THEORETICAL AND MASS SPECTROMETRIC STUDIES OF DAMAGED

NUCLEOBASES AND ANALOGS TOWARD UNDERSTANDING

GLYCOSYLASE MECHANISMS

by

ANNA MICHELSON

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Chemistry and Chemical Biology

written under the direction of

Professor Jeehiun K. Lee

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2012

ABSTRACT OF THE DISSERTATION

THEORETICAL AND MASS SPECTROMETRIC STUDIES OF DAMAGED NUCLEOBASES AND ANALOGS TOWARD UNDERSTANDING GLYCOSYLASE MECHANISMS

By ANNA MICHELSON

Dissertation Director:
Professor Jeehiun K. Lee

The focus of this thesis is the examination of the thermochemical properties, primarily the gas phase acidity, proton affinity, and leaving group (LG) ability of damaged nucleobases and related species via mass spectrometry (FT-ICR, ion trap) and theoretical studies (quantum mechanical calculations). Our main hypothesis is that the study of intrinsic, gas-phase properties of the damaged nucleobases will lend insight into the mechanism of their excision from DNA. We study damaged nucleobases and analogs that are cleaved from DNA by various glycosylases: uracil-DNA glycosylase (UDG), 3-methyladenine glycosylase II (AlkA), and MutY glycosylase.

The LG ability of the N1-deprotonated 3-methyluracil anion relative to the N1-deprotonated 3-methylthymine anion is examined in the context of the UDG enzymatic

reaction that excises uracil but not thymine from DNA. We confirmed that despite the close acidities uracil is a much better LG in the gas phase. Another interesting disparity between the LG ability and acidity is discovered for uracil substrates: when we examined hydrochloric acid and 3-methyluracil in the gas phase we found that despite similar acidities, chloride is a better LG than N1-deprotonated 3-methyluracil. We propose that the difference in LG ability is due to the different natures of the LGs (resonance vs. inductive stabilization). To test the hypothesis, a series of pyridone substrates were designed and examined.

AlkA is an enzyme that cleaves a wide range of damaged bases from DNA. Herein we examine 3- and 7-methylated AlkA purine substrates. The damaged nucleobases are found to be more acidic than the normal nucleobases. Because of this increased acidity, the damaged bases would be expected to be more easily cleaved from DNA by AlkA (their conjugate bases should be better LGs). We find that the acidity correlates to the AlkA excision rates, which lends support to an AlkA mechanism wherein the enzyme provides a nonspecific active site, and nucleobase cleavage is dependent on the intrinsic N-glycosidic bond stability.

The acidities and proton affinities of adenine and six adenine analogs that were designed to test various features of the enzyme MutY are also studied to allow better understanding of the mechanism of adenine removal by MutY.

DEDICATION

To my wonderful parents, Emil and Natalia, and to my dear husband Alexander

for their love, understanding, and endless support

ACKNOWLEDGEMENTS

I would like to thank my parents Natalia Zhachkina and Emil Dovidovich for their love and endless support not only through the years of my Ph.D. work but all my life, and for encouragement in my choosing chemistry career path.

And last but not least, I thank my beloved husband Alexander who was along my side all these graduate school years. I would never be able to do it without your constant support and understanding!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. Introduction

1.1. Overview

1.1.1. Normal and damaged nucleobases in DNA

Keeping the integrity of DNA, the repository of almost all organisms' genetic information, is essential to life.[1]

DNA, deoxyribonucleic acid, consists of two complementary polymeric chains (strands) composed of monomeric units called nucleotides; two chains form a double helix. Sugar (deoxyribose) molecules are linked together by phosphate groups via phosphodiester bonds, and form the backbone. A nucleic base is covalently attached to each sugar via an N-glycosidic bond (**Figure 1.1**). [2]
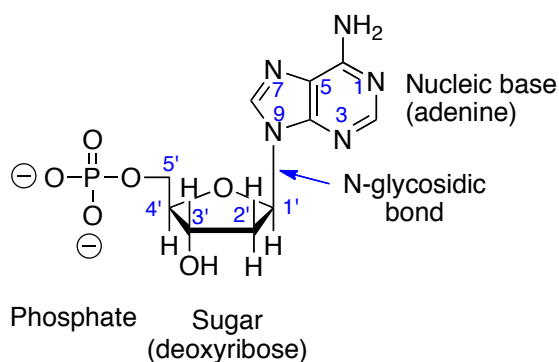
Figure 1.1. Nucleotide structure (with adenine as nucleic base as an example).

There are four nucleic bases that are normally present in DNA: two purines (adenine (A) and guanine (G)) and two pyrimidines (thymine (T) and cytosine (C)). Uracil (U), another pyrimidine, is mutagenic when it arises in DNA but normally occurs in RNA

(**Figure 1.2**; nucleobases are connected to sugar via N9-site (for purines) or N1-site (for pyrimidines)).
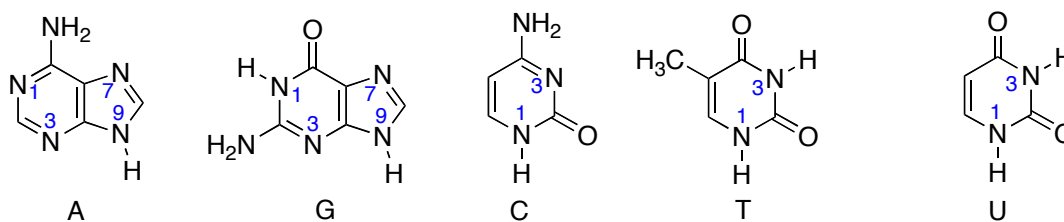
Figure 1.2. Structures of nucleobases common in DNA (A, G, C, T), and RNA base (U).

The sequence of nucleic bases is how information is encoded in DNA. Both endogeneous (cellular metabolites) and exogeneous agents (environmental mutagens, such as various chemicals, UV light, and ionizing radiation) constantly attack DNA, and cause nucleic base modification, and other types of DNA damage (like N-glycosidic bond hydrolysis, DNA strand breakage, the collapse of replication forks, etc.). The most common types of *nucleic base* damage are base deamination, alkylation, and oxidation, which lead to formation of so-called modified or damaged nucleic bases (**Figure 1.3**). These DNA lesions, left unrepaired, may interfere with DNA replication and transcription, and lead to mutation, carcinogenesis, aging, and cell death.[3-7]
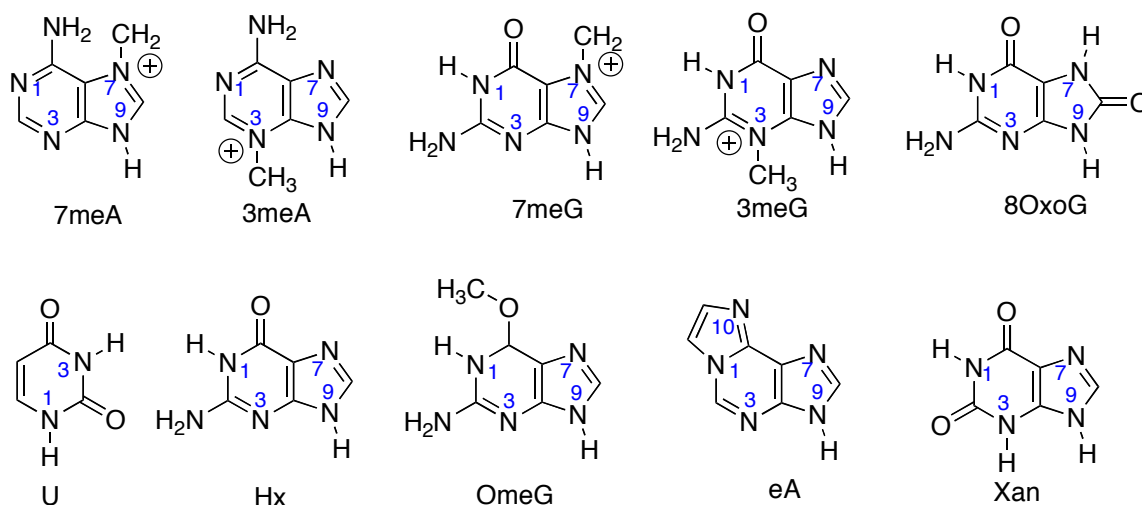
Figure 1.3. Examples of modified DNA bases.

1.1.2. DNA glycosylases and base excision repair

The base excision repair (BER) pathway is the main and most frequently used DNA repair mechanism in nature. BER is a complicated multistep process initiated by DNA repair enzymes called DNA glycosylases, which excise damaged (altered) nucleobases from DNA in free base form. The removal of an improper base is followed by subsequent incision of sugar phosphate backbone at the abasic site by endonuclease, removal of DNA terminus, and filling of the resulting gap by action of DNA polymerase and ligase.[4,8]

This study is focused on the mechanism of damaged nucleobase removal by several DNA glycosylases, and therefore the BER initial step is of our interest.

As the name suggests, DNA glycosylases are a family of DNA repair enzymes, which catalyze the hydrolysis of the N-glycosidic bond (**Figure 1.1**), the connection between the damaged base and the sugar. Monofunctional DNA glycosylases only remove the

base, while bifunctional glycosylases also have additional lyase activity (cleave DNA 3'

of the abasic site).[5]

Each organism has its own set of enzymes, therefore a set of DNA glycosylases found

in *E. coli* is different from a set found in human cells. The most common human and

bacterial DNA glycosylases and examples of substrates are listed in **Table 1.1**.[4]

Table 1.1. Human and *E. coli* DNA glycosylases and its substrates[4]

DNA glycosylase		Examples of substrates
Abbr.	Name	
in E. coli		
Ung	Uracil-DNA glycosylase	U
Mug	Mug-DNA glycosylase	U, T, or ethenocytosine opposite G
Fpg (MutM)	FaPy-DNA glycosylase	Oxidized and ring- opened purines (8-oxoG)
MutY	MutY-DNA glycosylase	A opposite 8-oxoG
Nth and Nei	Endonucleases III and VIII	Ring-saturated and fragmented pyrimidines
TagA (Tag)	3-Methyladenine-DNA glycosylase I	3-Methyladenine, 3-ethyladenine
AlkA	3-Methyladenine-DNA glycosylase II	3- and 7-methylpurines, 7-ethylpurines, eA, O^2-methylpyrimidines
in human cells		
UNG (UDG)	Uracil-DNA glycosylase	U
SMUG1	SMUG DNA glycosylase	U, 5-hydroxymethylU
MBD4 (MED1)	Methyl-binding domain glycosylase 4	U or T opposite G at cpG sequences, T opposite O^6-meG
TDG	Thymine-DNA glycosylase	U, T, or etheno-C opposite G
OGG1	8-oxoG-DNA glycosylase	8-oxoG, and other oxidized and ring-opened purines
MYH	MutY homolog DNA	A opposite 8-oxoG, 2-OH-A

	glycosylase	opposite G
MPG (AAG)	3-Methyladenine-DNA glycosylase I	3-Methylpurines, Hx, eA
NTHL1	Endonucleases III	Ring-saturated and fragmented pyrimidines
NEIL1, 2, and 3	Endonuclease-VIII-like DNA glycosylase 1, 2, and 3	Oxidized and fragmented pyrimidines, 8oxoG

Uracil-DNA glycosylase (UDG), 3-Methyladenine-DNA glycosylase II (AlkA), and MutY-DNA glycosylase (MutY) are three enzymes whose mechanisms we will discuss herein. Though all of these enzymes and their mechanisms have been studied for decades there are still many puzzling questions regarding specificity, ability to discriminate damaged and normal nucleobases, and some other mechanistic questions.

Uracil-DNA glycosylase (UDG) is the primary enzyme for the removal of uracil from DNA.[4,9] Uracil is a normal constituent of RNA. When it occurs in DNA uracil is not directly mutagenic (since U is able to pair with A as effectively as T can) however its presence affects the viability of cells. To understand why uracil is damaging when it occurs in DNA, one should look at its origin. Uracil arises in DNA from incorporation of dUTP, from exposure to DNA-damaging agents (such as nitrous acid or ionizing radiation), and spontaneous deamination of cytosine. If cytosine is converted to uracil it leads to the mispairing of U:G, which will lead to a CG to TA mutation after subsequent replication. Also the presence of high number of uracil residues in DNA changes the secondary structure of DNA, and affects important DNA-protein interactions.[4] Therefore the role of UDG is critical in maintaining DNA structure and the health of cells. One

should also note that UDG is highly specific; it cleaves uracil from DNA while leaving structurally similar thymine untouched.[10]

It has been suggested that UDG binds, kinks, and compresses the backbone of a DNA duplex while scanning the minor groove for a uracil residue. The enzyme uses a "push-pull" mechanism to extract the uracil nucleotide from the DNA base stack and positions it into its the active site. The energetic destabilization of the stacked conformation is equivalent to a "push", while the energetic stabilization of unstacked conformation is equivalent to a "pull".[9,10] Then, UDG hydrolyzes the N-glycosidic bond that links uracil to the deoxyribose backbone, leaving an abasic site, which can be repaired by the base excision repair pathway.[11]

Based on the X-ray crystal structure of enzyme-bound uracil, the mechanism of hydrolysis was proposed (**Scheme 1.1**).[12] The acidic residues (e.g. Asp 88) of the enzyme can activate water, and make it more nucleophilic. Water attacks C1' carbon of the ribose ring, and the uracil anion leaves.[12,13] This mechanism is believed to be a dissociative S_N2 (S_N1-like) process.[14]
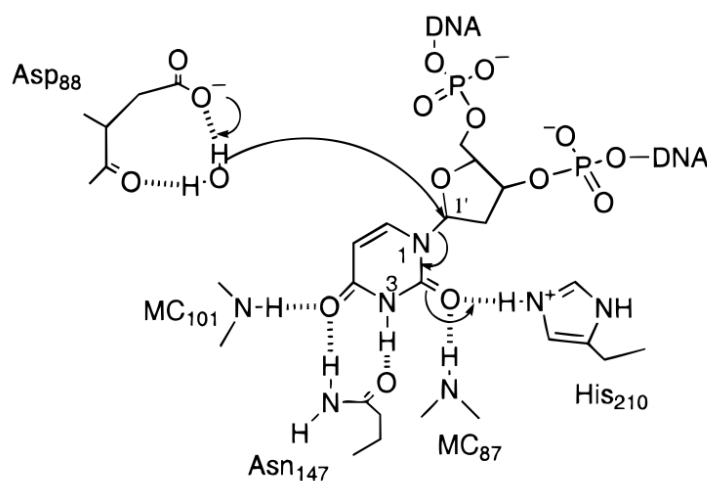
Scheme 1.1. A possible mechanism of uracil removal by UDG[12]

The structural basis for the exquisite recognition of uracil by UDG has been explored.[15,16] Based on the structure of human UDG bound to DNA it has been proposed that thymine is prevented from entering the binding site by the Tyr 147 residue of the enzyme, which packs against uracil C5. Since thymine has a methyl group at C5, the proposal is that thymine is too bulky to fit into the active site.[17,18] However, to our best knowledge, no binding studies of UDG with thymine have been done. So, it is actually unclear whether UDG binds to thymine but does not excise it from DNA, or if thymine just does not enter the enzyme's active site.

We will try to get insight into the puzzling behavior of UDG: How does it cleave the damaged base uracil from DNA, leaving normal thymine (which is present in DNA in great excess) untouched?

In contrast to the highly specific UDG, 3-methyladenine-DNA glycosylase II (AlkA) is an enzyme with very broad specificity.[4] In addition to 3-methyladenine AlkA catalyzes the excision of a wide variety of damaged bases.[19] Along with 3meA it is capable of the removal of a broad range of 3- and 7-alkylpurines (7meG, 7meA, 3meG), cyclic purine derivatives (like $1,N^6$-ethenoadenine), O^2-alkylated pyrimidines (O^2-methylthymine, O^2-methylcytosine), hypoxanthine, xanthine, oxanine and some other damaged bases in both the major and minor grooves of DNA.[4,20] Undamaged (normal) nucleic bases may also be excised but much less efficiently.[20,21]

The broad specificity of AlkA is especially surprising taking into account the ability of this enzyme to discriminate against normal bases. Normal bases are usually smaller than damaged bases, and therefore should be excised with the rates comparable to the rates of damaged base excision (if not faster) by a non-specific enzyme like AlkA, which is able

to accommodate variety of "shapes" of damaged bases (from purines to pyrimidines). It is interesting to note that AlkA can remove not only charged bases but neutral ones too. What is the mechanism of excision? What influences the specificity and discrimination for this enzyme? It is known that it has relatively open binding pocket.[21,22] The accepted theory is that the rate of excision is dictated *not* by catalytic recognition (size of substrate, shape of enzyme binding pocket) but by the reactivity of N-glycosidic bond of each substrate.[6,20,23] Alkylated bases and other AlkA substrates have decreased N-glycosidic bond stability, and are therefore readily excised. AlkA destabilizes the unstable N-glycosidic bond further (by stabilizing the TS for hydrolysis), and therefore provides preferential repair of substrates with weak N-glycosidic bonds.[6,20]

The question we will try to clarify in current studies is how AlkA removes such a broad range of damaged bases and does not cleave normal ones.

MutY-DNA glycosylase (MutY) is a unique glycosylase: it removes normal base adenine but only when it mispaired with 8-oxoguanine (8OxoG), preventing its potential mutagenic consequences.[4,6] 8OxoG is the most studied among oxidized guanine products; it is even used as a cellular biomarker of oxidative stress.[24-28] Along with being the most common oxidative stress product, 8OxoG also is one of the most deleterious lesions. The reason behind this harmfulness is structural, i.e. introduction of oxygen at C8 atom and addition of H at N7. These two structural changes allow OG in syn- conformation to mimic thymine (T) during replication (**Figure 1.2**, N7H of OG mimics N1H of T, and O6 of OG mimics O4 of T), and form stable OG (syn)-A (anti) base pair (**Figure 1.4**).[28] The failure to remove the OG causes the permanent G-C to T-A transversion mutation.
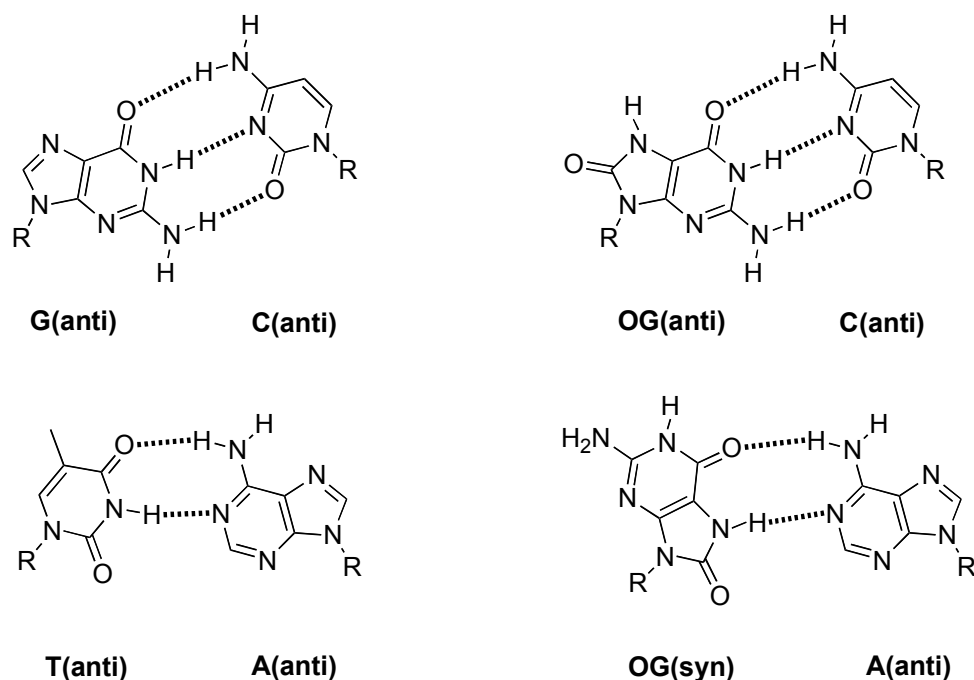
Figure 1.4. Normal and mutagenic base pairs relevant to adenine (A) removal by MutY

Concerted S_N2 and stepwise S_N1 mechanisms have been proposed earlier for adenine removal by MutY.[29] In both mechanisms the adenine leaving group is assumed to be protonated. The [15]N7 KIE observed for the *E. coli* MutY adenine excision is in agreement with N7 protonation.[29] Furthermore, TS analysis of MutY substrates indicates that protonation lowers the TS barrier by 7 kcal mol[-1] (corresponding to a 10^5 fold excision rate acceleration). All those studies, however, do not prove that N7, and only N7 site is protonated.[29,30] In fact, the mechanistic studies of MutY reveal that adenine should be involved in multiple hydrogen bonding interactions.[30] In particular, fluorinated lesion-recognition complex (FLRC) crystal structure has been solved, and it demonstrated that in this FLRC complex the substrate is buried deeper inside of the pocket, and has multiple direct contacts. Glu43 and Tyr126 residues contact N7; Arg31 hydrogen bonds to N1 of substrate; Glu188 hydrogen bonds to N6 and, Arg26 side chain forms water

mediated contact with N3. Hydrophobic side chains (Ile191, Leu46, Val51, Leu28, Trp30) surround extrahelical adenine, providing water exclusion and additional stabilization.[30] This data supports the possibility of multiple site protonation and/or hydrogen bonding of adenine in the MutY active site.

