig. 8A). While acid-catalysis of base departure has been proposed for many glycosylases that excise U or T, such a mechanism is unlikely due to the acidity of O2 and O4.^{33, 34} Moreover, acid catalysis is not required, because charge delocalization to the two carbonyl oxygens (O2, O4) stabilizes the departing anion and thereby improves its leaving-group quality (Fig. 8B).

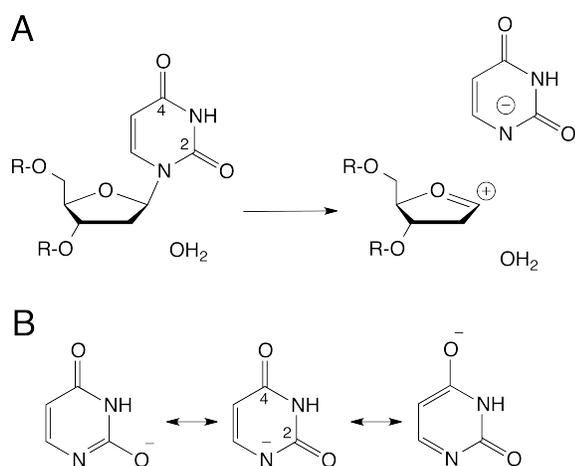


Fig. 8. (A) Hydrolysis of dU, dT, and dU analogs is not acid catalyzed at neutral pH, for enzymatic or non-enzymatic reactions, because the carbonyl oxygens are highly acidic ($pK_a < -3$), so the base departs as an anion. (B) Resonance stabilization of the uracil anion involves the two carbonyl oxygens (O2, O4).

Given that a stepwise ($D_N^*A_N$) mechanism is observed for glycosidic bond hydrolysis reactions in deoxynucleotides (see above), the rate can be expected to depend on the leaving-group quality of departing base, at least for reactions in which the D_N step is rate limiting. One indicator of leaving-group quality is the acidity of the glycosidic nitrogen (N1 for pyrimidines), where greater acidity

(lower pK_a) indicates a more stable anion, a weaker nucleophile, and thus a better leaving group.

Supporting this idea, a Brønsted-type linear free energy relationship (LFER) is observed for the dependence of reaction rate ($\log k_{obs}$) on leaving-group acidity (N1 pK_a) for non-enzymatic hydrolysis of 5-substituted deoxyuridines, including dU, dT, 5-F-dU, 5-Cl-dU, and 5-Br-dU (Fig. 9).^{38, 39, 83} The electron-withdrawing ($\sigma_m > 0$) halogen substituents stabilize the uracil anion, enhancing its leaving-group quality and increasing the reaction rate.

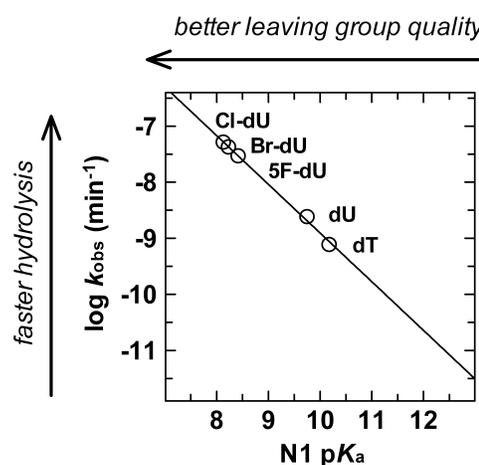


Fig. 9. Brønsted-type linear free energy relationship (LFER) for non-enzymatic hydrolysis of 5-substituted deoxyuridines.⁸³ The dependence of $\log k_{obs}$ on N1 pK_a (acidity of the leaving-group nitrogen) for hydrolysis of dT, dU, 5-F-dU, 5-Br-dU, and 5-Cl-dU gives a slope of $\beta_{lg} = -0.87 \pm 0.03$. The k_{obs} values were extrapolated to 22 °C using values at higher temperatures,^{39, 87} as described.⁸³ For 5-Br-dU, a corrected N1 pK_a of 8.24 was used⁸³ rather than the previously reported 8.49.³⁹

Enzymatic hydrolysis of dU and its analogs – We next consider how enzymes facilitate leaving group departure during glycosidic bond hydrolysis of dU, dT, and related nucleotides in DNA. As noted above, UNG catalyzes the hydrolytic excision of uracil from DNA. In the stepwise ($D_N^*A_N$) UNG reaction,^{23, 31} uracil departs as an anion, stabilized by hydrogen bonds to O2 and O4 from backbone and side-chain groups. Most notable among these is the imidazole group of a His residue that is strictly conserved in UNG enzymes and some other

members of the UDG superfamily.^{32, 84} The catalytic imidazole is neutral through all steps of the UNG reaction, as demonstrated by NMR, and it provides a strong hydrogen bond to uracil O2, stabilizing the anion by ~ 5 kcal/mol (Fig. 10A).^{33-35, 85} Remarkably, UNG binds uracil as an anion in a ternary product complex with abasic DNA at neutral pH (Fig. 10A).³⁵ Thus, UNG stabilizes the uracil anion by about 5 kcal/mol, reducing its N1 pK_a from 9.8 in solution to 6.4 in the enzyme-product (EP) complex,³⁵ due largely to the strong hydrogen bond noted above.⁸⁵ The uracil anion is even further stabilized ($pK_a^{N1} < 4.5$) when bound in an EP complex that includes a cationic mimic of the oxocarbenium ion (Fig. 10B).³⁶ These observations suggest that the catalytic His residue provides a strong hydrogen bond to stabilize the departing uracil anion in the chemical transition state(s) of the UNG reaction. Notably, the catalytic His residue is also conserved in SMUG1 enzymes,⁸⁶ which belong to the UDG superfamily, though its role in these enzymes has not been rigorously examined.

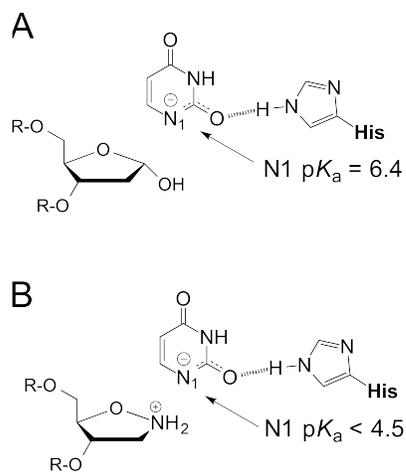


Fig. 10. Enzymatic stabilization of anionic leaving group. UNG provides a strong hydrogen bond to stabilize the uracil anion in the product complex, as indicated by a reduced N1 pK_a of 6.4 relative to that of 9.8 for uracil in solution. (B) The uracil N1 pK_a is even further reduced in a ternary complex with UNG and DNA that contains a mimic of the cationic oxocarbenium ion.

The UNG superfamily includes the TDG-MUG family of enzymes, which can excise many

pyrimidines from DNA, but do not have the catalytic His residue that facilitates leaving-group departure in UNG enzymes. The findings discussed above for glycosidic bond hydrolysis in deoxynucleotides suggest that the TDG-MUG enzymes follow a stepwise ($D_N^*A_N$), or perhaps a highly dissociative (A_ND_N) mechanism. As such, excision of U, T, or U analogs by TDG-MUG enzymes likely involves expulsion of an anionic leaving group. Consistent with this idea, the activity of TDG (k_{obs}) depends on the N1 acidity of the excised pyrimidine base, as demonstrated by a Brønsted-type LFER for the dependence of $\log k_{obs}$ on pK_a^{N1} for uracil and cytosine bases harboring various C5-substituents (Fig. 11).⁸³ Moreover, differences in leaving-group quality are enhanced in the hydrophobic TDG active site compared to aqueous solution, as indicated by the steeper slope of the LFER for TDG ($\beta_{lg} = -1.6$) relative to non-enzymatic reactions ($\beta_{lg} = -0.87$).⁸³ As noted previously, a steeper (more negative) β_{lg} can be expected for reactions in environments that are less-polar than aqueous solution, because charge development in the transition state of the latter is stabilized to a greater extent by solvent.^{88, 89}