Our studies herein explore the possibilities of various sites of protonation for adenine in the MutY site through model studies of unnatural MutY substrates. This project is done in collaboration with Prof. Sheila David (UC Davis, CA).

1.1.3. Gas phase acidity and proton affinity of nucleobases and related compounds

The interior of proteins, in particular the interior of active sites, is rarely aqueous in nature but rather nonpolar.[31,32] The gas phase is the "ultimate" nonpolar environment, which therefore can provide a valuable environment for the modeling of intrinsic reactivity, and reactivity inside of an enzyme's active site.

The gas-phase acidities and proton affinities are largely unknown in contrast to solution values, pK_as. The gas phase acidity (ΔG_{acid} or ΔH_{acid}) for a compound AH is defined as the Gibbs energy change, or the enthalpy change of the deprotonation reaction yielding H^+ and A^- (**Eq. 1.1**). The proton affinity (PA) is defined as the negative value of the enthalpy change associated with protonation of a chemical species B to form HB^+ (**Eq. 1.2**). The gas-phase basicity (GB) is the corresponding Gibbs energy value. The gas-phase acidities (ΔH_{acid}) and PAs at 298 K are mostly used.[33,34]

$$HA \rightarrow H^+ + A^- \qquad\qquad \textbf{Eq. 1.1}$$

$$B + H^+ \rightarrow HB^+ \qquad\qquad \textbf{Eq. 1.2}$$

The gas-phase acidity (ΔH_{acid}) scale runs approximately from 314 (strong acids, like HI) to 417 kcal mol^{-1} (alkanes). The higher the ΔH_{acid} value, the lower the acidity is. The proton affinity varies approximately from 130 (alkanes) to 291 kcal mol^{-1} (inorganic oxides, like BaO); the higher the value, the more basic the compound.[33]

In previous work, our research group has reported the gas phase thermochemical properties (both experimental and computational) of various normal[13,35-39] and a few damaged nucleobases (like 3-methyladenine,[36,37] 1,N^6-etheonadenine,[40] and hypoxanthine[41]). This work continues the survey of the damaged base (and their analogues) gas phase thermochemical properties, in the hope to elucidate acidity/PA trends, and to illuminate the mechanism of the removal of damaged nucleobases by various enzymes.

1.1.4. Proton transfer (PT) and S_N2 reactions in the gas phase

In order to evaluate proton affinities and acidities of nucleobases or model damaged nucleobase removal by the glycosylase enzyme (attack of water on C1' of sugar, and damaged nucleobase departing) one should study the corresponding gas phase reactions (PT or S_N2 reactions).

Eq. 1.3 and **Eq. 1.4** represent PT reactions between a cation or an anion correspondingly, and a neutral molecule.

$$HA^+ + B \rightleftharpoons HB^+ + A \qquad \textbf{Eq. 1.3}$$

$$HA + B \rightleftharpoons HB + A^- \qquad \textbf{Eq. 1.4}$$

Lets consider an example of the gas phase exothermic PT reaction between a neutral molecule HA and anion B$^-$ (**Eq. 1.4**). In the initial step, when the ion and molecule get close to each other, and collide, they form an ion-molecule complex [HA B$^-$] with low

energy (at the bottom of the potential energy well; **Figure 1.5**).[42] The energy of formation

of such complexes is roughly from 10 to 20 kcal mol^{-1} (for complexes involving ion-

dipole and ion-induced dipole interactions). This excess of energy is enough to overcome

the possible intrinsic barrier for PT between an ion and molecule, and makes the overall

PT reaction virtually barrierless.[43]. PT reaction leads to formation of the second ion-

molecular complex [HB A$^-$], and its separation to products HB and A$^-$ (which is of

course "uphill from the ion-molecule complex; **Figure 1.5**).[44,45]
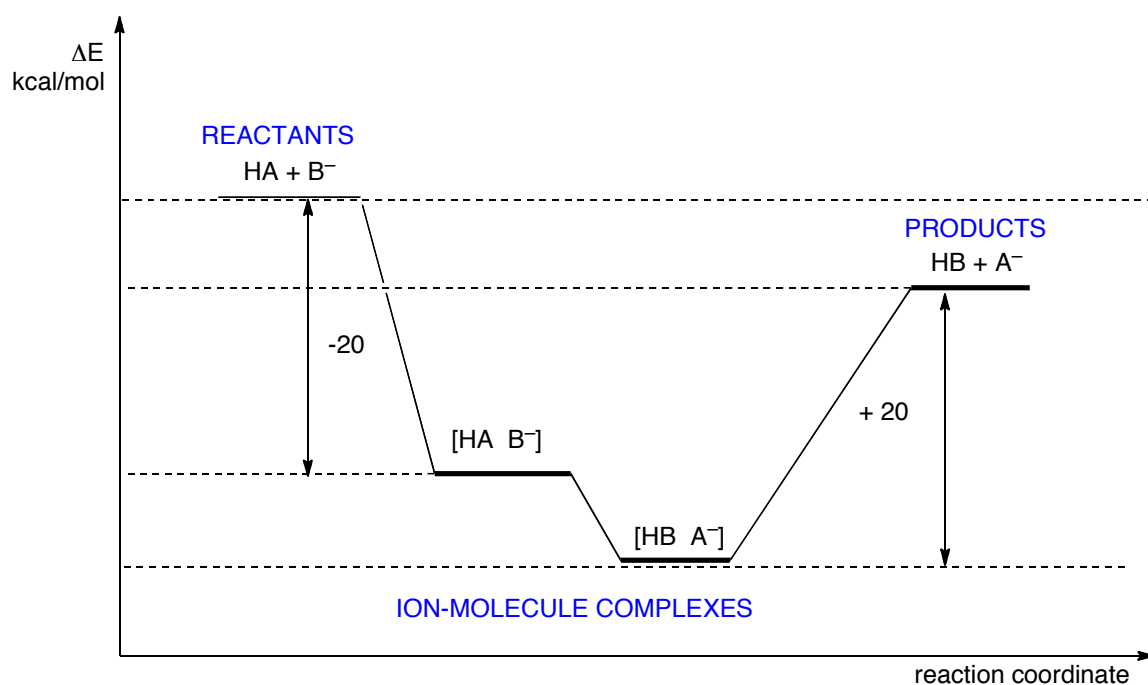
Figure 1.5. Representative potential energy diagram for exothermic proton transfer

reaction (**Eq. 1.4**).

The overall barrierless feature of PT reactions is important, as it allows one to elucidate

thermochemical properties of the compounds by measuring kinetics of the corresponding

reactions (more in **Methodology** section).[43]

The collision of an ion with a molecule in the gas phase may result in one or several of the possible pathways, such as multiple PT reactions, elimination, substitution, or dissociation of the ion-molecule complex back into reactants.

In the current work we study S_N2 reactions in the gas phase both theoretically and experimentally.

S_N2 reactions in the gas phase have drawn attention for years and have been the topic of numerous gas phase studies.[46] Unlike the "bell" shaped solution potential energy surface (PES),[47] the PES for S_N2 reaction in the gas phase is "double-well" shaped.[46,48] The "double-well" PES was introduced in 1977 by Olmsted and Brauman.[49] When the reactants (substrate and nucleophile) approach each other the energy drops due to ion-dipole interactions, and eventually a complex is formed (similar to ion-molecular complex formation in gas phase PT reactions). Then the energy rises up towards the transition state – the gas phase S_N2 reactions do have a barrier, unlike PT reactions. As the reaction progresses to product formation first another drop in energy occurs, and another complex is formed (now between the product ion and molecule). The rise in energy accompanies separation of the complex into final products (**Figure 1.6**).[46,49]
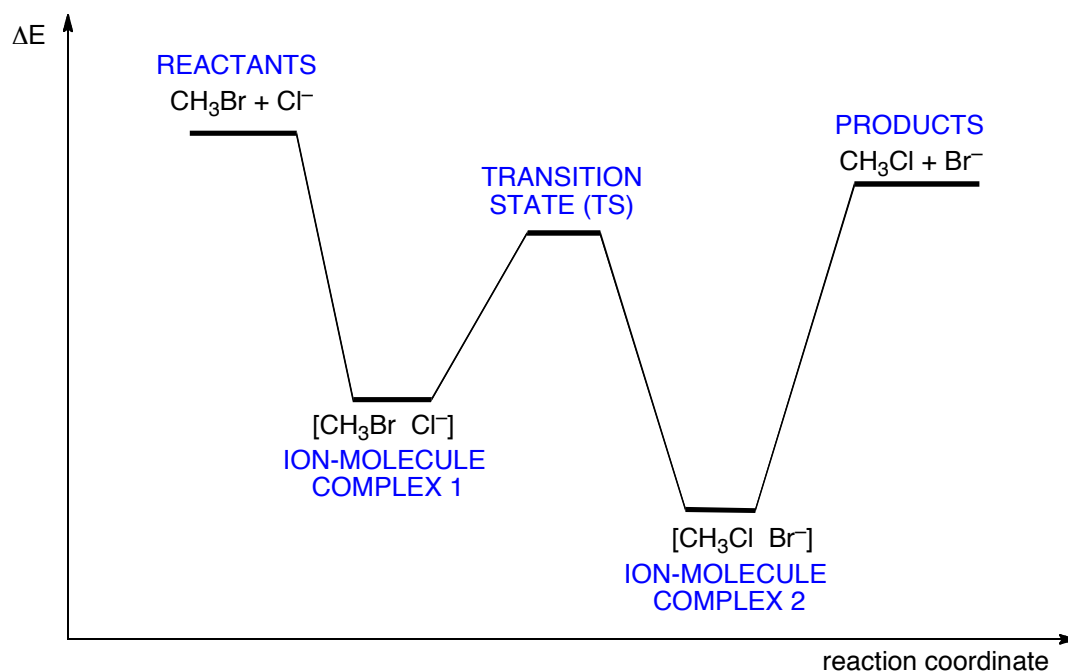
Figure 1.6. The "double-well" potential energy surface for the reaction of methylbromide with chloride anion (classical example of the gas phase S_N2 "double-well").[46]

Early studies focused on simple substrates such as methyl halides to preclude any complications or competition from elimination.[46,48-55] Rate constants and efficiencies for S_N2 reactions of methyl chloride with different anionic nucleophiles have been measured using both flowing afterglow (FA) and ion-cyclotron resonance (ICR) mass spectrometry (**Table 1.2**). The efficiencies measured by different techniques can vary. Rate constants and efficiencies obtained by the FA method are sometimes higher, which may be due to the fact that reactions are sometimes not fully thermalized under ICR conditions.[46,48]

Table 1.2. Rate constants and efficiencies for reactions with methyl chloride[46]

Nucleophile	PA, kcal mol^{-1}	Rate, x10^{-10} cm^3/molecule/s	Efficiency, %
NH$_2^-$	404	15	63

CH_3NH^-	403	17	85
Ph^-	402	8.7	54
H^-	400	30	35
OH^-	391	20	84
$O^{\cdot-}$	382	17	71
CH_3O^-	382	13	65
$PhCH_2^-$	381	0.15	1
HCC^-	378	1.3	6.2
$(CH_3)_3CO^-$	374	1.6	10
F^-	371	13	56
$CF_3CH_2O^-$	362	2.2	15
CH_3S^-	358	1.1	6.5
$O_2^{\cdot-}$	356	7.4	39
HS^-	351	0.12	0.6
$S^{\cdot-}$	351	0.3	1.6
Cl^-	333	0.00035	0.002

The careful gas-phase S_N2 studies previously accomplished show approximate correlation between proton affinity (PA) of the nucleophile and efficiency of the S_N2 process: the larger the PA, the higher the efficiency of the S_N2 gas phase reaction (**Table 1.2**).[46,50] There are also other factors that influence nucleophilicity, including charge delocalization and steric hindrance. For example, bulky *tert*-butoxide is much less efficient in S_N2 reactions than F^- though the latter is a weaker base.[50,53] Charge delocalization in the benzyl anion leads to low reactivity despite the relatively high basicity of this anion.[51] Carbon-centered nucleophiles are less reactive than nucleophiles of the same basicity with a more electronegative reaction center. This fact is explained by the relationship we observe between the electronegativity of the atom and its charge density: the more electronegative the atom, the higher the charge density, and consequently, the more compact its lone pairs.[46]

The choice of appropriate nucleophile is critical for the gas phase S_N2 studies. Though, in general, it seems that the strongest nucleophiles (and therefore *usually*, the most basic ones) are the best candidates, one should remember that basic nucleophiles can also

participate in proton transfer reactions, i.e. deprotonation of the acidic sites of the substrates. Proton transfer reactions are barrierless, and therefore fast, and will compete with S_N2 reactions. To avoid this complication one should always choose nucleophiles with proton affinities lower than the most acidic site of the substrate (see more on the choice of nucleophile in **Chapter 2**).

1.2. Instrumentation

Since the first experiments in the beginning of the 20th century mass spectrometry has come a long way, and emerged as one of the most powerful and widely used analytical tools.[56] High sensitivity, low detection limits, high resolution, ability to couple with various separation instruments, and diverse applicability are among the remarkable advantages of mass spectrometry.[56] On top of the advantage of numerous analytical applications mass spectrometry is also a powerful tool for the gas phase reaction studies, including kinetics, mechanisms, product distribution, and thermodynamic parameters elucidation.[46,56]

The development of new mass spectrometers (including flow instruments (e.g. selected-ion flow tube (SIFT)), trapping instruments (Fourier transform mass spectrometer (FTMS) and quadrupole ion trap mass spectrometer (QITMS)), and high-pressure mass spectrometers (HPMS)) allowed the investigation of gas phase experiments under highly controlled conditions.[46] In our lab, we use two types of mass spectrometers mentioned above: FTMS and QITMS. The details of the instrumentation, and advantages they provide for our studies are described below.

1.2.1. Fourier transform mass spectrometer (FTMS)

Fourier transform ion cyclotron resonance (FT-ICR), also known as the Fourier transform mass spectrometer (FTMS), is a type of mass analyzer for mass-to-charge ratio (m/z) determination based on the cyclotron frequency of the ions in a fixed magnetic field. FTICR mass spectrometry is known for its very high resolution, sensitivity, and accuracy, as well as capability of coupling with various ionization techniques (such as electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI)). The last feature mentioned allows the use of FTMS for studies of molecules with large molecular weight, which was impossible to get into the gas phase otherwise. Now applications of FTMS range from analysis of small molecules (such as drug metabolites or free nucleobases) to large biomolecules (such as peptides or oligonucleotides).[57]

Fourier transform mass spectrometry evolved from ion cyclotron resonance (ICR) spectrometry, which was first developed in 1930s by Lawrence[58] for the fundamental study of the atom, and then, in 1950s incorporated into mass spectrometer by Sommer[59] and co-workers. In 1974, inspired by recent development of FT-NMR technique, Comissarow and Marshall[60] applied the Fourier transform method to ICR-MS, and built the very first FTMS instrument.[57]

All FTMS instrument have four common features: a magnet, an analyzer cell (or cells), an ultra-high vacuum system and, a sophisticated data-processing system.[57] High operational cost (cryogens for a magnet), and necessity of an experienced operator made this complicated, but this highly sensitive and accurate instrument unjustifiably rare in non-academic environment.