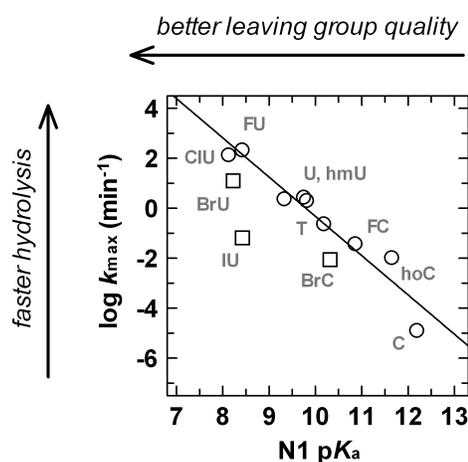


Fig. 11. Brønsted-type LFER for enzymatic (TDG) hydrolysis of 5-substituted dU and dC nucleotides in DNA as reported by Bennett et al.⁸³ The dependence of $\log k_{obs}$ on N1 pK_a for the base gives $\beta_{lg} = -1.6 \pm 0.2$; included in the fitting are U, T, FU, CIU, hoU, hmU, FC, hoC, and C (O). Data for BrU, IU, and BrC (□) were not included in the fitting, because k_{max} values

suggested limited access of these bases to the TDG active site.⁸³

Hydrolysis of deoxycytidine and its analogues –

As noted above, protonation of the departing base is a common mechanism for catalyzing glycoside hydrolysis, because it enhances leaving-group quality. Non-enzymatic hydrolysis of deoxycytidine (dC) is acid catalyzed via protonation at N3, which ionizes with a pK_a of 4.3.^{38, 90, 91} Leaving-group quality is greatly improved for N3-protonated relative to neutral dC, but the rate of non-enzymatic dC hydrolysis is still very slow at physiological pH because the N3-protonated species is only sparsely populated.

Non-enzymatic hydrolysis of 5-Br-dC is also acid catalyzed by N3 protonation, but the electron-withdrawing halogen substituent enhances acidity at both N1 and N3, with counteracting effects on the reaction rate. While the 5-Br substituent enhances leaving-group quality by increasing N1 acidity ($\Delta pK_a^{N1} = -2$), it also increases N3 acidity ($\Delta pK_a^{N3} = 1.6$) such that acid catalysis is less effective.^{92, 93} Thus, leaving-group quality is better for 5-Br-C versus C, but dC is hydrolyzed as rapidly as 5-Br-dC (95 °C, pH 5) because acid catalysis is more effective for dC than 5-Br-dC. The effect of electron-withdrawing substituents on acidity (N1, N3) and glycosidic bond hydrolysis for dC analogues is an important factor when considering the enzymatic removal of the oxidized forms of 5-methylcytosine found in mammalian DNA, as discussed below.

Enzymatic excision of cytosine and its analogues from DNA is relatively rare and has not been extensively studied. A variant of UNG has aberrant activity for removing C from normal GC pairs;^{94, 95} a conserved Asn that contacts the Watson-Crick region of uracil is replaced by Asp, such that favorable contacts can be made with cytosine. Similar to the non-enzymatic reaction, the Asn→Asp UNG variant employs acid catalysis, with the Asp facilitating cytosine N3-protonation.⁹⁵

Although no glycosylases in animals can excise mC, this activity is possessed by the DEMETER (DME) and ROS1 glycosylases in plants, and these enzymes perform key functions in epigenetic regulation.¹⁶⁻¹⁸ While some mechanistic insight has been obtained, their approach for catalyzing leaving group departure is unknown. Given that C and mC exhibit very weak N1 acidity ($pK_a^{N1} > 12$) and the N1 anions have poor leaving group quality,^{83, 92} it seems reasonable that DME and ROS1 enzymes may activate mC by general acid catalysis via N3-protonation. Two residues that are conserved in DME and ROS1 enzymes and essential for mC excision activity include an Asp and a His,⁹⁶ which appear to be positioned on opposite sides of the flipped mC base. It is conceivable that one of these residues could facilitate N3-protonation to activate mC.