Our custom-modified dual-cell Finnigan 2001 FTMS is equipped with 3.3 Tesla non-shielded superconducting magnet and vacuum system (mechanic and diffusion pumps) capable to pump to baseline pressure of 10^{-9} Torr. Two adjoined 2-inch cubic cells is the heart of the instrument. Our FTMS also features batch inlets, pulse and leak valves, and solids probe, which allow us to introduce several compounds into the cell simultaneously, and conduct the gas phase studies (**Figure 1.7**, see more details in **Methodology**).
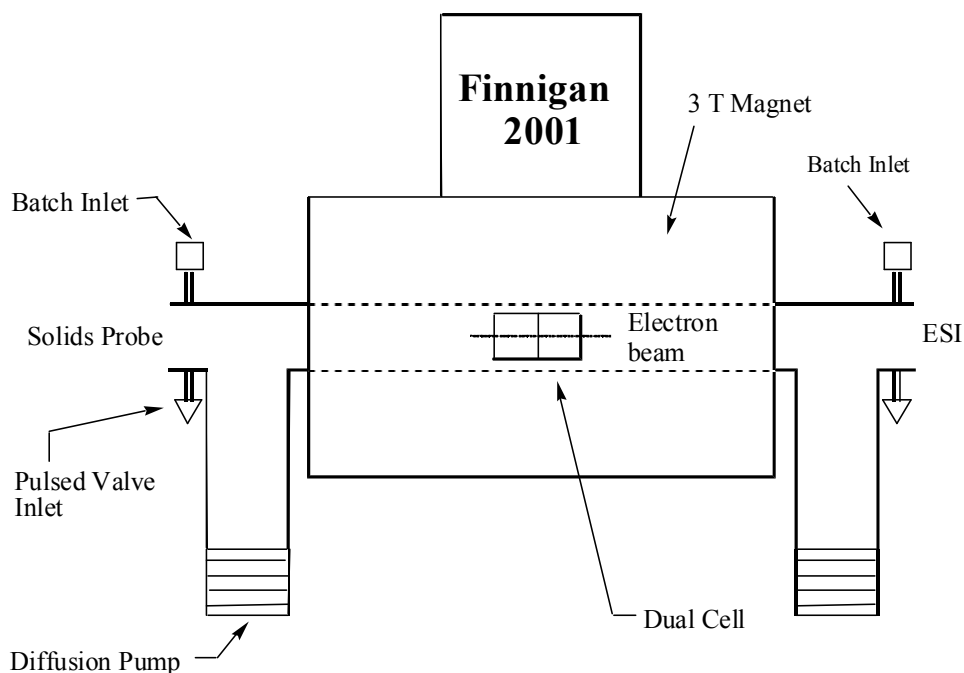
Figure 1.7. Finnigan 2001 FTMS

The theory behind FT-ICR is the principle of ion motion in a magnetic field, *B*. An ion moving in a uniform magnetic field B executes cyclotron motion in a plain perpendicular to the applied field (**Figure 1.8**). Herein, *v* is velocity, *r* is radius of the circular orbital, F_1 is Lorentz force (centripetal or magnetic force), F_2 is centrifugal force.
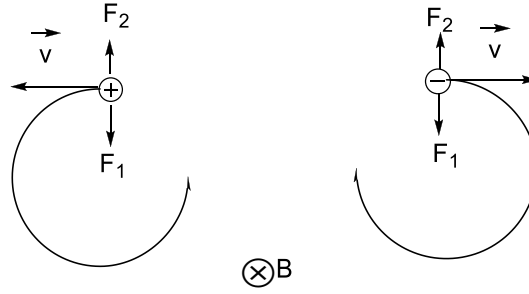
Figure 1.8. The motion of a charged ions (cation (on the left) and anion (on the right)) in a magnetic field, B[57,61]

$$F_1 = qvB \qquad \text{Eq. 1.5}$$

$$F_2 = mv^2/r \qquad \text{Eq. 1.6}$$

The ion motion can be described via **Eq.1.5-1.6**, where q is the charge on the ion, m is the mass of the ion. When the two forces are balanced (**Eq. 1.7**), ions can be stabilized on a circular trajectory with an ion cyclotron orbital frequency f (**Eq. 1.8-1.9**).[57,61]

$$qvB = mv^2/r \qquad \text{Eq. 1.7}$$

$$f = v/2\pi r \qquad \text{Eq. 1.8}$$

$$f = qB/2\pi m \qquad \text{Eq. 1.9}$$

The ions with the same mass-to-charge ratio rotate with the same cyclotron frequency, which depends only on the magnetic field strength (B) but not an ion velocity. However, the motion of ions in the plane *parallel* to the direction of the magnetic field is unconstrained. Without supplementary potential ions would drift along the magnetic field axis, which would create a detection problem. To resolve this issue McIver (1970) included a small trapping voltage into ICR cubic cell design: a small symmetric voltage is

applied to the pair of trapping plates (positive voltage for positive ions, negative – for negative; **Figure 1.9**) perpendicular to the direction of magnetic field, which allows to trap ions within a cell.[57]
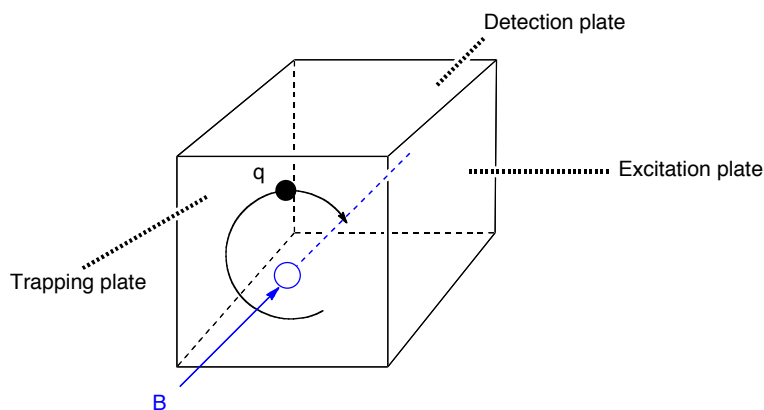
Figure 1.9. Scheme of cubic FT-ICR MS cell (each plate has a matching pair at the opposite end of the cube). [57,61]

Alternating current applied to a pair of excitation plates excites ions (cause the increase of their kinetic energy) and, therefore, increases the radius of their circular orbits. All ions with the same mass-to-charge ratio will be excited coherently. When the packet of excited ions passes the electrode (detection plate), it produces an alternating image current on the detection plates (**Figure 1.9**). The detected image current is then amplified and digitized. The frequency components of the signal (and corresponding mass-to-charge ratios) are obtained by applying a Fourier transform to the time domain signal.[57]

The details of FTMS bracketing and S_N2 experiments are described in the **Methodology** section.

1.2.2. Quadrupole ion trap mass spectrometer and electrospray ionization

A quadrupole ion trap (QIT) mass spectrometer is one of the most common benchtop instruments. It became popular in the analytical field due to the combination of the high sensitivity and specificity it provides, a tandem mass capability, and ability of convenient coupling with chromatographic component (LC) with a relatively low cost and its simplicity of operation.[62] Recently, it also became routine instrument for gas phase reaction studies.[63]

Quadrupole ion trap analyzer uses the combination of a constant DC and a radio frequency (RF) oscillating AC electric fields to trap ions. QIT was invented by Paul[64,65] in 1953 (and therefore referred as Paul trap) and modified by Stafford et al. (Finnigan)[66] in 1980s to a commercial mass spectrometer. The 3D trap consists of three hyperbolic electrodes (one ring and two endcap electrodes, **Figure 1.10**).
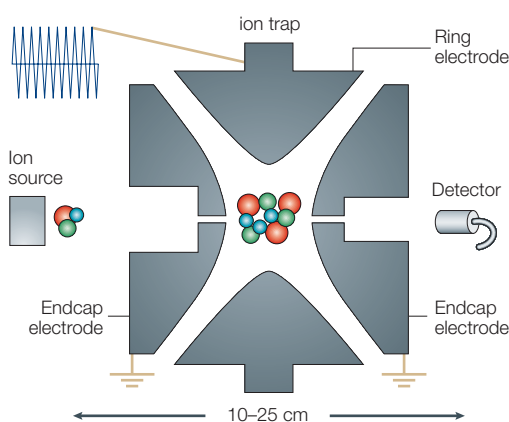
Figure 1.10. Quadrupole ion trap[67]

The ESI-QIT has four main regions: ion generation, ion focusing, ion analysis, and ion detection (**Figure 1.11**). Ions are generated by electrospray (ESI) and focused using two octapole transmission systems (ion optics), and then the ions can be trapped, excited and

ejected in the 3D cavity formed by three electrodes. Each endcap electrode has a small hole through which ions can travel. Damping (buffer) gas (Helium, ~1 mTorr) is filled in the ion trap to dampen the kinetic energy of "hot" ions, and to "focus" ions more, making resolution and sensitivity higher. An AC applied to the ring electrode forces the ions to move toward the center of ion trap. Then by altering the amplitude of the AC, the ions are destabilized and ejected through the hole in the endcap electrode to be detected.[56]

A Finnigan LCQ (DUO or Deca) we used to run Cooks kinetic experiment described herein is QIT mass spectrometer equipped with ESI source.
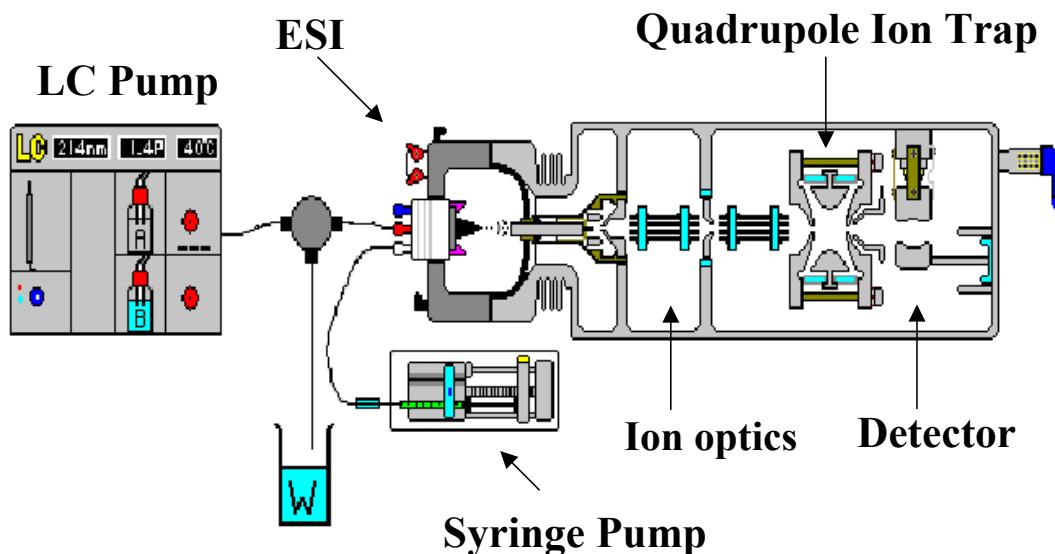
Figure 1.11. Scheme of a quadrupole ion trap mass spectrometer (coupled with LC)[68]

Electrospray ionization (ESI)[69] is an atmospheric pressure ionization (API) source, which ionizes the sample (at atmospheric pressure), and transfers the ions into the mass spectrometer. ESI is "soft" ionization technique, which means the sample typically does not fragment during electrospray. This is one of the reasons why ESI is the most popular ionization method in modern mass spectrometry.[56,70-72]
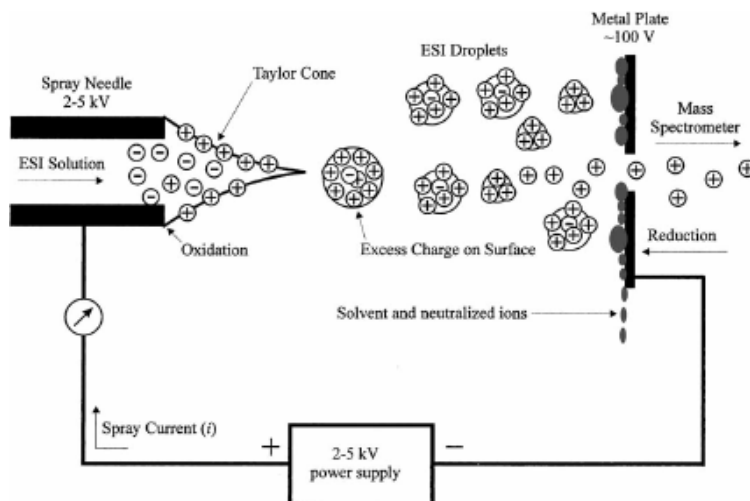
Figure 1.12. Electrospray process

The process of ESI is illustrated in **Figure 1.12.** A sample is dissolved in a polar and volatile solvent (like water, methanol or acetonitrile, etc.; mixture of solvents can be used too), and injected through a narrow capillary (silica or stainless steel) at a relatively slow flow rate (usually 1 μL/min to 1 mL/min). By applying a strong electric field (3-6 kV) at the capillary tip under atmospheric pressure, a charge accumulation at the liquid's surface is induced. Cations (in positive mode) are enriched on the liquid's surface and anions move towards the conductive tip.[73,74] When the effect of the electric field overcomes the surface tension, the highly charged droplet deforms to a conical ("Taylor cone") rather than spherical shape.[75] Then the tip of the cone elongates into a filament, which breaks apart and emits a stream of charged droplets. Increased charged density and coulombic repulsion forces caused by evaporation of solvent lead to the formation of finer droplets. This process of evaporation and repulsion repeats until fully desolvated cations are released. Both single and multiple charged ions (either cations or anions depending on operation mode) can be formed during electrospray process.[70,72]

1.3. Methodology

The methods used in current studies include acidity and proton affinity measurements via bracketing method and Cooks kinetic method, S_N2 reaction studies, and computational studies. Bracketing and supplementary computational studies were applied for nucleobase studies in our lab successfully for more than a decade.[10,13,35-41,76] Cooks kinetic method and S_N2 reaction studies are a newer endeavor. My S_N2 project was the first of its kind in the Lee lab.[77,78]

1.3.1. Bracketing method

Acidity and proton affinity bracketing experiments were conducted using a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FTMS) with a dual cell setup, which has been described previously and in **Instrumentation** section above.[10,13,36,38-41] In our FTMS, two adjoining 2-in. cubic cells are positioned collinearly with the magnetic field produced by a 3.3 T superconducting magnet. Traditionally, the cells are called the source cell (on the "left" side as you face the instrument) and the analyzer cell (on the "right" side; **Figure 1.13**).
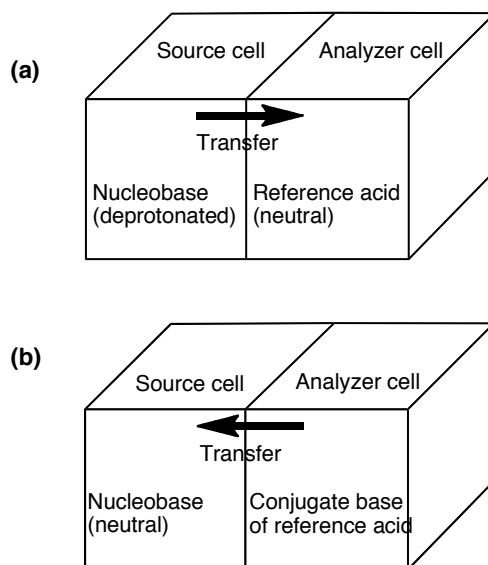
Figure 1.13. Acidity bracketing experiments in FT-ICR MS dual cell

The pressure of the dual cell is pumped down to less than 1×10^{-9} Torr. Solid substrates (normal or damaged nucleobases and analogs) are introduced into the cell via a heatable solids probe. Hydroxide or hydronium ions are generated from water pulsed into the cell, and ionized by an electron beam (typically 8 eV (for OH⁻), or 20 eV (for H_3O^+) and 6 µA, ionization time 0.5 s). Liquid reference acids or bases are introduced via a batch inlet system or a leak valve, and allowed to react with either hydroxide (for acidity measurement) or hydronium ions (for proton affinity (PA) measurement).

The typical protocol for bracketing experiments has been described previously by our lab.[10,13,36,38-41] Briefly, ions are generated from reference acids/bases or nucleobase (or analog), selected, transferred to another adjoining cell via a 2-mm hole in the center of the central trapping plate, cooled by a pulse of argon (that raises the cell pressure to 10^{-5} Torr), and allowed to react with neutral nucleobase or reference/base. Proton transfer (PT) reactions are conducted in both directions (**Figure 1.13**). The occurrence of PT is

regarded as evidence that the reaction is exothermic ("+" in the Tables). Reaction efficiency is used to assess occurrence or non-occurrence of a PT reaction (**Eq. 1.10**); the cutoff is 10%.

$$\text{Efficiency, } \% = k_{exp}/k_{coll} \times 100\% \qquad \textbf{Eq. 1.10}$$

The theoretical ion-molecule collision rate constant k_{coll} is obtained from the "ADO" program, utilizing parametrized trajectory theory.[79,80] "ADO" program estimates k_{coll} based on the dipole moment (μ_D), polarizability (α), and mass of the neutral molecule and the mass-to-charge ratio of the ion (**Eq. 1.11**):[79,80]

$$k_{coll} = (2\pi q/\mu^{1/2})[\alpha^{1/2} + C\,\mu_D\,(2/\pi kT)^{1/2}] \qquad \textbf{Eq. 1.11}$$

In **Eq. 1.11** μ is the reduced mass of ion molecule system, q is charge of ion, C is the dipole locking constant, and T is temperature (usually 298 K). Dipole moments (μ_D) are calculated at B3LYP/6-31+G(d) (or experimentally known); polarizability (α) was estimated using the method of Miller and Savchik.[81]

We run bracketing reactions under pseudo-first order conditions, where the amount of the neutral substrate is in excess relative to the reactant ions. Lets consider an example of acidity bracketing for nucleobase (AH). In order to "bracket" the acidity of AH we perform a series of reactions with reference acids (B$_i$H), and evaluate its efficiencies. For each reference acid we run reactions in both directions (**Figure 1.13**, **Eq. 1.12 (a)** and **Eq. 1.13 (b)**)

$$A^- + BH \rightarrow B^- + AH \qquad \textbf{Eq. 1.12}$$

$$B^- + AH \rightarrow A^- + BH \qquad \textbf{Eq. 1.13}$$

Rate of a reaction (v) of deprotonated nucleobase with neutral reference (Eq. 1.12) can be expressed by **Eq. 1.14**.

$$v = k\,[A^-][BH] \qquad\qquad \textbf{Eq. 1.14}$$

Under our experiment conditions neutral reference acid BH present in excess when we run reaction of A^- with BH (**Figure 1.13 (a)**), $[BH] \gg [A^-]$, therefore $[BH]$ can be used as a constant, $k_{obs} = k[BH]$ and Eq. 1.14 can be re-written as **Eq. 1.15.**

$$v = k_{obs}\,[A^-] \qquad\qquad \textbf{Eq. 1.15}$$

Also, rate can be expressed as **Eq. 1.16.**

$$v = -d[A^-]/dt \qquad\qquad \textbf{Eq. 1.16}$$

Then,

$$\ln[A^-]_t - \ln[A^-]_0 = \ln([A^-]/\ln[A^-]_0) = -k_{obs}t \qquad\qquad \textbf{Eq. 1.17}$$

Therefore, if we plot the $\ln[A^-]$ versus reaction time (t), we should get a straight line with a slope of k_{obs}. Since $k_{obs} = k_{exp}[BH]$, then the actual rate constant k (k_{exp}) could be determined (if $[BH]$ is known). Due to the nature of the gas phase experiments it is more convenient to operate in terms of pressure rather than concentration. So if we plot the natural log of relative intensity of reactant ions vs. time, then **Eq. 1.18** express the relationship between k_{exp} and the *slope* of this line. Φ is the conversion factor = 3.239×10^{16} molecule·cm^{-3}·torr^{-1}.

$$k_{exp} = -\,slope\,/\,(P_{BH} \times \Phi) \qquad\qquad \textbf{Eq. 1.18}$$

The simplest and the most commonly used method to obtain neutral pressure P_{BH} is just to read the pressure from an ion gauge. However, it is often unreliable, both because of the gauge's remote location and the varying sensitivity for different substrates.[37,82] We therefore "back out" the neutral pressure from a control reaction.[38-41,79,83]

Taking acidity measurements as an example, we "back out" the neutral pressure P_{BH} from the control reaction where hydroxide reacts with neutral substrate (BH in our example). Because hydroxide is very basic, we assume this reaction proceeds at the theoretical collision rate $k_{exp}' = k_{coll}'$ (note that $k_{coll}' \neq k_{coll}$ because it refers to different reactions), which can be calculated by "ADO" program. We can then use the calculated k_{coll}' to "back out" neutral pressure P_{BH} **(Eq. 1.19).** The *slope'* here corresponds to the slope of the line of the disappearance of hydroxide ions (natural log of relative intensity) vs. reaction time.

$$P_{BH} = -\ slope' \ / \ (k_{coll}' \times \Phi) \qquad \textbf{Eq. 1.19}$$

1.3.2. Gas phase S_N2 experiments

The gas-phase S_N2 experiments, which are the focus of **Chapter 2**, were conducted using FTMS with a dual cell setup, which has been described previously and in the **Instrumentation** section above.[10,13,36,38-41] The protocol for the S_N2 experiment is similar to that for acidity bracketing, but only in one direction (**Figure 1.13b**). The nucleophile, usually the deprotonated reference acid, was generated in the analyzer cell, and transferred into source cell, where it is allowed to react with neutral substrate (1,3-dimthyluracil, 1,3-dimethylthymine, or methyl chloride). Neutral substrates were introduced via solids probe (if solids) or leak valve (if gas) on the source side; reference

acids – via the batch inlet on the analyzer side. Kinetics of the reactions was carefully monitored, and efficiencies calculated (see bracketing section for details). Efficiencies (%) are reported herein.

1.3.3. Cooks kinetic method

We also used the Cooks kinetic method in a quadrupole ion trap (LCQ) mass spectrometer[84-87] (described in **Instrumentation** section above) to measure the acidities and proton affinities of nucleobases and analogs.

The Cooks kinetic method involves the formation of a proton-bound complex, or dimer, of the unknown AH and a reference acid B_iH of known acidity (**Eq. 1.20**).

$$[AHB_i]^{\ominus} \begin{array}{c} \xrightarrow{k_1} [B_iH] + A^{\ominus} \\ \xrightarrow{k_2} [AH] + B_i^{\ominus} \end{array}$$

Eq. 1.20

$$[AH] + B_i^{\ominus} \underset{}{\overset{K}{\rightleftharpoons}} [B_iH] + A^{\ominus}$$

Eq. 1.21
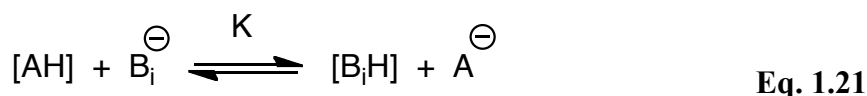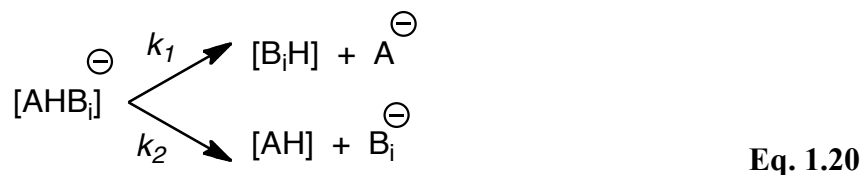
$$K \approx k_1/k_2$$

Eq. 1.22

$$\Delta H_{BiH} - \Delta H_{AH} \approx RT_{eff} \ln K$$

Eq. 1.23

$$\ln(k_1/k_2) = (1/RT_{eff})(\Delta H_{BiH} - \Delta H_{AH})$$

Eq. 1.24

The proton bound dimer $[AHB_i]^-$ is dissociated via collision-induced dissociation (CID). The rate constants k_1 and k_2 are for the two different dissociation pathways. The relationship of these rate constants to ΔH_{acid} is shown in **Eq. 1.22-1.24**. R is the gas constant and T_{eff} is the effective temperature[88] of the activated dimer.[84-87] The ratio of the intensities of the two deprotonated products yields the relative acidity of the two compounds of interest (**Eq. 1.24**), assuming the dissociation has no reverse activation

energy barrier and that the dissociation transition structure is late and therefore indicative of the stability of the two deprotonated products. These assumptions are generally true for proton bound systems.[43,87,89] In addition, entropy effects should be negligible, which means the two acids (our unknown and reference acid) should have similar structures.

To obtain the acidity of compound AH, the natural logarithm of the relative intensity ratios is plotted versus the acidities for a series of reference acids, where the slope is $(1/RT_{eff})$ and the y-intercept is $(-\Delta H_{AH}/RT_{eff})$. The T_{eff} is obtained from the slope. The acidity of compound AH, (ΔH_{AH}) is calculated from either eq. 2 or the y-intercept.

The same procedure can be applied for proton affinity measurements (via positively charged proton bound dimers).

The proton-bound complex ions are generated by electrospray (ESI) of 100–500 μM solutions of an unknown and a reference acid (or base, for PA measurement). Methanol or water–methanol (20%) solution are used as a solvent.[72] Addition of one drop of acetic acid or ammonium hydroxide sometimes used to promote dimer formation. An electrospray needle voltage of ~4 kV and the flow rate of 25 μL/min is applied. The proton-bound complex ions are isolated and then dissociated by applying collision-induced dissociation (CID); the complexes are activated for about 30 ms. Finally, the dissociation product ions are detected to give the ratio of the deprotonated (or protonated) analyte and deprotonated (or protonated) reference acid. A total of 40 scans are averaged for the product ions.

1.3.4. Computational method

Theoretical studies were conducted in order to predict acidities, proton affinities and tautomeric composition of nucleobases and analogs. In **Chapter 2** and **Chapter 3** transition states for the gas phase S_N2 reactions, and potential energy surfaces (PES) were also calculated.

The gas phase calculations were conducted using Gaussian03[90] and Gaussian09[91] programs. The method of choice is B3LYP/6-31+G(d),[92-94] which has been previously shown to be reasonably accurate for the gas phase acidity and proton affinity calculations of nucleobases.[10,13,35-37,40,41,76] The geometries are fully optimized and the frequencies are calculated (unless indicated otherwise). No scaling factor is applied. All the gas phase values reported are ΔH at 298 K.

Other more sophisticated (and therefore expensive and time consuming) methods like M06-2X/6-311+G(2df,2p),[95,96] MP2/6-31+G(d,p),[97-102] and CBS-QB3[103,104] were used as needed (see **Chapters 2, 3, 4 and 5** for details).

Dielectric medium calculations were done using the conductor-like polarizable continuum solvent model (CPCM, full optimization; UAKS cavity) at B3LYP/6-31+G(d) as implemented in Gaussian03.[105,106,107] The "total free energy in solution" values are reported. For more details see **Chapters 4** and **5**.

Chapter 2. The Gas Phase S$_N$2 Reaction and UDG Mechanism Implications for Electron Delocalization in Uracil and Thymine Leaving Groups

2.1. Introduction

Uracil and thymine are pyrimidine nucleobases that differ in structure by only a methyl group at C5 (**1a** and **3a**). Uracil naturally occurs in RNA, while thymine is its DNA counterpart.

Although uracil and thymine are very similar in structure, the presence of uracil in DNA is problematic.[10,108-110] Uracil can arise in DNA from cytosine deamination, which is mutagenic; uracil can also be misincorporated into DNA, leading to cytotoxic uracil•adenine base pairs.[8,108-111] Uracil is removed from the genome by the enzyme uracil DNA glycosylase (UDG).[5,8,9,15,17,112-114]

The mechanism of UDG has been shown to involve N1-deprotonated uracil as the leaving group (LG).[8,11-14,16-18,115-125] Deprotonated uracil as a leaving group seems somewhat surprising, and begs the question: *How good of a leaving group is deprotonated uracil*? The N1-H pK_a in water is 9.8, which would indicate a fair or mediocre leaving group ability. In the enzyme, uracil has a depressed pK_a of 6.4.[116,117]

Examining properties in the gas phase is useful for elucidating inherent reactivity in the absence of solvent.[10,13,35,36,39-41,43,46,76,126,127] In previous studies we calculated and measured the gas phase acidity of uracil and found it to be as acidic as hydrochloric acid, indicating that in the gas phase, deprotonated uracil might be, relatively speaking, a good LG.[10] Furthermore, because enzyme environments are sometimes quite nonpolar, reactivity in the gas phase -- a kind of "ultimate" nonpolar medium -- can yield insight into biological reactivity.[10,13,31,32,35-37,39-41,76] Uracil is particularly intriguing, we found, because the acidities at the N1 and N3 sites are very different in the gas phase, but coalesce in aqueous solution.[10,13]

a $R_1 = R_2 = H$
b $R_1 = H, R_2 = CH_3$
c $R_1 = CH_3, R_2 = H$
d $R_1 = R_2 = CH_3$
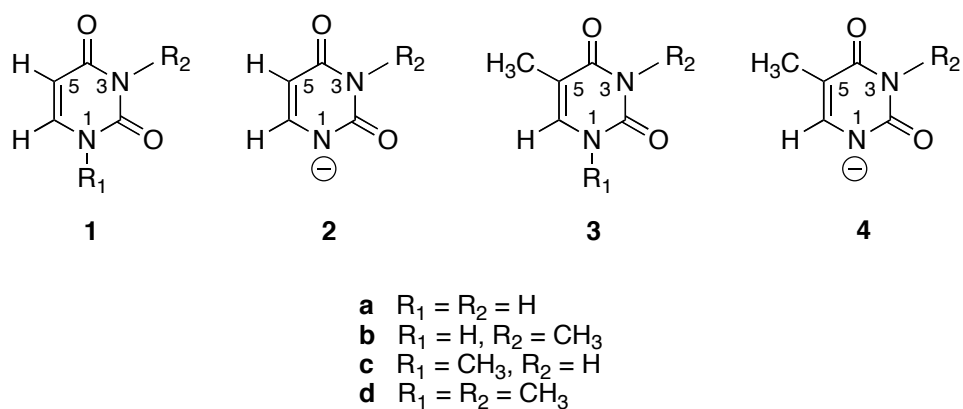
Figure 2.1. Di- and mono-substituted uracil and thymine studied herein

Having established that uracil (**1a**) is quite acidic in the gas phase, we now examine the leaving group ability of the conjugate base, N1-deprotonated uracil (**2a**) in a substitution reaction. Since uracil is as acidic as HCl, is deprotonated uracil as good of a leaving group as chloride in the gas phase?

We are also interested in comparing the leaving group abilities of deprotonated uracil versus deprotonated thymine. The occurrence of uracil in the human genome is typically one uracil per $> 10^7$ normal DNA base pairs. The ability of the UDG enzyme to find and

excise the few uracils present while leaving the structurally similar thymine untouched is of interest.[128] A recent study indicates that the uracil versus thymine discrimination could be due in part to base-pair dynamics.[110] We wanted to probe the possibility that another contribution to the favorable excision of uracil over thymine could be due to the relative leaving group abilities of the corresponding conjugate bases (deprotonated at N1). If N1-deprotonated thymine (**4a**) is a poorer leaving group than N1-deprotonated uracil (**2a**), this could presumably contribute to the favorable excision of uracil over thymine.[10,13,39,76] Leaving group ability has been implicated in the discrimination of substrates with other glycosylases.[36,37,40,41,129]

2.2 Experimental section

1,3-dMU (**1d**) and 3-MeT (**3b**) are commercially available, and were used as received. 1,3-dMT (**3d**) was synthesized from thymine (**3a**). The procedure used is similar to that reported in the literature for 1,3-dMU synthesis and described below (**Figure 2.2**).[130,131] The product was purified and the identity confirmed by ^{1}H and ^{13}C NMR.

All nucleophiles except CH_3S^- were generated from commercially available neutral reference acids by deprotonation with hydroxide ions. CH_3S^- was generated via the elimination reaction of hydroxide plus dimethyldisulfide, which is a well-known source of the methyl thiolate anion.[132-134]

2.2.1. 1,3-Dimethylthymine (1,3-dMT, 3b) synthesis

Potassium carbonate (0.553 g, 4 mmol) was added to 15 ml of dimethyl acetamide (DMA), and gently heated to dissolve the potassium carbonate. The mixture was cooled to room temperature, and 0.126 g (1 mmol) of thymine was added followed by 0.190 ml

(3 mmol) of methyl iodide. The reaction mixture was stirred for 48 hours at room temperature. Next, 10 ml of distilled water was added. The aqueous solution was extracted with chloroform (3 x 40 ml). The organic phase was washed with distilled water (6 x 40 ml), and then dried with Na_2SO_4. Solvent was removed by rotovap, and the solid product was dried on high vacuum overnight. 0.114 g (74%) of pale yellow crystals were formed. ^1H NMR (400 MHz, CDCl_3, ppm) : 6.98 (1H, s, -CH), 3.37, 3.38 (6H, unresolved s, -NMe), 1.94 (3H, s, Me); ^{13}C NMR (500 MHz, CDCl_3, ppm): 164.3 (C4=O); 152.1 (C2=O); 139.2 (C6); 109.7 (C5); 36.9 (N1-Me); 28.1 (N3-Me); 13.2 (C5-Me).
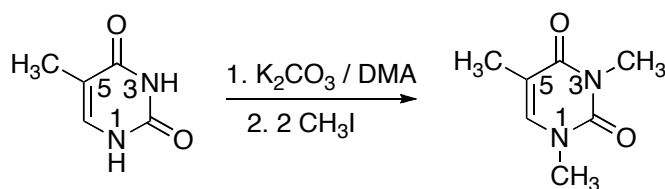
Figure 2.2. 1,3-dMT synthesis

2.2.2. Gas-phase S_N2 and acidity measurement experiments

All S_N2 reaction experiments and 3-MeT acidity bracketing were conducted using a dual cell Fourier transform ion cyclotron resonance mass spectrometer (FTMS), according to the protocols, which has been described previously, and in **Chapter 1** in details.[10,13,36,39-41] Proton transfer reactions were conducted in both directions (deprotonated 3-MeT with neutral reference acids and the conjugate bases of neutral reference acids plus 3-MeT). The occurrence or non-occurrence of proton transfer is denoted as a "+" or "−" in **Table 2.1.** Each kinetics experiment was run at least three times; reported values are the average and standard deviation.

The Cooks kinetics method[39,40,84-87,135] was used to conduct a relative acidity study of

3-MeT (**3b**) and 3-MeU (**1b**). We use a quadrupole ion trap (LCQ) mass spectrometer to conduct the experiments. The detailed protocol for Cooks kinetic experiments conducted in our lab has been described previously, and in **Chapter 1**.[40,41] Briefly, the "relative" Cooks kinetic method we use herein involves formation of a proton-bound dimer of the two species of interest (3-MeU and 3-MeT herein). The dimer is isolated and dissociated via CID. The ratio of intensities of the two deprotonated substrates yields the ratio of rate constants of the two possible dissociation pathways, which yields the relative acidities of the two substrates.

Proton-bound dimers were generated by electrospray (ESI) from the 1:1 mixture of 250 μM 3-MeT (**3b**) and 250 μM 3-MeU (**1b**) solutions in 20% methanol-water. A needle voltage of 4 kV, capillary temperature 150 $^{\circ}$C, and a flow rate of about 25 mL/min were used. The proton bound complex [(3-MeU)$^{-}$•H^{+}•(3-MeT)$^{-}$] was isolated and activated for about 30 ms. 40 scans were averaged for the product ions, and the experiment was repeated 3 times. A T$_{eff}$ of 420 K, obtained from a calibration experiment with 3-MeT (**3b**), was used.

2.2.3. Calculations

Calculations were conducted at B3LYP/6-31+G(d) using Gaussian03;[90,92-94] the geometries were fully optimized and frequencies were calculated. This method has been shown to be reasonable for calculating S$_{N}$2 reaction potential energy surfaces.[136-139] All the values reported are at 298 K. No scaling factor was applied. All calculated TS structures have one negative frequency. Partial charges were calculated using CHELPG as implemented in Gaussian03.[140] As described in the paper, for some substrates we used the CBS-QB3 model chemistry.[103,104]

2.3. Results

S_N2 reactions in the gas phase have been studied for more than two decades. They follow the "classic" gas phase double-well potential energy surface, where an initial ion-molecule complex is formed that can either dissociate back to reactants or react to products.[46,48-55,141-147] Previously measured second order reaction rate constants and efficiencies for S_N2 reactions of methyl chloride with different anionic nucleophiles are shown in **Table 1.2** (see **Chapter 1**; reaction efficiency is defined as the ratio of the observed rate constant to the estimated collision rate constant calculated by parametrized trajectory calculations).[34,79,80]

2.3.1. Model systems for gas phase study

The excision of uracil from DNA involves nucleophilic attack of the C1' of ribose (**Figure 2.3**). Kinetic isotope effects point to a "dissociative S_N2" reaction mechanism (D_N*A_N).[8,14,121] Our interest is in testing the leaving group ability of deprotonated uracil in a substitution reaction. The simplest model would be to examine reactivity at the N1-CH_3 group of 1-methyluracil (**1c**). However, experiments with 1-methyluracil are limited by the acidity of the N3-H, which has been measured to be 348 ± 3 kcal mol^{-1} in the gas phase.[10,13] Therefore, any nucleophile with a proton affinity (PA) of 348 or greater will likely deprotonate the N3-H. Proton transfers are enthalpically generally barrierless ($\Delta H^{\ddagger}$ = 0) and will, if exothermic, compete with the S_N2 reaction.[43] Using nucleophiles with PAs less than 348 kcal mol^{-1} would be too limiting; for substitution reactions with CH_3Cl, anions with a PA at or below 348 kcal mol^{-1} yield very low efficiencies (**Table 1.2**). Therefore, in order to allow the use of more basic nucleophiles, we chose to examine the

1,3-dimethyl substrates (**1d**, **3d**) wherein the methyl group at N3 acts as a sort of "protecting group".
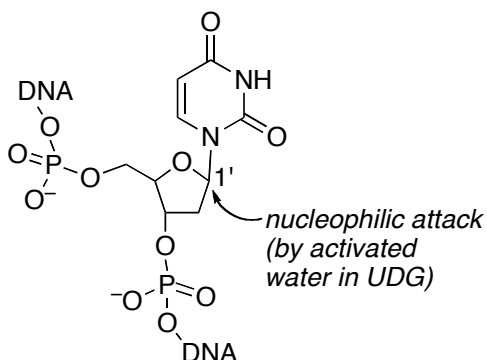
Figure 2.3. Nucleophilic attack at C1' to excise uracil.[8]

2.3.2. Acidity studies

With the chosen model systems, our leaving groups are no longer the deprotonated uracil and deprotonated thymine, but rather the 3-methyl derivatives (**Figure 2.4**).
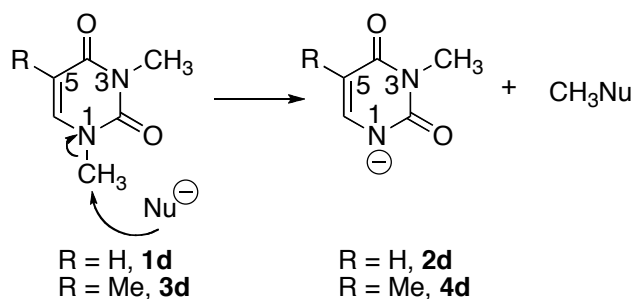
Figure 2.4. S_N2 reactions studied.