Endo III-Nth and Nei enzyme families excise some oxidized forms of dC, including 5-hydroxy-dC and 5,6-dihydroxy-dC, and they remove a broad range of other lesions, but it is not presently clear how they catalyze leaving group departure.⁵

In contrast to other DNA glycosylases, TDG excises many cytosine analogues,^{83, 91} some of which are also removed by the related *E. coli* MUG enzymes.^{97, 98} Most of these TDG substrates have an electron-withdrawing ($\sigma_m > 0$) substituent at C5, which enhances the otherwise dismal leaving-group quality of cytosine. These substrates include 5-fluoro-C (5FC), 5-bromo-C (5BrC), and 5-hydroxy-C (5hoC).³¹ TDG also excises fC and caC,^{45, 46} two recently-discovered components of DNA that arise by Tet-mediated oxidation of mC (Fig. 1).¹³

As shown in Fig. 12, fC is acidic (N1) relative to the established TDG substrates U and T, indicating it is a much better leaving group than cytosine.⁹¹ This can be explained by the electronic effect of the formyl substituent ($\sigma_m = 0.35$)⁹⁹ and resonance stabilization of the fC anion via charge delocalization to the formyl and O2 oxygens. By contrast, the electron-donating carboxylate group ($\sigma_m = -0.10$)⁹⁹ decreases the N1 acidity of caC, which exists predominantly as an anion at

physiological pH (Fig. 12).^{91, 100} As such, the caC anion is less acidic (N1) than C or hmC, bases that are not removed by TDG (Fig. 12).⁹¹ This is consistent with the idea that a dianionic base would be a very poor leaving group. However, the neutral forms of caC, particularly the N3-protonated forms, are much more acidic than the anion (Fig. 12), indicating they are much better leaving groups. This suggests that enzymatic excision of the caC anion from DNA requires acid catalysis, such that the base departs as a monoanion rather than a dianion.

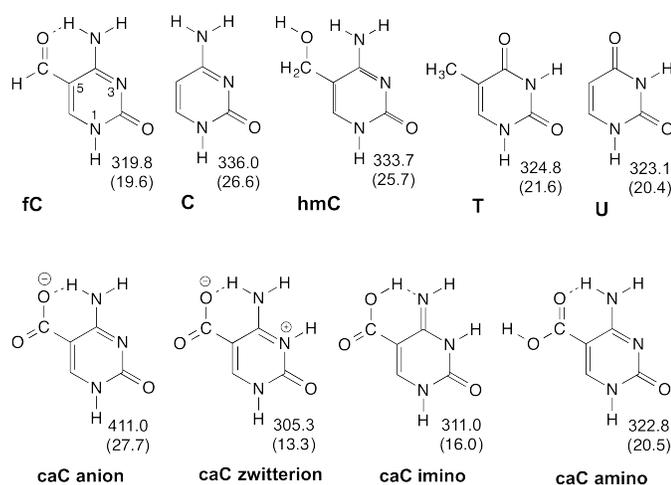


Fig. 12. Calculated N1 acidity for pyrimidines are given as the free energy (ΔG , kcal mol⁻¹) of deprotonation in the gas phase and bulk solution (parentetical values).⁹¹ Lower values indicate greater N1 acidity and better leaving-group quality.

Consistent with this idea, TDG excision of caC is acid catalyzed (Fig. 13), though it does not appear to involve a general acid from the enzyme. Rather, caC ionizes with an apparent pK_a of 5.8 in the TDG active site, likely at N3, facilitated by a conserved Asn side chain.⁹¹ In contrast, no evidence of acid catalysis is observed for TDG excision of fC (Fig. 13)⁹¹ 5fC, BrC, or hoC.⁸³ However, this is not unexpected. While N3-protonation would likely enhance TDG excision of these cytosine analogues it is unlikely to be observed because the N3 pK_a values are far below the pH range for which TDG is active.^{90, 92, 101} Thus, acid catalysis is not observed for TDG excision of cytosines bearing an electron-

withdrawing substituent that enhances leaving-group quality (lowers N1 pK_a), but it is needed for excision of caC, which has an electron-donating substituent that weakens leaving-group quality (raises N1 pK_a). A remaining question is how TDG avoids N3-protonation and subsequent excision of C, mC, hmC, which have N3 pK_a values close to that of caC and should be similarly activated by N3-protonation.⁹¹

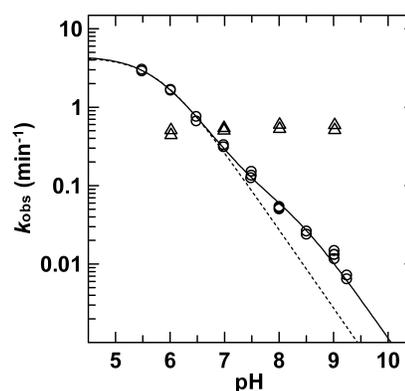


Fig. 13. pH dependence of TDG activity for G·fC (Δ) and G·caC (O) substrates.⁹¹ The G·caC data is best fitted to a model for ionization of one essential protonated group ($pK_a^1 = 5.75 \pm 0.03$) and a second group ($pK_a^2 = 8.2 \pm 0.7$) that is not essential but confers higher activity when deprotonated (solid line). Fitting is poor for a model with a single essential protonated group (dotted line).