As a starting point toward ascertaining whether N1-deprotonated-3-methyluracil is a better leaving group than N1-deprotonated 3-methylthymine, we assessed the acidity of the N1-H proton in 3-methyluracil (3-MeU, **1b**) and compared it to the acidity of the N1-H proton in 3-methylthymine (3-MeT, **3b**). The acidity of 3-MeU has been calculated previously by us (332.8 kcal mol^{-1}, **Figure 2.5**); herein we calculate the ΔH_{acid} of 3-MeT

to be 1.4 kcal mol^{-1} less than the ΔH_{acid} of 3-MeU (332.8 vs. 334.2 kcal mol^{-1}, Figure 2.5).[13] Based on these acidities, we would expect deprotonated 3-methyluracil to be a better leaving group than deprotonated 3-methylthymine.
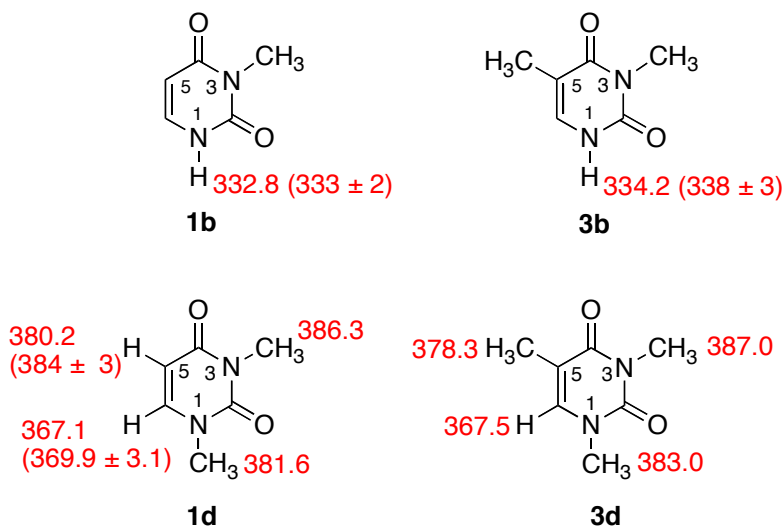
Figure 2.5. Calculated N1-H acidities of 3-methyluracil[13] (3-MeU, **1b**) and 3-methylthymine (3-MeT, **3b**) and all the sites of 1,3-dimethyluracil (**1d**) and 1,3-dimethylthymine (**3d**) (B3LYP/6-31+G(d), ΔH_{298K}, kcal mol^{-1}). Experimental values (if known) are in parentheses.[13,89]

We have also previously measured the gas phase ΔH_{acid} of 3-MeU (**1b**) to be 333 ± 2 kcal mol^{-1}.[13] We bracket the acidity of 3-MeT herein (**3b**) (**Table 2.1**). We find that while the conjugate base of 3-MeT deprotonates 2-chloropropanoic acid (ΔH_{acid} = 337.0 ± 2.1 kcal mol^{-1}) and acids with lower ΔH_{acid} values, it cannot deprotonate trifluoro-*m*-cresol (ΔH_{acid} = 339.3 ± 2.1 kcal mol^{-1}) or reference acids with higher ΔH_{acid} values. Consistent with this, 2-chloropropanoate cannot deprotonate 3-MeT, but trifluoro-*m*-cresolate can. We therefore bracket the ΔH_{acid} of 3-MeT to be 338 ± 3 kcal mol^{-1} (ΔG_{acid} (3-MeT) = 331 ± 3 kcal mol^{-1}).

Table 2.1. Summary of results of acidity bracketing of 3-methylthymine (3-MeT, **3b**)

Reference compound	$\Delta H_{acid}{}^a$	$\Delta G_{acid}{}^a$	Proton transfer[b]	
	(kcal mol^{-1})	(kcal mol^{-1})	Ref. acid	Conj. base
2,4-pentanedione	343.8 ± 2.1	336.7 ± 2.0	−	+
methyl cyanoacetate	340.80 ± 0.60	334.5	−	+
trifluoro-*m*-cresol	339.3 ± 2.1	332.4 ± 2.0	−	+
2-chloropropanoic acid	337.0 ± 2.1	330.4 ± 2.0	+	−
malononitrile	335.8 ± 0.1	328.1 ± 2.0	+	−
pyruvic acid	333.5 ± 2.9	326.5 ± 2.8	+	−
difluoroacetic acid	331.0 ± 2.2	323.8 ± 2.0	+	−
1,1,1-trifluoro-2,4-pentadione	328.3 ± 2.9	322.0 ± 2.0	+	−

[a]Acidities are in kcal mol^{-1}.[33] [b]A "+" indicates the occurrence and a "−" indicates the absence of proton transfer.

We also conducted Cooks kinetic method experiments to measure the relative acidity of 3-MeT (**3b**) and 3-MeU (**1b**).[84-87,135] We accomplished this by dissociating the [(3-MeU)$^-$•H$^+$•(3-MeT)$^-$] dimer. These experiments indicate that 3-MeU (**1b**) is 2-3 kcal mol^{-1} more acidic than 3-MeT (**3b**).

The experiments therefore indicate a difference in acidity between 3-MeU and 3-MeT that is on the order of 2-5 kcal mol^{-1}. Although this is a rather large range, what is consistent is that 3-MeU is more acidic than 3-MeT. This could make the N1$^-$ conjugate base of 3-MeU a potentially better leaving group than that of 3-MeT.

Another important reason to probe acidities is to establish the upper limit of proton affinity for the nucleophiles to be studied experimentally. As we noted earlier, we are using the N3-methyl substrates since the N3-H of 1-methyluracil has a ΔH_{acid} of 348 kcal mol^{-1}, and to avoid competition between S$_N$2 reaction at the N1-CH$_3$ and proton transfer

at the N3-H, we would be limited to using nucleophiles with proton affinities (PA) under 348 kcal mol^{-1}.[10,13] Now that we have established that we will be using the 1,3-dimethyl derivatives **1d** and **3d**, we need to assess the acidities of all the sites of those substrates (**Figure 2.5**). The C6 protons of both these derivatives are quite acidic, with values just below that of acetone (the C6 and C5 of 1,3-dMU have also been previously measured by the Gronert lab and our lab).[13,145] Therefore, to avoid competition from deprotonation, nucleophiles with proton affinities below 367 kcal mol^{-1} will be utilized.

2.3.3. S$_N$2 reaction studies -- calculations.

2.3.3.1. 1,3-Dimethyluracil (1,3-dMU)

We next calculated the energetics associated with S$_N$2 reactions of 1,3-dMU. We chose formate and methyl thiolate as the nucleophiles. Formate was chosen as a probable slow reaction example, based on the acidity of formic acid (ΔH_{acid} = 346.2 ± 1.2 kcal mol^{-1}) and the known methyl chloride data for nucleophiles in that acidity range (**Table 1.2**). Methyl thiolate (ΔH_{acid} = 357.6 ± 2.0 kcal mol^{-1}) was chosen as a faster (though still moderate) reaction example.

The reaction of formate with 1,3-dMU first forms the expected reactant ion-molecule complex, a process that is 23.1 kcal mol^{-1} exothermic (Figure 2.6). The $\Delta H^{\ddagger}$ barrier to the S$_N$2 reaction is 32.3 kcal mol^{-1} from this complex (and 9.2 kcal mol^{-1} from the separated reactants **(Figure 2.6)**). The product ion-molecule complex is 9.3 kcal mol^{-1} more stable than the separated reactants, but the separated products are 2.7 kcal mol^{-1} higher in energy than the separated reactants. Given that both the transition state (TS) and the

separated products are higher in energy than the separated reactants, this reaction is not likely to proceed significantly under our gas-phase conditions.
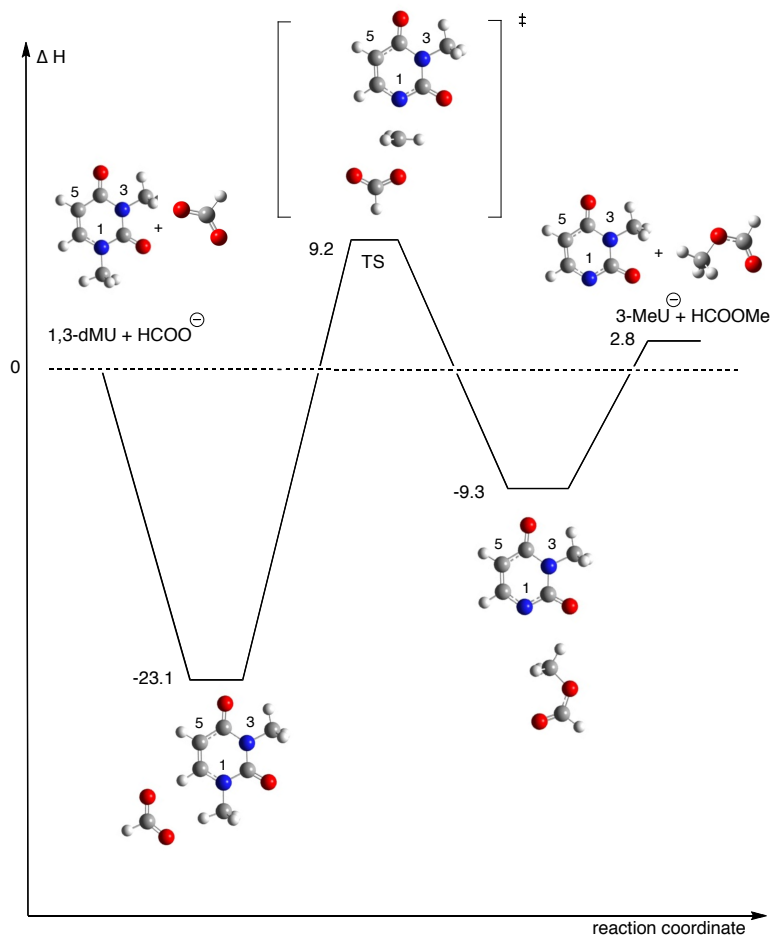
Figure 2.6. Calculated (B3LYP/6-31+G(d)) potential energy diagram of the reaction of 1,3-dMU (**1d**) with formate (ΔH, 298 K)

For the reaction of methyl thiolate with 1,3-dMU (**Figure 2.7**), the transition state is calculated to be just 4.3 kcal mol^{-1} higher than the energy of the separated reactants. The energy of the transition structure must be below the energy of the separated reactants for reaction to be observed, so we could see reaction depending on how accurate the calculations are, and how entropically unfavorable the process is.[46]

43
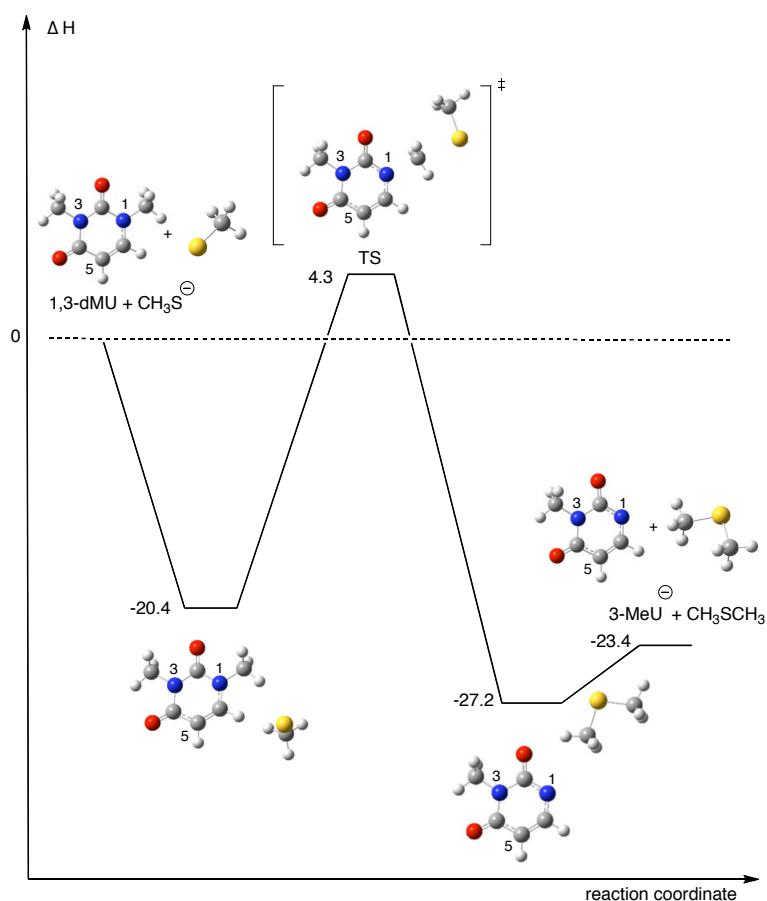
Figure 2.7. Calculated (B3LYP/6-31+G(d)) potential energy diagram of the reaction of 1,3-dMU (**1d**) with methyl thiolate (ΔH, 298 K)

2.3.3.2. Methyl chloride

We also calculated the energetics for the reaction of formate (**Figure 2.8**) and methyl thiolate (**Figure 2.9**) with methyl chloride, to provide a benchmark for comparison. Interestingly, the methyl chloride reactions are consistently more exothermic and have lower barriers than the uracil reactions. With formate, the $\Delta H^{\ddagger}$ is 1.9 kcal mol^{-1} *below* the separated reactants (**Figure 2.8**). For methyl thiolate, the $\Delta H^{\ddagger}$ is 7.5 kcal mol^{-1} below the separated reactants (**Figure 2.9**).
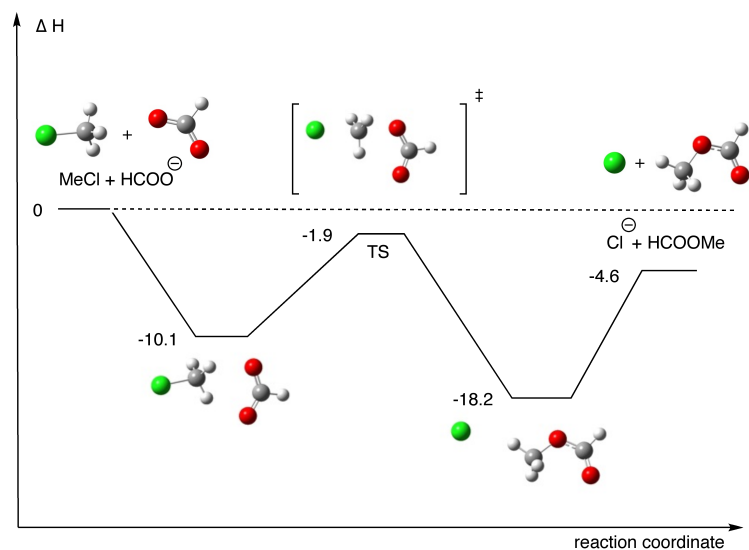
Figure 2.8. Calculated (B3LYP/6-31+G(d)) potential energy diagram of the reaction of

methyl chloride with formate (ΔH, 298 K)
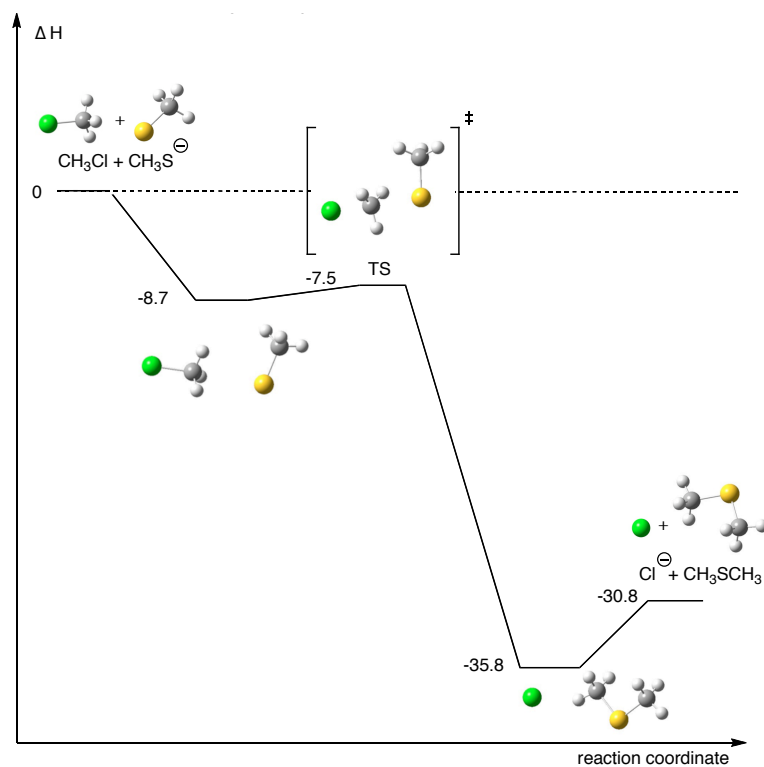
Figure 2.9. Calculated (B3LYP/6-31+G(d)) potential energy diagram of the reaction of

methyl chloride with methyl thiolate (ΔH, 298 K)

Our calculations therefore indicate that although 3-MeU and HCl have similar acidities, the 1,3-dMU S_N2 reactions are expected to be slower than those of CH_3Cl.

2.3.3.3. 1,3-Dimethylthymine (1,3-dMT)

Our calculations indicate that in keeping with the fact that thymine is less acidic than uracil, the S_N2 reactions wherein deprotonated thymine is a leaving group do have slightly higher barriers than the corresponding reactions with deprotonated uracil. Using 1,3-dMT as the model system, the energy surfaces for reaction with formate and methyl thiolate were calculated (**Figures 1.10** and **1.11**).
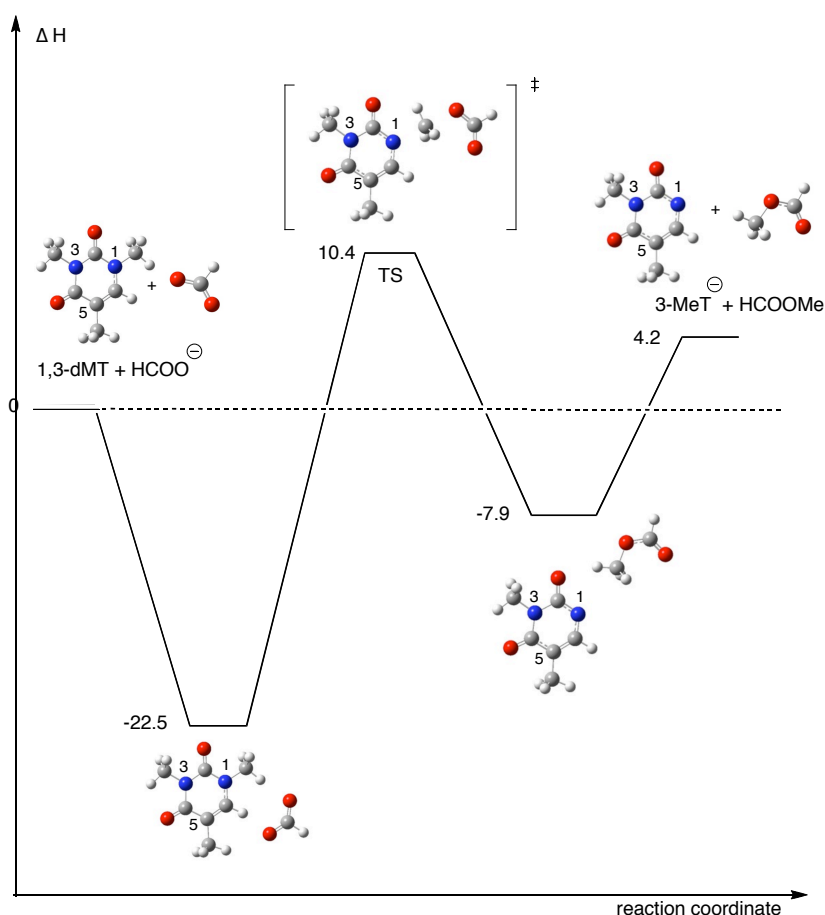
Figure 2.10. Calculated (B3LYP/6-31+G(d)) potential energy diagram of the reaction of 1,3-dMT (**3d**) with formate (ΔH, 298 K)
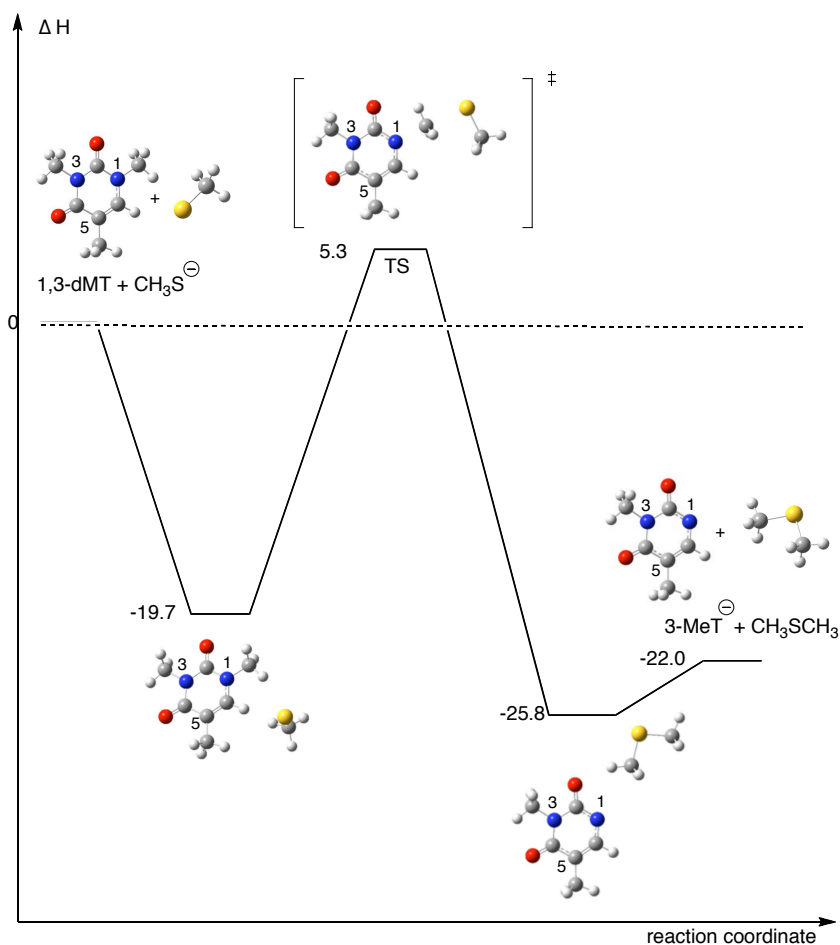
Figure 2.11. Calculated (B3LYP/6-31+G(d)) potential energy diagram of the reaction of 1,3-dMT (**3d**) with methyl thiolate (ΔH, 298 K)

For the formate reaction, the barrier is 1.2 kcal mol^{-1} higher with 1,3-dMT than with 1,3-dMU (relative to separated reactants) and is also more endothermic. For the methyl thiolate reaction, the barrier for 1,3-dMT is 1 kcal mol^{-1} higher than that with 1,3-dMU. We also find computationally that for a wide range of nucleophiles (formate, acetate, n-pentylthiolate, methyl thiolate, and anilide), the exothermicities of the 1,3-dMU versus the 1,3-dMT reactions are always roughly 1.4 kcal mol^{-1} apart (with 1,3-dMU being more exothermic; **Table 2.2**).