Concluding Remarks

It was noted by Berti and McCann in 2006 that 2'-deoxynucleosides are poorly represented in transition-state (TS) analysis of *N*-glycosidic bond cleavage.⁵ While subsequent studies have been reported, the situation has not changed dramatically. This is particularly true for non-enzymatic reactions, which include only one purine, dAMP,²⁶ and no pyrimidines. TS analysis of non-enzymatic hydrolysis reactions for additional deoxynucleotides would likely be highly informative, including dTMP, 5-methyl-dCMP and some of its oxidized derivatives (5-formyl-dCMP, 5-ca-dCMP) and for some important lesions such as 8-oxo-dGMP. Likewise, TS analysis of enzymatic *N*-glycosidic bond cleavage includes only two DNA glycosylases (UNG, MutY), plus the "moonlighting"

reactions catalyzed by ricin and TP (which do not normally act on deoxynucleotides in DNA). While TS analyses of related reactions can be informative, much remains to be learned about the chemical mechanism of DNA glycosylases. However, TS analysis is challenging and is pursued by a limited number of groups, so progress will likely take time.

Of course, many other approaches are used to elucidate enzyme mechanisms, including structural biology, NMR and other spectroscopic methods, transient kinetics, equilibrium binding, structure-activity relationships, mutational analysis, and, increasingly, computational methods. These various tools have already provided much information regarding the mechanism of DNA glycosylases, as discussed above, and a multi-disciplinary approach will be needed to rigorously address many of the remaining questions, some of which are noted above. We anticipate substantial future progress toward understating the chemical mechanism of DNA glycosylases.

REFERENCES

1. T. Lindahl, *Nature*, 1993, **362**, 709-715.
2. K. Hitomi, S. Iwai and J. A. Tainer, *DNA Repair (Amst)*, 2007, **6**, 410-428.
3. S. C. Brooks, S. Adhikary, E. H. Rubinson and B. F. Eichman, *Biochim. Biophys. Acta*, 2013, **1834**, 247-271.
4. J. T. Stivers and Y. L. Jiang, *Chem. Rev.*, 2003, **103**, 2729-2759.
5. P. J. Berti and J. A. McCann, *Chem. Rev.*, 2006, **106**, 506-555.
6. S. S. David, V. L. O'Shea and S. Kundu, *Nature*, 2007, **447**, 941-950.
7. D. Cortazar, C. Kunz, Y. Saito, R. Steinacher and P. Schar, *DNA Repair (Amst)*, 2007, **6**, 489-504.
8. C. S. Nabel, S. A. Manning and R. M. Kohli, *ACS Chem. Biol.*, 2012, **7**, 20-30.
9. W. A. Pastor, L. Aravind and A. Rao, *Nat. Rev. Mol. Cell Biol.*, 2013, **14**, 341-356.
10. R. M. Kohli and Y. Zhang, *Nature*, 2013, **502**, 472-479.
11. M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind and A. Rao, *Science*, 2009, **324**, 930-935.
12. S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang, *Nature*, 2010, **466**, 1129-1133.
13. S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang, *Science*, 2011, **333**, 1300-1303.
14. Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He and G. L. Xu, *Science*, 2011, **333**, 1303-1307.
15. T. Pfaffeneder, B. Hackner, M. Truss, M. Munzel, M. Muller, C. A. Deiml, C. Hagemeyer and T. Carell, *Angew Chem Int Ed Engl*, 2011, **50**, 7008-7012.
16. Z. Gong, T. Morales-Ruiz, R. R. Ariza, T. Roldan-Arjona, L. David and J. K. Zhu, *Cell*, 2002, **111**, 803-814.
17. Y. Choi, M. Gehring, L. Johnson, M. Hannon, J. J. Harada, R. B. Goldberg, S. E. Jacobsen and R. L. Fischer, *Cell*, 2002, **110**, 33-42.
18. T. Morales-Ruiz, A. P. Ortega-Galisteo, M. I. Ponferrada-Marin, M. I. Martinez-Macias, R. R. Ariza and T. Roldan-Arjona, *Proc Natl Acad Sci U S A*, 2006, **103**, 6853-6858.
19. A. Maiti and A. C. Drohat, *J Biol Chem*, 2011, **286**, 35334-35338.
20. V. L. Schramm, *Accounts Chem. Res.*, 2003, **36**, 588-596.
21. R. D. Guthrie and W. P. Jencks, *Accounts Chem. Res.*, 1989, **22**, 343-349.
22. P. J. Berti and K. S. E. Tanaka, *Adv. Phys. Org. Chem.*, 2002, **37**, 239-314.
23. R. M. Werner and J. T. Stivers, *Biochemistry*, 2000, **39**, 14054-14064.
24. X. Y. Chen, P. J. Berti and V. L. Schramm, *J. Am. Chem. Soc.*, 2000, **122**, 6527-6534.
25. J. A. B. McCann and P. J. Berti, *J Am Chem Soc*, 2008, **130**, 5789-5797.
26. J. A. B. McCann and P. J. Berti, *J. Am. Chem. Soc.*, 2007, **129**, 7055-7064.