Table 2.2. Calculated enthalpies of S_N2 reactions of 1,3-dMU and 1,3-dMT with series of nucleophiles

Nucleophile	PA, kcal mol^{-1}	ΔH_{rxn}, kcal mol^{-1} Substrate	
		1,3-dMU	1,3-dMT
HCOO$^-$	346.2	2.7	4.2
AcO$^-$	348.1	-6.3	-4.9
n-C$_5$H$_{11}$S$^-$	352.5	-20.3	-18.9
n-PrS$^-$	354.2	-20.5	-19.1
CH$_3$S$^-$	357.6	-23.4	-22.0
PhNH$^-$	366.4	-30.5	-29.1

2.3.3.4. Methyl chloride versus pyrimidine derivatives: acidity calculations revisited

Experimentally, HCl and uracil have similar acidities. However, our calculations indicate that although the N1-deprotonated uracil derivatives are still fairly good leaving groups (for example, the reaction of methyl thiolate, which is not that basic, with 1,3-dMU is exothermic by 23.4 kcal mol^{-1} and has a barrier of only 4.3 kcal mol^{-1} above the separated reactants), chloride is still better.

To assess whether this calculated difference in reactivity might be a computational artifact, we compared the computed acidities of 3-MeU and HCl. We know by experiment that the two have comparable acidities.[10,13] However, we find that the ΔH_{acid} of HCl calculates to 325.1 kcal mol^{-1} at B3LYP/6-31+G(d), 7.7 kcal mol^{-1} more acidic than the calculated value for 3-methyluracil, and 8.3 kcal mol^{-1} more acidic than the known measured ΔH_{acid} of HCl (333.4 ± 0.1 kcal mol^{-1}). Essentially, although 3-methyluracil and HCl have the same experimental acidity, the B3LYP/6-31+G(d) calculations are not accurate: HCl appears to be more acidic. To assess whether the

calculated differences in S_N2 reaction barriers for the methyl chloride and pyrimidine derivatives is due to a related computational artifact or truly reflects the reactivity, we conducted experiments, which are described later in the paper.

2.3.3.5. N1 vs. N3 attack

One possible complication for examining these reactions experimentally is a potential competition between N1-CH$_3$ attack (which we wish to see) and N3-CH$_3$ attack (which is not of biological interest; **Figure 2.12**).
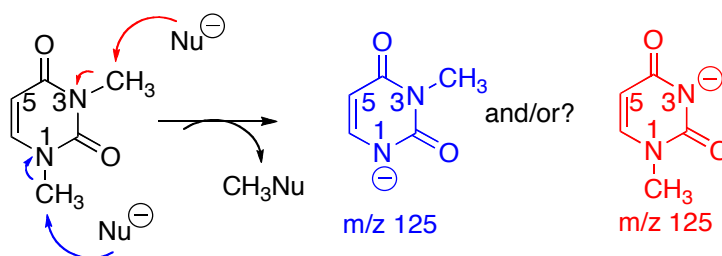
Figure 2.12. N1 versus N3 attack of 1,3-dimethyluracil.

We computationally examined nucleophilic attack at the N1-CH$_3$ versus N3-CH$_3$ for a series of nucleophiles (formate, acetate, n-pentyl thiolate, n-propyl thiolate, methyl thiolate and anilide) and find that the N1-CH$_3$ S_N2 reaction is consistently on the order of 12 kcal mol^{-1} more exothermic than the N3-CH$_3$ path (**Table 2.3**).

Table 2.3. Calculated enthalpies of S_N2 reactions of 1,3-dMU (**1d**) with a series of nucleophiles: N1-methyl vs. N3-methyl group attack.

Nucleophile	PA[33], kcal mol^{-1}	ΔH_{rxn}, kcal mol^{-1}	
		N1 attack	N3 attack
HCOO$^-$	346.2	2.7	15.1
AcO$^-$	348.1	-6.3	6.0
n-C$_5$H$_{11}$S$^-$	352.5	-20.3	-7.9
n-PrS$^-$	354.2	-20.5	-8.2
CH$_3$S$^-$	357.6	-23.4	-11.1

PhNH⁻	366.4	-30.5	-18.1

We also calculated the activation enthalpies ($\Delta H^{\ddagger}$) for the reactions with formate and methyl thiolate; for both reactions, the barrier for attack at the N1-methyl group is 10 kcal mol^{-1} lower than that at the N3-methyl. These differences in barrier are significant enough that experimentally, S_N2 reaction at N1-CH$_3$ should be considerably favored over S_N2 reaction at N3-CH$_3$. These results are as expected, given that the N1-H is more than 10 kcal mol^{-1} acidic than the N3-H.[10,13]

2.3.4. S_N2 reaction studies -- experiments

The calculated barriers for the reaction of formate with 1,3-dMU versus methyl chloride indicates that the methyl chloride reaction should be faster (**Figures 2.6 and 2.9**). Furthermore, **Table 1.2** indicates that a nucleophile with PA of 361.7 (CF$_3$CH$_2$O⁻) reacts with methyl chloride with an efficiency of just 11%. We would therefore expect that the S_N2 reactions of nucleophiles whose PAs are less than the ΔH_{acid} of C6-H ($\Delta H_{acid} \sim$ 370 kcal mol^{-1}) with 1,3-dMU and 1,3-dMT will all be relatively slow.

To establish that we can see S_N2 reactivity under our conditions, we repeated the known reactions of 2,2,2-trifluoroethoxide and methyl thiolate with methyl chloride. Our efficiency values for these S_N2 reactions are comparable to those obtained previously (for 2,2,2-trifluoroethoxide: 13.3% (current work) vs. 11% (literature); for methyl thiolate 8.0% (current work) vs. 4.7% (literature)).[46,50,52]

The reactions of a series of nucleophiles with 1,3-dMU and 1,3-dMT were studied (**Table 2.4**). The S_N2 reaction product was observed for the reaction of 1,3-dMU (**1d**) with nucleophiles ranging from *m*-CF$_3$PhO⁻ (PA=339 kcal mol^{-1}) to HO⁻ (PA=390 kcal

mol^{-1}). However, when using nucleophiles with PA greater than about 365 kcal mol^{-1}, both proton transfer and S_N2 reactions were observed. Proton transfer is presumably the result of deprotonation of the most acidic (C6-H) site of 1,3-dMU (calculated ΔH_{acid} = 367.1 kcal mol^{-1}; experimental ΔH_{acid} = 369.9 ± 3.1 kcal mol^{-1}).[13,89] The S_N2 reaction efficiencies of the reactions of nucleophiles with PA less than 365 kcal mol^{-1} with 1,3-dMU are all fairly low (less than 1%) (**Table 2.4**). For nucleophiles with PA lower than 339 kcal mol^{-1} the efficiencies were less than 0.01%.

Table 2.4. S_N2 Reactions of 1,3-dMU (**1d**) and 1,3-dMT (**3d**)

PA[33] (nucleophile), kcal mol^{-1}	Nucleophile, A$^-$	Substrate (efficiency of S_N2 reaction, %)	
		1,3-dMU (**1d**)	1,3-dMT (**3d**)
364.1	PhNEt$^-$	0.38±0.07	0.006±0.003
361.7	$CF_3CH_2O^-$	0.84±0.38	0.004±0.001
359.5	$C_4H_4N^-$	0.35±0.21	0.033±0.019
357.6	CH_3S^-	0.23±0.14	0.009±0.003
354.2	n-PrS$^-$	0.22±0.07	0.014±0.009
353.4	i-PrS$^-$	0.23±0.21	0.004±0.001
352.5	n-$C_5H_{11}S^-$	0.28±0.04	0.008±0.006
348.1	AcO$^-$	0.11±0.09	no S_N2 reaction
346.2	HCOO$^-$	0.08±0.04	no S_N2 reaction
339.3	m-CF_3-PhO$^-$	0.04±0.02	no S_N2 reaction

We also examined the reactions of the same series of nucleophiles with 1,3-dMT (**3d**). For nucleophiles with PA higher than 365 kcal mol^{-1}, both S_N2 reactions and proton transfer were observed (as we saw with 1,3-dMU). S_N2 reaction products were observed for nucleophiles with PAs as low as ~352 kcal mol^{-1}. Generally, the S_N2 reaction efficiencies for 1,3-dMT (**3d**) are lower than those efficiencies for 1,3-dMU (**1d**). Such

small efficiencies are quite challenging to measure, and therefore the precision for the 1,3-dMT (**3d**) measurements varies from 0.001 to 0.02%. For the reaction of nucleophiles with PA smaller than 352.5 kcal mol^{-1} with 1,3-dMT (**3d**), no S$_N$2 reaction was observed (**Table 2.4**).

It therefore appears that the S$_N$2 reaction proceeds for both 1,3-dMU (**1d**) and 1,3-dMT (**3d**), and that the efficiencies observed for 1,3-dMT (**3d**) are lower than those for 1,3-dMU (**1d**).

2.4. Discussion

2.4.1. Acidity

The experimental measurements of acidity indicate that while 3-MeU and HCl have comparable acidities (around 333 kcal mol^{-1}), 3-MeT is slightly less acidic, by 2-5 kcal mol^{-1}. The B3LYP/6-31+G(d) calculations involving HCl are not quite in agreement with experiment. The computed (B3LYP/6-31+G(d)) acidity for HCl is 325.1 while that for 3-MeU is 332.8 kcal mol^{-1}. Thus, by calculation, the acidity difference is more than 7 kcal mol^{-1} while by experiment that difference is much less. The calculations are likely to be in error, since the measured acidity of HCl is very well known to be ΔH_{acid} = 333.4 ± 0.1 kcal mol^{-1}, and our previous bracketing studies show that reaction of HCl and 3-MeU (i.e. the conjugate base of one with the acid of the other and vice versa) proceeds in both directions.[13] To establish that the discrepancy is due to a computational artifact, we calculated the acidity of HCl and 3-MeU using the CBS-QB3 (complete basis set) model chemistry, which has been shown to accurately calculate thermochemical values.[103,104,139,148,149] Using this method, we find that the calculations and experimental

data are in much better agreement: for HCl: ΔH_{acid} = 332.2 kcal mol^{-1} (calc) and 333.4 ±

0.1 kcal mol^{-1} (expt); for 3-MeU: ΔH_{acid} = 335.5 kcal mol^{-1} (calc) and 333 ± 2 kcal mol^{-}

^{1} (expt). Therefore, it does appear that the discrepancy between calculations and

experiments is a computational issue. We therefore rely on our experimental data; the

experimental acidity studies indicate that Cl^{-} and the conjugate base of 3-MeU might be

comparable leaving groups, based on their acidities; 3-MeT would be a slightly worse

leaving group.

2.4.2. S_N2 reactions. Methyl chloride vs. pyrimidines

The calculated energy diagrams for the reactions of formate and methyl thiolate with

CH$_3$Cl and 1,3-dMU are superimposed in **Figures 2.13** and **2.14**. We are not indicating

that the methyl chloride and dimethyluracil systems start with the exact same total energy;

we plot them in such a way that the differences in reactivity are easier to see. We leave

off 1,3-dMT to keep the diagrams uncluttered. In both the formate and methyl thiolate

reactions, the methyl chloride energetics are more favorable than the 1,3-dMU energetics

(lower transition state energy and more exothermic). The relevant values for discussion

are the differences in the transition state energies for reaction of a given nucleophile with

1,3-dMU versus methyl chloride (11.1 kcal mol^{-1} for formate (**Figure 2.13**); 11.8 kcal

mol^{-1} for methyl thiolate (**Figure 2.14**)) and the differences in product energies for

reaction of a given nucleophile with 1,3-dMU versus methyl chloride (7.3 kcal mol^{-1} for

formate (**Figure 2.13**); 7.4 kcal mol^{-1} for methyl thiolate (**Figure 2.14**)). We know from

our acidity calculations that the differences in the product energies are probably due to a

computational artifact: HCl calculates to be 7.7 kcal mol^{-1} less acidic than 3-MeU even

though by experiment they have comparable acidities (*vide supra*). This energy

difference between the acidity values of HCl and 3-MeU corresponds to the ~7 kcal mol^{-1} difference in product energies for the reactions in which chloride and deprotonated 3-MeU are the leaving groups (**Figures 2.13** and **2.14**). The transition state energies, however, show a larger difference for the 1,3-dMU versus methyl chloride reactions (11.1 kcal mol^{-1} for formate and 11.8 kcal mol^{-1} for methyl thiolate, favoring the CH$_3$Cl reaction). The accuracy of these S$_N$2 reaction transition state energies is certainly called into question based on the failure of B3LYP/6-31+G(d) to correctly predict the acidity of HCl. However, DFT methods have been shown to be reasonable for S$_N$2 reaction energetics in degenerate reactions involving chloride as the nucleophile and leaving group.[136-139,150] The consistent *qualitative* conclusion from **Figures 2.13** and **2.14** is that the methyl chloride S$_N$2 reaction appears to have a lower barrier than that of 1,3-dMU. That is, although chloride and deprotonated 3-MeU have similar basicities, chloride is the better leaving group.
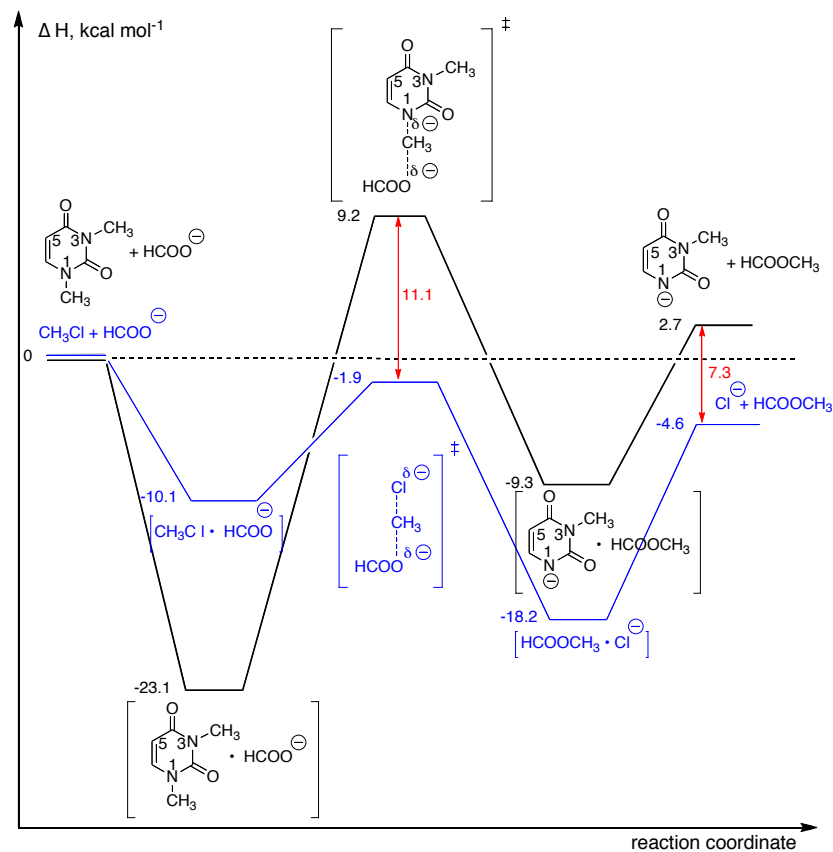
Figure 2.13. Superimposed energy diagrams for reactions of 1,3-dMU and methyl chloride with formate (B3LYP/6-31+G(d), 298 K)
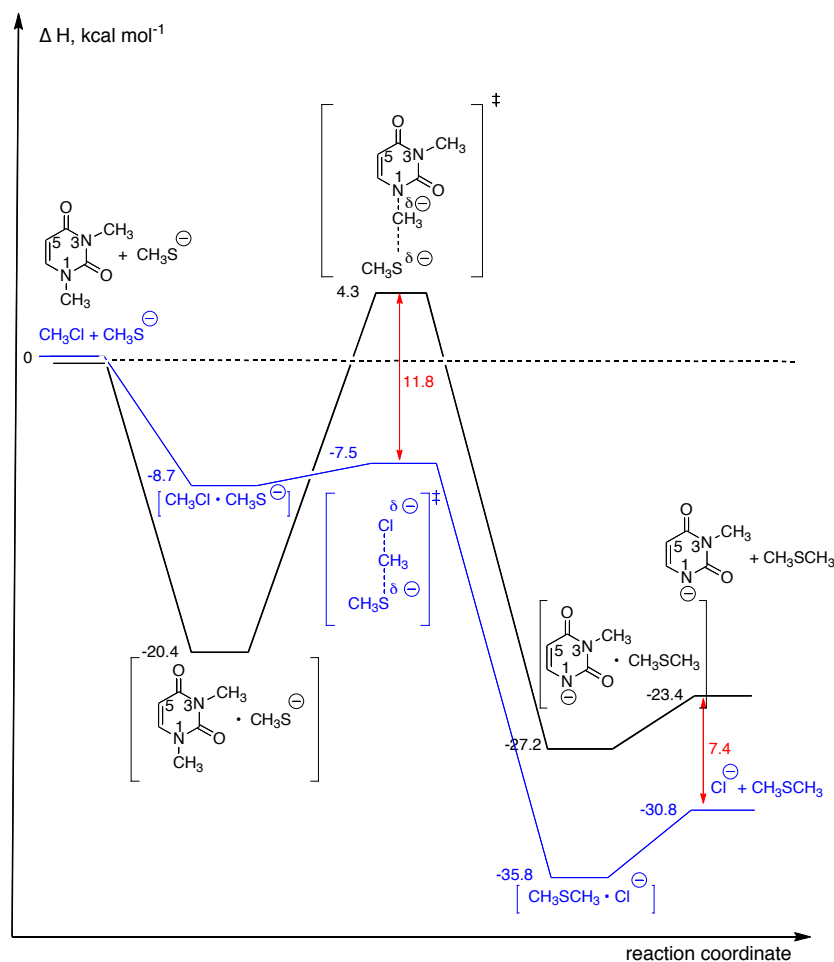
Figure 2.14. Superimposed PES for reactions of 1,3-dMU and methyl chloride with methyl thiolate (B3LYP/6-31+G(d), 298 K)

The experimental S_N2 reaction data qualitatively support the calculations (**Tables 1.2** and **2.4**). For example, for the reaction of 1,1,1-trifluoroethoxide with methyl chloride, the efficiency is 11% (literature) to 13% (our lab). With 1,3-dMU, the efficiency with the same nucleophile is less than 1% (a rate difference of about 15x, corresponding to about a 1.5 kcal mol^{-1} difference in barrier). Thus, both calculations and experiments do indicate that although HCl and 3-MeU have comparable acidities in the gas phase, the conjugate bases do not appear to be equivalently good leaving groups. It appears that the calculations may overestimate that difference (11 kcal mol^{-1} by calculation versus about 2

kcal mol^{-1} by experiment). We understand that our efficiency values for the pyrimidine reactions are incredibly low, and it is in fact almost impossible to fully discount the possibility that the 1,3-dMU sample (and the 1,3-dMT sample) are not contaminated with some monomethyl substrate. Should, for example, 3-methyluracil be present in the 1,3-dMU sample, then deprotonation at N1-H of that contaminant 3-MeU would result in the same product ion expected from the S_N2 reaction. The 1,3-dMU purchased from Sigma Aldrich has a purity of 99%. We synthesized the 1,3-dMT sample from thymine. The compounds appear pure by NMR and also by our mass spectrometric studies of reaction with H_3O^+, where we see only the mass-to-charge ratio corresponding to protonated 1,3-dMU and 1,3-dMT, and do *not* see the m/z ratio corresponding to protonated monomethylated compound. We are therefore quite confident that what we see are very slow S_N2 reactions but cannot discount the possibility that deprotonation of a trace amount of monomethylated compound contributes to the observed product ions. However, if there is such a contaminant, that would mean that the contribution of the S_N2 reaction to the total reaction efficiency would be even lower than the reported values. Therefore, the reported efficiency values for the pyrimidine reactions are an upper limit. Our data indicate that the methyl chloride reactions are on the order of 15 times faster than the 1,3-dMU reactions; this is a lower limit in the sense that our S_N2 reaction efficiencies might be lower than reported.

The reason for this is probably related to the nature of the respective leaving groups: deprotonated uracil is a delocalized ion while chloride is not. In the neutral uracil, the N1 electrons can delocalize into the pi system, but such delocalization separates charge and therefore those resonance structures probably contribute less to the actual structure.

Deprotonation results in a negative charge that can be stabilized by pushing that charge into the carbonyl oxygens (**Figure 2.15**). This is the same argument that, for example, could be used to explain why *N*-bromosuccinimide is an effective brominating agent.

neutral uracil

N1-deprotonated uracil anion
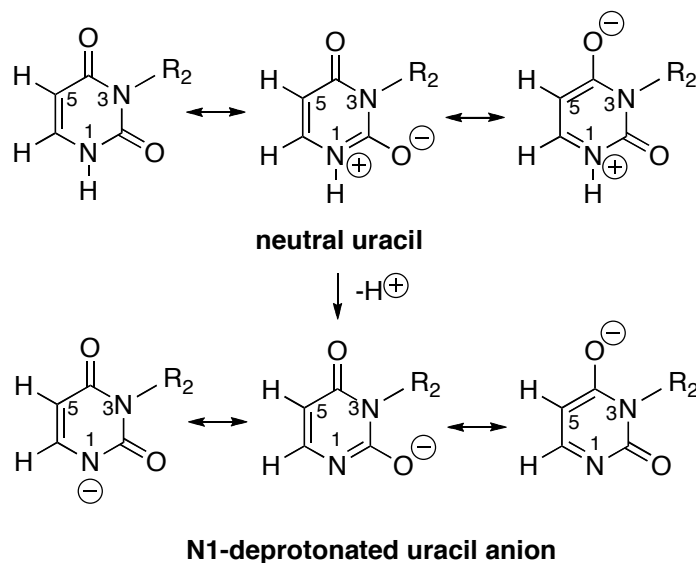
Figure 2.15. Resonance structures of neutral and N1-deprotonated uracil.

Acidity is, of course, a thermodynamic property, while leaving group ability is essentially a kinetic property: how fast is the S_N2 reaction? In the S_N2 reaction transition state, since the N1-deprotonated uracil anion is not yet fully formed, the stability provided by electron delocalization in the product is not yet completely in place. Chloride, on the other hand, is a polarizable entity whose good leaving group properties are likely present in the transition state; little electron reorganization is required. In other words, the special stability enjoyed by the deprotonated uracil N1⁻ anion, while making uracil as thermodynamically acidic as HCl, helps only partially in the leaving group ability: in the S_N2 reaction transition state, the N1 anion is not fully formed and therefore not fully delocalized.[151-153] (Presumably the carbonyl groups also stabilize the anion by induction, but one would expect that effect to be seen in both the acidity and S_N2

reactivity). In summary, delocalization in an anion may make its conjugate acid acidic, but may not affect leaving group ability as much as expected.[151-154]

We examined the bond lengths and charge distribution (CHELPG) for the participants in the S_N2 reaction of formate and methyl thiolate with 1,3-dMU to probe these ideas further. In Table 2.5, we list the distance for the breaking $N1$-CH_3 bond (r1), the distance for the forming Nu-CH_3 bond (r2), and the charges on the O2 and O4. For the reactant, negative charges of –0.560 and –0.579 reside on the O2 and O4, respectively. Those values increase (negatively) to –0.767 and –0.735 in the product. For the reaction with formate, the transition state is predictably a little late (based on the r1 length, relative to the methyl thiolate reaction); the Nu-CH_3 bond is more formed (1.94 Å) and the $N1$-CH_3 bond is quite elongated (2.03 Å). The charge distributions are more negative than in the reactant, but not nearly as negative as in the product (–0.670 and –0.686 for the O2 and O4, respectively). For the reaction with methyl thiolate, the transition state is earlier than that for formate, which would be expected since the reaction with methyl thiolate is more exothermic (methyl thiolate is a more basic nucleophile). The breaking $N1$-CH_3 bond is shorter (1.94 Å) than the forming Nu-CH_3 bond (2.52 Å). Consistent with the earlier TS, the oxygens are not as negative as they are in the formate transition state (–0.641 and –0.683 for methyl thiolate, versus –0.670 and –0.686 for the O2 and O4, respectively, for formate). There does seem to be an interesting paradox for reactions where the leaving group gains stability from delocalization: as the nucleophile becomes more basic, the reaction becomes more exothermic, which is consistent with a slightly lowered barrier (as we see with the formate reaction, which has a $\Delta H^{\ddagger}$ of 9.2 kcal mol^{-1} above the separated reactants versus that of methyl thiolate, which is only 4.3 kcal mol^{-1}).[155,156] However, as

the reaction becomes more exothermic, the transition state may move earlier, where

charge might be less delocalized, which would have an *opposing* effect on rate.
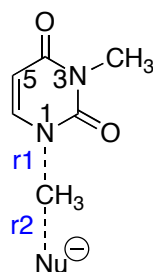
Table 2.5. Calculated (B3LYP/6-31+G(d)) distances and charges (CHELPG) for S_N2

reactions of 1,3-dMU.

SUBSTRATE	$r1$ (N1–CH$_3$ distance) (Å)	$r2$ (CH$_3$–Nu distance) (Å)	charge, O2	charge, O4
1,3-dMU (reactant)	1.47	N/A	–0.560	–0.579
N1-deprotonated 3-MeU (product)	N/A	N/A	–0.767	–0.735
TS with Nu⁻ = HCOO⁻	2.03	1.94	–0.670	–0.686
TS with Nu⁻ = MeS⁻	1.94	2.52	–0.641	–0.683
HCOOCH$_3$ (product)	N/A	1.44	N/A	N/A
CH$_3$SCH$_3$ (product)	N/A	1.83	N/A	N/A

In summary, although uracil is as acidic as HCl, deprotonated uracil is not as good of a

leaving group as chloride due to electron delocalization-related issues. This is not to say

that deprotonated uracil is necessarily a terrible leaving group; it is clearly the species

that leaves in the UDG reaction.[8,115,117,121]

In this study, we are also interested in comparing 1,3-dMU to 1,3-dMT. Since UDG

cleaves uracil but not thymine, we wanted to probe whether there was an intrinsic

reactivity component that would favor uracil cleavage. The calculations for the formate

and methyl thiolate reactions with 1,3-dMU versus 1,3-dMT indicate that the barriers for

the 1,3-dMU reactions are usually about 1 kcal mol⁻¹ lower (relative to the separated

reactants) than the reactions with 1,3-dMT. The measured reaction efficiencies for the uracil reactions are consistently higher than for the analogous thymine reactions (**Table 2.4**). The efficiency values are very small, but the 1,3-dMU reactions are always at least ten times higher in efficiency than the corresponding 1,3-dMT reactions. Assuming that the efficiency results are not due in part to deprotonation reactions of a monomethylated contaminant, it does appear that the uracil reactions are faster than the thymine reactions, which is consistent with the calculations and the measured acidities (wherein 3-MeU is more acidic than 3-MeT, *vide supra*). Thus, intrinsically, deprotonated uracil is more easily cleavable than deprotonated thymine, which could be one factor (of many) aiding in the discrimination between the two by UDG in DNA.

Last, although we cannot be sure that nucleophilic attack occurs at N1 (rather than N3) of 1,3-dMU and 1,3-dMT (**Figure 2.12**), our calculations indicate that the difference in reactivity of those sites is so great (more than 10 kcal mol^{-1}, *vide supra*) that attack at N1 is likely. Future studies with appropriately deuterated substrates were proposed.[157]

2.5. Conclusions

We find that although the acidity of HCl and 3-methyluracil are comparable in the gas phase, the leaving group abilities of chloride and deprotonated 3-methyluracil are different, with chloride being slightly better. The basis of this difference lies in the fact that deprotonated 3-methyluracil is thermodynamically very acidic due to delocalization, which does not yield as large of a beneficial effect in the S_N2 reaction transition state.

Comparison of calculations to experiments indicate that the B3LYP/6-31+G(d) estimates the acidity of the pyrimidine derivatives well, but not that of HCl (the error is on the order of 7 kcal mol^{-1}, with the calculations yielding too low of a value). In terms

of the S_N2 reactions, the calculations predict that those with methyl chloride would be favored over those of 3-methyluracil. The experiments also indicate that the methyl chloride reactions are faster than those of 1,3-dimethyluracil, with rates of at least 15-20 times more.

We also compared the leaving group ability of 1,3-dimethyluracil versus 1,3-dimethylthymine. Our calculations and experiments indicate that the N1-deprotonated uracil derivative is a slightly better leaving group than the N1-deprotonated thymine derivative, which is consistent with the cleavage of the former but not the latter from DNA.

Note: Major parts of the following chapter have been published: Michelson A. Z.; Petronico, A.; Lee, J. K. *J. Org. Chem.* **2012**, *77,* 1623–1631.

Chapter 3. 2-Pyridone and Derivatives: Gas Phase Acidity, Proton Affinity, Tautomer Preference and Leaving Group Ability

3.1. Introduction

We have long been interested in the leaving group ability of nucleobases, particularly damaged ones, in relation to the mechanism of the enzymes that remove such bases from DNA. Many of our recent studies have focused on uracil, which is an RNA base that can be mutagenic when it occurs in DNA.[5,10,13,77,122]

Uracil is removed from the genome by the enzyme uracil DNA glycosylase, which has been shown to involve N1-deprotonated uracil as the leaving group (**Figure 3.1**).[5,122] In earlier work described in **Chapter 2** and previous Lee and co-workers papers, we examined the properties of uracil in the gas phase; with a dielectric of 1, the gas phase is an "ultimate" nonpolar environment and can therefore potentially lend insight into reactivity in other nonpolar media, including enzyme active sites.[10,13,77]

We found that the intrinsic, gas-phase acidity of uracil is comparable to that of hydrochloric acid. Because the strength of an acid generally correlates with the leaving group ability of its conjugate base in the gas phase, deprotonated uracil may be a very good leaving group, comparable to chloride.
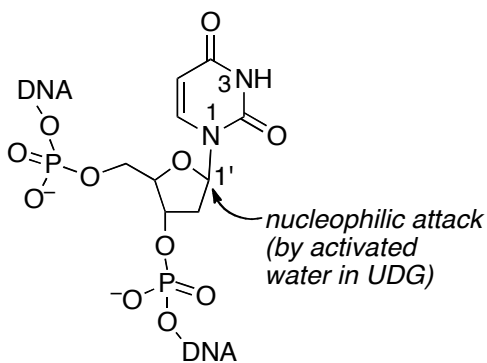
Figure 3.1. Uracil is removed from a genome by uracil DNA glycosylase (UDG).

However, when we examined hydrochloric acid and 3-methyluracil in the gas phase, we found that despite similar acidities, chloride is a better leaving group than N1-deprotonated 3-methyluracil (**Figure 3.2**).[77,158] We proposed that the reason for the disparity between acidity and leaving group ability could be due to the resonance delocalization in the N1-deprotonated 3-methyl uracil anion versus that is absent in the chloride ion. Deprotonated 3-methyluracil is thermodynamically stable due to delocalization by resonance (**Figure 3.2c**); however, that delocalization might not be fully realized in an S_N2 transition state. Therefore, the stabilizing benefit of resonance delocalization is not as evident in leaving group ability, and deprotonated 3-methyl uracil is not as good a leaving group as chloride ion.
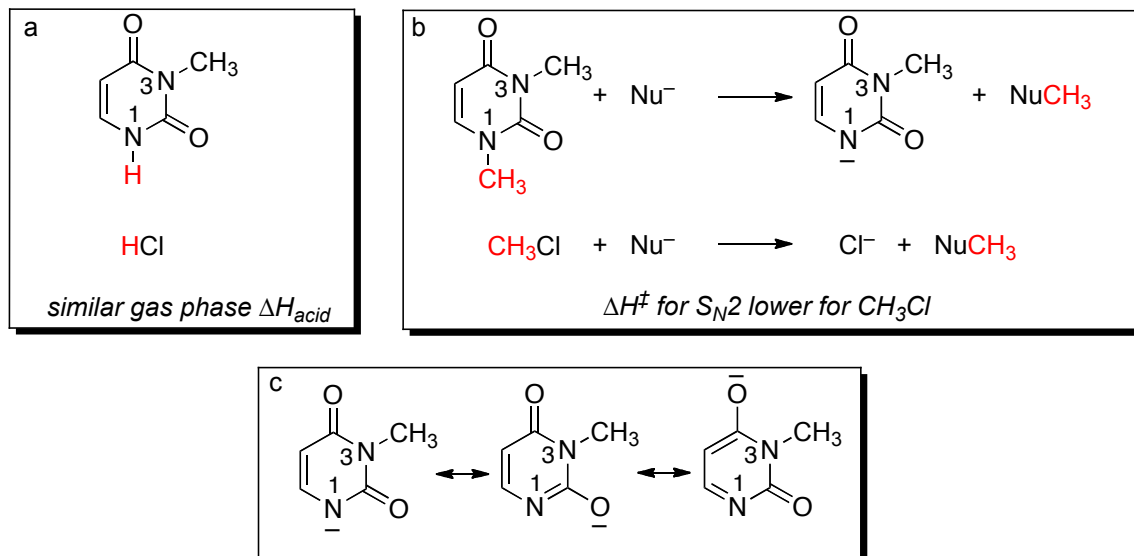
Figure 3.2. Comparison of **(a)** acidity and **(b)** leaving group ability for 3-methyluracil and hydrogen chloride. Resonance delocalization in N1-deprotonated 3-methyluracil anion is also shown **(c).**

To further probe this hypothesis, we would need a model system where we could systematically compare resonance-stabilized and non-resonance-stabilized anionic leaving groups. We would expect a closer correlation between acidity and leaving group ability for the non-resonance-stabilized anions than the resonance-stabilized anions. Toward that end, we decided to examine the effect of substitution on a series of 2-pyridones (**Figure 3.3**). We chose the pyridone system for various reasons, including simplicity, resemblance to uracil, and the plan that changing substituent "X" would allow us to probe effects systematically.
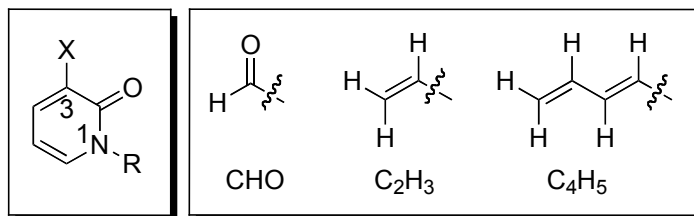
X = H, Cl, Br, CHO, C$_2$H$_3$, C$_4$H$_5$
R = H (for acidity studies); CH$_3$ (for S$_N$2 studies)

Figure 3.3. 2-Pyridone and 3-substituted derivatives studied herein.

Although understanding the substituent effects on leaving group ability was the initial motivation for this study, pyridones are also of interest in their own right. The keto-enol tautomerism of the parent 2-pyridone has been much studied in the last century; pyridone/hydroxypyridine is considered a prototypical model for hydrogen bonding, tautomerization, and proton shuttling in both chemical and biological systems, including those involving nucleobases.[159-163] Aqueous studies point to the keto form (**1a**); gas phase studies indicate a mixture, but with a 2-hydroxypyridine (**1b**) preference.[164-170] In this study, we measure the acidity and proton affinity of various derivatives not heretofore examined, which establishes fundamental properties as well as giving insight to tautomer presence.
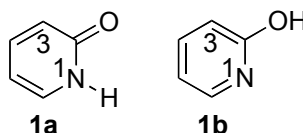
3.2. Experimental Section

All pyridones as well as reference acids and bases are commercially available and were used as received.

3.2.1. Bracketing experiments

Acidity and proton affinity bracketing experiments were conducted using a Finnigan 2001 Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR) with a dual cell set up according to the protocol described in details previously and in **Chapter 1**.[10,13,35,36,38-41] Proton transfer reactions were conducted in both directions. Solid pyridones were introduced to the cell via a solids probe and slightly heated if necessary. Liquid reference acids or bases were introduced via a heatable batch inlet system. Hydroxide or hydronium ions were produced from water pulsed into the cell, and ionized by an electron beam (typically 8 eV (for OH^-), 20 eV (for H_3O^+), 6 μA, 0.5 s).

Bracketing experiments were run under pseudo-first order conditions since the amount of the neutral reactant was always in excess, relative to the reactant ions. Reading the pressure of the neutral compounds from the ion gauges is not always accurate; therefore we "back out" the neutral substrate pressure from fast control reactions (described previously and in **Chapter 1**).[38-41,79,80]

3.2.2. Cooks kinetic method

We also used the Cooks kinetic method in a Finnigan quadrupole ion trap (LCQ) mass spectrometer[84-87,135] to measure proton affinity and acidity. The proton-bound complex ions are generated by electrospray (ESI).[72] For each experiment, a solution of the pyridone (250 μM) and reference acid or base (250 μM) is prepared (in 20% methanol/80% water). An electrospray needle voltage of ~4 kV was used. The flow rate is 25 μL/min. The proton-bound complex ions were isolated and then dissociated by applying collision-

induced dissociation (CID); the complexes were activated for about 30 ms. A total of

forty scans was averaged for the product ions.