27. S. J. Admiraal and P. J. O'Brien, *Biochemistry*, 2010, **49**, 9024-9026.
28. D. W. Parkin, H. B. Leung and V. L. Schramm, *J Biol Chem*, 1984, **259**, 9411-9417.
29. F. Mentch, D. W. Parkin and V. L. Schramm, *Biochemistry*, 1987, **26**, 921-930.
30. X. Y. Chen, P. J. Berti and V. L. Schramm, *J Am Chem Soc*, 2000, **122**, 1609-1617.
31. A. R. Dinner, G. M. Blackburn and M. Karplus, *Nature*, 2001, **413**, 752-755.
32. S. S. Parikh, G. Walcher, G. D. Jones, G. Slupphaug, H. E. Krokan, G. M. Blackburn and J. A. Tainer, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 5083-5088.
33. A. C. Drohat, G. Xiao, M. Tordova, J. Jagadeesh, K. W. Pankiewicz, K. A. Watanabe, G. L. Gilliland and J. T. Stivers, *Biochemistry*, 1999, **38**, 11876-11886.
34. A. C. Drohat, J. Jagadeesh, E. Ferguson and J. T. Stivers, *Biochemistry*, 1999, **38**, 11866-11875.
35. A. C. Drohat and J. T. Stivers, *J. Am. Chem. Soc.*, 2000, **122**, 1840-1841.
36. Y. L. Jiang, A. C. Drohat, Y. Ichikawa and J. T. Stivers, *J. Biol. Chem.*, 2002, **277**, 15385-15392.
37. M. A. Bianchet, L. A. Seiple, Y. L. Jiang, Y. Ichikawa, L. M. Amzel and J. T. Stivers, *Biochemistry*, 2003, **42**, 12455-12460.
38. R. Shapiro and M. Danzig, *Biochemistry*, 1972, **11**, 23-29.
39. R. Shapiro and S. Kang, *Biochemistry*, 1969, **8**, 1806-1810.
40. P. A. Schwartz, M. J. Veticatt and V. L. Schramm, *J Am Chem Soc*, 2010, **132**, 13425-13433.
41. M. R. Birck and V. L. Schramm, *J Am Chem Soc*, 2004, **126**, 2447-2453.
42. J. A. Zoltewicz, D. F. Clark, T. W. Sharpless and G. Grahe, *J Am Chem Soc*, 1970, **92**, 1741-1749.
43. S. Sharma and J. K. Lee, *J. Org. Chem.*, 2002, **67**, 8360-8365.
44. M. C. Ho, M. B. Sturm, S. C. Almo and V. L. Schramm, *Proc Natl Acad Sci U S A*, 2009, **106**, 20276-20281.
45. Y. Kim and J. D. Robertus, *Protein Eng*, 1992, **5**, 775-779.
46. A. F. Monzingo and J. D. Robertus, *J Mol Biol*, 1992, **227**, 1136-1145.
47. A. Z. Michelson, A. Rozenberg, Y. Tian, X. Sun, J. Davis, A. W. Francis, V. L. O'Shea, M. Halasyam, A. H. Manlove, S. S. David and J. K. Lee, *J Am Chem Soc*, 2012, **134**, 19839-19850.
48. Y. Guan, R. C. Manuel, A. S. Arvai, S. S. Parikh, C. D. Mol, J. H. Miller, S. Lloyd and J. A. Tainer, *Nat Struct Biol*, 1998, **5**, 1058-1064.
49. S. Lee and G. L. Verdine, *Proc Natl Acad Sci U S A*, 2009, **106**, 18497-18502.
50. M. K. Brinkmeyer, M. A. Pope and S. S. David, *Chem Biol*, 2012, **19**, 276-286.
51. S. L. Porello, S. D. Williams, H. Kuhn, M. L. Michaels and S. S. David, *J Am Chem Soc*, 1996, **118**, 10684-10692.
52. A. L. Livingston, V. L. O'Shea, T. Kim, E. T. Kool and S. S. David, *Nat Chem Biol*, 2008, **4**, 51-58.
53. A. W. Francis, S. A. Helquist, E. T. Kool and S. S. David, *J Am Chem Soc*, 2003, **125**, 16235-16242.
54. P. J. O'Brien and T. Ellenberger, *Biochemistry*, 2003, **42**, 12418-12429.
55. A. Y. Lau, M. D. Wyatt, B. J. Glassner, L. D. Samson and T. Ellenberger, *Proc Natl Acad Sci U S A*, 2000, **97**, 13573-13578.
56. A. Z. Michelson, M. Chen, K. Wang and J. K. Lee, *J Am Chem Soc*, 2012, **134**, 9622-9633.
57. K. H. Scheller and H. Sigel, *J. Am. Chem. Soc.*, 1983, **105**, 3005-3014.
58. M. Liu, M. Xu and J. K. Lee, *J. Org. Chem.*, 2008, **73**, 5907-5914.
59. X. J. Sun and J. K. Lee, *J. Org. Chem.*, 2007, **72**, 6548-6555.
60. C. Cao, K. Kwon, Y. L. Jiang, A. C. Drohat and J. T. Stivers, *J. Biol. Chem.*, 2003, **278**, 48012-48020.

61. A. C. Drohat, K. Kwon, D. J. Krosky and J. T. Stivers, *Nat. Struct. Biol.*, 2002, **9**, 659-664.
62. A. H. Metz, T. Hollis and B. F. Eichman, *EMBO J.*, 2007, **26**, 2411-2420.
63. K. G. Berdal, R. F. Johansen and E. Seeberg, *EMBO J.*, 1998, **17**, 363-367.
64. P. J. O'Brien and T. Ellenberger, *J. Biol. Chem.*, 2004, **279**, 9750-9757.
65. A. Prakash, S. Doublet and S. S. Wallace, *Prog. Mol. Biol. Transl. Sci.*, 2012, **110**, 71-91.
66. J. C. Fromme and G. L. Verdine, *J. Biol. Chem.*, 2003, **278**, 51543-51548.
67. Y. H. Jang, W. A. Goddard, K. T. Noyes, L. C. Sowers, S. Hwang and D. S. Chung, *Chem. Res. Toxicol.*, 2002, **15**, 1023-1035.
68. K. Bialkowski, P. Cysewski and R. Olinski, *Z. Naturforsch. C*, 1996, **51**, 119-122.
69. P. A. van der Kemp, D. Thomas, R. Barbey, R. de Oliveira and S. Boiteux, *Proc Natl Acad Sci U S A*, 1996, **93**, 5197-5202.
70. H. M. Nash, R. Lu, W. S. Lane and G. L. Verdine, *Chem Biol*, 1997, **4**, 693-702.
71. S. D. Bruner, D. P. Norman and G. L. Verdine, *Nature*, 2000, **403**, 859-866.
72. D. O. Zharkov, T. A. Rosenquist, S. E. Gerchman and A. P. Grollman, *J Biol Chem*, 2000, **275**, 28607-28617.
73. J. C. Fromme, S. D. Bruner, W. Yang, M. Karplus and G. L. Verdine, *Nat Struct Biol*, 2003, **10**, 204-211.
74. A. A. Sartori, S. Fitz-Gibbon, H. Yang, J. H. Miller and J. Jiricny, *EMBO J*, 2002, **21**, 3182-3191.
75. L. Dong, R. Mi, R. A. Glass, J. N. Barry and W. Cao, *DNA Repair (Amst)*, 2008, **7**, 1962-1972.
76. R. Mi, L. Dong, T. Kaulgud, K. W. Hackett, B. N. Dominy and W. Cao, *J Mol Biol*, 2009, **385**, 761-778.
77. H. W. Lee, A. R. Brice, C. B. Wright, B. N. Dominy and W. Cao, *J Biol Chem*, 2010, **285**, 41483-41490.
78. H. W. Lee, B. N. Dominy and W. Cao, *J Biol Chem*, 2011, **286**, 31282-31287.
79. K. N. Rogstad, Y. H. Jang, L. C. Sowers and W. A. Goddard, 3rd, *Chem Res Toxicol*, 2003, **16**, 1455-1462.
80. K. B. Roy and H. T. Miles, *Nucleosides Nucleotides*, 1983, **2**, 231-242.
81. V. Vongchampa, M. Dong, L. Gingipalli and P. Dedon, *Nucleic Acids Res*, 2003, **31**, 1045-1051.
82. G. E. Wuenschell, T. R. O'Connor and J. Termini, *Biochemistry*, 2003, **42**, 3608-3616.
83. M. T. Bennett, M. T. Rodgers, A. S. Hebert, L. E. Ruslander, L. Eisele and A. C. Drohat, *J Am Chem Soc*, 2006, **128**, 12510-12519.
84. R. Savva, K. McAuley-Hecht, T. Brown and L. Pearl, *Nature*, 1995, **373**, 487-493.
85. A. C. Drohat and J. T. Stivers, *Biochemistry*, 2000, **39**, 11865-11875.
86. J. E. Wibley, T. R. Waters, K. Haushalter, G. L. Verdine and L. H. Pearl, *Mol. Cell*, 2003, **11**, 1647-1659.
87. A. Van Schepdael, N. Ossembe, P. Herdewijn, E. Roets and J. Hoogmartens, *J. Pharm. Biomed. Anal.*, 1993, **11**, 345-351.
88. W. P. Jencks, *Cold Spring. Harb. Symp. Quant. Biol.*, 1972, **36**, 1-11.
89. F. Hollfelder and D. Herschlag, *Biochemistry*, 1995, **34**, 12255-12264.
90. C. J. La Francois, Y. H. Jang, T. Cagin, W. A. Goddard and L. C. Sowers, *Chem. Res. Toxicol.*, 2000, **13**, 462-470.
91. A. Maiti, A. Z. Michelson, C. J. Armwood, J. K. Lee and A. C. Drohat, *J Am Chem Soc*, 2013, **135**, 15813-15822.
92. I. Wempen and J. J. Fox, *J. Med. Chem.*, 1963, **122**, 688-693.
93. I. Wempen and J. J. Fox, *J. Am. Chem. Soc.*, 1964, **86**, 2474-2477.
94. B. Kavli, G. Slupphaug, C. D. Mol, A. S. Arvai, S. B. Peterson, J. A. Tainer and H. E. Krokan, *EMBO J*, 1996, **15**, 3442-3447.
95. K. Kwon, Y. L. Jiang and J. T. Stivers, *Chem. Biol.*, 2003, **10**, 351-359.
96. S. C. Brooks, R. L. Fischer, J. H. Huh and B. F. Eichman, *Biochemistry*, 2014, **53**, 2525-2532.

97. R. J. O'Neill, O. V. Vorob'eva, H. Shahbakhti, E. Zmuda, A. S. Bhagwat and G. S. Baldwin, *J Biol Chem*, 2003, **278**, 20526-20532.
98. S. Morera, I. Grin, A. Vigouroux, S. Couve, V. Henriot, M. Sapparbaev and A. A. Ishchenko, *Nucleic Acids Res*, 2012, **40**, 9917-9926.
99. C. Hansch, A. Leo and R. W. Taft, *Chem. Rev.*, 1991, **91**, 165-195.
100. M. Sumino, A. Ohkubo, H. Taguchi, K. Seio and M. Sekine, *Bioorg Med Chem Lett*, 2008, **18**, 274-277.
101. A. Maiti and A. C. Drohat, *DNA Repair*, 2011, **10**, 545-553.