Cooks kinetic method involves the formation of a proton bound complex, or dimer, of

the conjugate bases of the unknown AH and a reference acid B_iH of known acidity (**Eq.**

1). The same can be done for proton affinity, where a positively charged proton-bound

dimer is formed.

$$[AHB_i]^{\ominus} \overset{k_1}{\underset{k_2}{\diagdown}} \begin{array}{c} [B_iH] + A^{\ominus} \\ \\ [AH] + B_i^{\ominus} \end{array} \qquad \text{eq. 1}$$
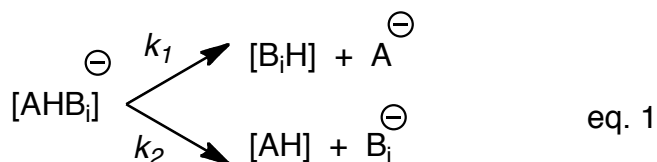
$$\ln(k_1/k_2) = (1/RT_{eff})(\Delta H_{BiH} - \Delta H_{AH}) \qquad \text{eq. 2}$$

The proton bound dimer $[AHB_i]^-$ is dissociated via collision-induced dissociation

(CID). The rate constants k_1 and k_2 are for the two different dissociation pathways. The

relationship of these rate constants to ΔH_{acid} is shown in **Eq. 2**. R is the gas constant and

T_{eff} is the effective temperature[88] of the activated dimer.[84-87,135] The ratio of the amounts

(intensities) of the two deprotonated products yields the relative acidity of the two

compounds of interest, assuming the dissociation has no reverse activation energy barrier

and that the dissociation transition structure is late and therefore indicative of the stability

of the two deprotonated products. These assumptions are generally true for proton bound

systems. To obtain the acidity of compound AH, the natural logarithm of the relative

intensity ratios is plotted versus the acidities for a series of reference acids, where the

slope is $(1/RT_{eff})$ and the y-intercept is $(-\Delta H_{AH}/RT_{eff})$. The T_{eff} is obtained from the slope.

The acidity of compound AH, (ΔH_{AH}) is calculated from either **Eq. 2** or the y-intercept.

3.2.3. Calculations

Calculations were conducted at B3LYP/6-31+G(d)[92-94] and M06-2X/6-311+G(2df,2p)[95,96] level as implemented in Gaussian09.[91] The geometries were fully optimized and frequencies were calculated. No scaling factor was applied. All the values reported are at ΔH at 298 K. The acidity and PA values include the enthalpy of the proton at 298 K (1.5 kcal mol^{-1}). All calculated transition state structures have one negative frequency.

3.3. Results and Discussion.

The substrates we considered are shown in **Figure 3.3**. We chose substitution at the 3-position as this allowed for a negative charge on N1 to delocalize into resonance-stabilizing groups ($–C_2H_3$, C_4H_5, CHO). Chloride and bromide serve as electron-withdrawing groups into which charge cannot delocalize by resonance.

Our first goal was to benchmark calculations by examining the commercially available parent, 3-chloro- and 3-formyl-2-pyridones experimentally and theoretically. These three were chosen as models for substrates with no substitution, substitution with a moiety that does not provide resonance delocalization for an anion at N1, and substitution with a resonance-stabilizing group, respectively.

3.3.1. 2-Pyridone.

3.3.1.1. Calculations: 2-pyridone tautomers, acidity, proton affinity.

The keto-enol tautomerism of the parent pyridone system has been theoretically examined quite extensively in the past several decades.[162] The two tautomers, 2-pyridone (PY, **1a**) and 2-hydroxypyridine (HP, **1b**) appear to have less than a 1 kcal mol^{-1}

difference in stability in the gas phase (with HP being more stable), which makes it a challenging computational system to examine.

Although DFT methods generally are known to reverse the PY/HP tautomer relative energies, they do generate reliable molecular structures.[171-177] Incorrect DFT energies have been shown to arise from the exchange potentials.[171]

Because we are interested not just in the relative tautomer stabilities, but also the thermochemical properties (proton affinity and acidity, which in our previous studies of nucleobases are well calculated by DFT methods), we calculated the possible tautomers of pyridone using B3LYP/6-31+G(d) (**Figure 3.4**).[10,13,35-41,76] As expected, B3LYP/6-31+G(d) incorrectly predicts that PY tautomer **1a** should be more stable than HP tautomer **1b** (**Figure 3.4**). (The other possible enol structure **1b'** is 7 kcal mol^{-1} less

| (0) | (1.6) | (7.1) |
| **1a** | **1b** | **1b'** |

stable than **1a**, **Figure 3.4**).
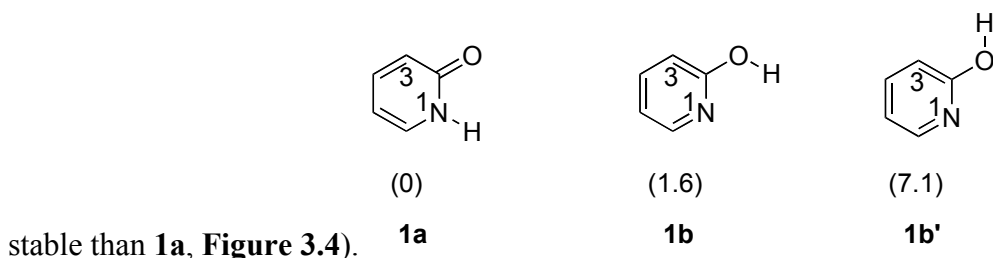
Figure 3.4. Pyridone calculations at B3LYP/6-31+G(d). Values in parentheses are relative stabilities. All are ΔH_{298} values, in kcal mol^{-1}.

We also calculated the pyridone stabilities and properties using M06-2X/6-311+G(2df,2p) (**Figure 3.5**).[95,96,148] Tautomerism of PY/HP has not been examined by this relatively new suite of density functionals. Because recent papers have shown the accuracy of this method for predicting a wide range of chemistry, we wished to probe whether it could serve as a reasonable method for predicting stability and thermochemical properties in these systems.[95,96,148] M06-2X/6-311+G(2df,2p) correctly predicts the

higher stability of the enol form HP (**1b**, **Figure 3.5**). (The other possible enol structure, with the proton pointing "toward" C3 (**1b'**) is 5 kcal mol^{-1} less stable than **1b** at this level). The relative energy of the keto (PY) form **1a** (+1.6 kcal mol^{-1}) is slightly higher than that found by gas-phase experiments (which predict less than 1 kcal mol^{-1}), but is still a fairly reasonable calculational estimate.[163,165,178,179] More computationally intensive methods (G3, G4, CBS-APNO) yield more accurate values, but the faster M06-2X method is surprisingly quite comparable to CBS-APNO (which gives a relative stability of HP to PY as 1.3 kcal mol^{-1}).[162]

The acidity (ΔH_{acid}) and proton affinity (PA, which is $-\Delta H$ for protonation)[33] of the 2-pyridone structures at M06-2X/6-311+G(2df,2p) are also shown in **Figure 3.5**. In terms of acidity, the more stable enol **1b** and the keto structure **1a** have similar values (347.9 versus 346.3 kcal mol^{-1}). The PAs, however, may allow for differentiation between the two tautomers: the most basic site of enol **1b** is calculated to be 213.0 kcal mol^{-1}, while for keto **1a** it is 218.7 kcal mol^{-1}.[180] The enol **1b'** is significantly higher in energy than the other two structures and is unlikely to be present.

Figure 3.5. 2-Pyridone calculations at M06-2X/6-311+G(2df,2p). Values in parentheses are relative stabilities. Proton affinity values are in blue; acidity values are in red. All are ΔH_{298} values, in kcal mol^{-1}.

3.3.1.2. Experiments: 2-pyridone acidity.

We measured the acidity of 2-pyridone using acidity bracketing (details in experimental section).[181] In the bracketing experiment (**Table 3.1**), a proton transfer occurs from acetic acid (ΔH_{acid} = 347.4 ± 0.5 kcal mol^{-1}) to deprotonated pyridone; the opposite reaction also occurs (that is, acetate deprotonates 2-pyridone), placing the acidity (ΔH_{acid}) of 2-pyridone at 347 ± 3 kcal mol^{-1}.

Table 3.1. Summary of results for acidity bracketing of 2-pyridone (**1**).

Reference compound	$\Delta H_{acid}{}^{a}$	Proton transfer[b]	
		Ref. acid	Conj. base
i-propylthiol	354.6 ± 0.5	−	+
n-pentanethiol	352.5 ± 2.3	−	+
m-cresol	349.6 ± 2.1	−	+
acetic acid	347.4 ± 0.5	+	+
butyric acid	346.8 ± 2.0	+	−
formic acid	346.0 ± 0.5	+	−
2,4-pentanedione	343.8 ± 2.1	+	−

$^{a}\Delta H_{acid}$ is in kcal mol^{-1}.[33,182] ^{b}A "+" indicates the occurrence and a "−" indicates the absence of proton transfer.

3.3.1.3. Experiments: 2-pyridone proton affinity.

We measured the proton affinity (PA) of 2-pyridone using two complementary methods: PA bracketing and the Cooks kinetic method (details in experimental section). In the bracketing experiment, the reaction is found to proceed in both directions for *N*-

methylaniline (PA = 219.1 ± 2.0 kcal mol^{-1}), yielding a PA of 219 ± 3 kcal mol^{-1} (**Table 3.2**).

Table 3.2. Summary of results for proton affinity bracketing of 2-pyridone (**1**).

Reference compound	PA^a	Proton transfer^b	
		Ref. base	Conj. acid
N-ethylaniline	221.0 ± 2.0	+	−
n-butylamine	220.2 ± 2.0	+	−
N-methylaniline	219.1 ± 2.0	+	+
3-methylpyrazole	216.5 ± 2.0	−	+
2-chloropyridine	215.3 ± 2.0	−	+
o-toluidine	212.9 ± 2.0	−	+
pyrrole	209.2 ± 2.0	−	+

^{a} PA is in kcal mol^{-1}.[33] ^{b}A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

Using the Cooks kinetic method with reference bases 4-methylpyrazole (PA = 216.7 ± 2.0 kcal mol^{-1}), *N,N*-dimethylacetamide (PA = 217.0 ± 2.0 kcal mol^{-1}), *N*-benzylamine (PA = 218.3 ± 2.0 kcal mol^{-1}), *N*-methylaniline (PA = 219.1 ± 2.0 kcal mol^{-1}), L-phenylalanine (PA = 220.6 ± 2.0 kcal mol^{-1}), and cyclohexylamine (PA = 223.3 ± 2.0 kcal mol^{-1}) yields a PA of 218 ± 3 kcal mol^{-1}.

3.3.1.4. Tautomer composition: 2-pyridone.

Therefore, our experiments indicate a ΔH$_{acid}$ of 347 kcal mol^{-1}. At M06-2X/6-311+G(2df,2p), both the keto (**1a**) and enol (**1b**) tautomers have acidities close to this

value (346.3 and 347.9 kcal mol⁻¹), so the experimental acidity cannot be used to ascertain what tautomers are present.

The M06-2X/6-311+G(2df,2p) calculated PA for the enol tautomer **1b** is 213.0 kcal mol⁻¹ (for the most basic site, which is the ring nitrogen); for the keto tautomer **1a**, the computed PA (at the carbonyl) is 218.7 kcal mol⁻¹. We find that calculated proton affinities using this same method and level for a series of model compounds whose PAs are well known experimentally (cyclohexanone, *N*-methyl-2-pyridone, *N*-methylacetamide, 3,5,5-trimethyl-2-cyclohexen-2-one) are accurate to within 1 kcal mol⁻¹ (**Figure 3.6**).[33] Therefore, the bracketed PA value of 219 kcal mol⁻¹ implies that under our conditions, the keto tautomer is present.
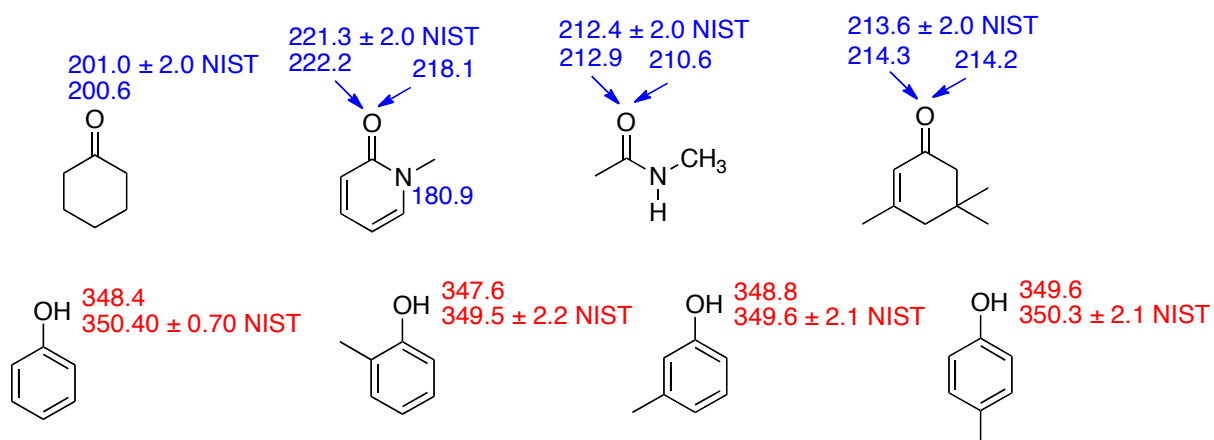
Figure 3.6. Benchmarking calculations results (M06-2X/6-311+G(2df,2p)) and comparison with literature (NIST) values. Values in blue are gas phase PA, values in red are gas phase acidities (in kcal mol⁻¹); 298 K.

As mentioned earlier, previous gas phase experiments indicate a mixture of the keto and enol tautomers. In the PA bracketing experiment, as long as the neutral keto tautomer **1a** is present, it will deprotonate protonated reference bases with PAs around 219 kcal

mol^{-1} and lower ("+" in the rightmost column of **Table 3.2**). We cannot know whether the enol **1b** is also present; the enol may contribute to the "+" reactivity at lower PA values but there is no way to discern that.[183] In the opposite direction, we find that *only* bases with PAs of 219 kcal mol^{-1} and higher deprotonate protonated 2-pyridone. This experimental result implies the presence of the protonated structure **2** but not **3** (**Figure 3.7**), since if **3** were present, one would expect reference bases in the 213-215 kcal mol^{-1} range to deprotonate the protonated 2-pyridone. At M06-2X/6-311+G(2df,2p), **2** is more stable than **3** by 4.2 kcal mol^{-1}. Presumably, once 2-pyridone is protonated, reaction (**Scheme 3.1**) or rearrangement to the more stable form **2** may occur, leading to deprotonation only by bases with PAs higher than 219 kcal mol^{-1}.[10] Practically speaking, we bracket the more basic tautomer present and cannot be sure that the tautomer with a lower PA is not also present.

Figure 3.7. Possible structures for protonated 2-pyridone. Values in parentheses are relative stabilities. Enthalpy required to deprotonate protons are in blue. All are ΔH_{298} values, in kcal mol^{-1}, calculated at M06-2X/6-311+G(2df,2p).

Scheme 3.1
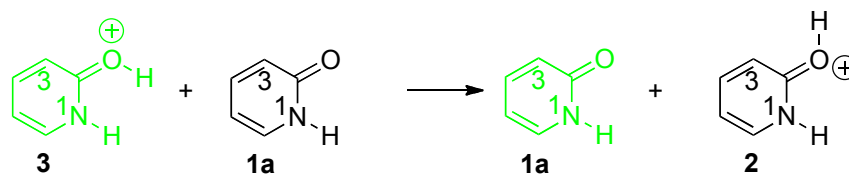
The Cooks kinetic experiment is interesting as a complementary method in that the pyridone is not vaporized - rather, a proton-bound dimer of the pyridone and reference base (**4**) is electrosprayed from aqueous (20% methanol, 80% water) solution. The proton-bound dimer is then isolated in the mass spectrometer and energy is applied (collision-induced dissociation, CID, details in experimental section). The measured PA value of 218 kcal mol^{-1} implies that we measure the proton affinity on the "C3 side" (the face of the oxygen facing C3, not N1) of the keto structure **1a**.[40] Because the proton-bound dimer is electrosprayed from a water solution, this result indicates that pyridone probably exists as the keto tautomer in aqueous solution, which is consistent with previous solution phase experimental data.[162-164,184,185]

$$\left[\text{pyridinone}\cdots\text{H}\cdots\text{reference base}\right]^{\oplus}$$
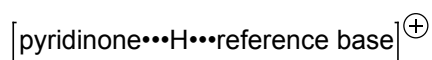
4

Thus, in our experiments, whether we vaporize 2-pyridone from the solid phase or electrospray a proton-bound dimer of 2-pyridone with a reference base, we measure a PA that is consistent with the calculated PA of the keto structure. This result does not discount the possibility of a keto-enol mixture; we can only say that the keto form is present.

To validate the comparison of calculations to experiment, we also calculated the PA of *N*-methyl-2-pyridone (**5**) at M06-2X/6-311+G(2df,2p); methylation of the N removes the possibility of multiple tautomers and "locks" the pyridone into keto form. The calculated PA of the most basic site is 222.2 kcal mol^{-1}. The literature value is 221.3 $\pm$ 2.0 kcal mol^{-1}[33,89] which we also confirmed by bracketing the PA in our FTMS (**Table 3.3**). The calculated and measured values are therefore consistent, and support our conclusion for

the *parent* (N-H) 2-pyridone: the measured PA of 219 kcal mol^{-1} corresponds to the keto

form (calculated PA of 218.7 kcal mol^{-1}).
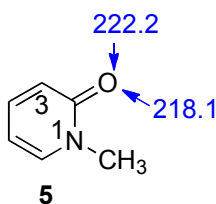
Table 3.3. Summary of results for proton affinity bracketing of N-methyl-2-pyridone

(**5**).

Reference compound	PA^a	Proton transfer	
		Ref. base	*Conj acid*
3-picoline	225.5 ± 2.0	+	−
cyclohexylamine	223.3 ± 2.0	+	−
pyridine	222.0 ± 2.0	+	+
N-ethylaniline	221.0 ± 2.0	+	+
n-butylamine	220.2 ± 2.0	−	+
N-methylaniline	219.1 ± 2.0	−	+

^{a} *PA* is in kcal mol^{-1}. ^{b}A "+" indicates the occurrence and a "−" indicates the absence of

proton transfer

3.3.2. 3-Chloro-2-pyridone.

3.3.2.1. Calculations: 3-chloro-2-pyridone tautomers, acidity, proton

affinity.

The calculated values at M06-2X/6-311+G(2df,2p) for the acidity and proton affinity

for the possible tautomers of 3-chloro-2-pyridone are shown in **Figure 3.8.** As with the

parent pyridone, calculations predict that the more stable enol tautomer (**6b**, with the proton "pointing" toward N1) will be slightly more stable than the keto **6a**, by 1.2 kcal mol^{-1}. Both tautomers have similar acidity (338-339 kcal mol^{-1}), but the proton affinities of the most basic sites differ by 8 kcal mol^{-1} (208.0 (enol) versus 215.9 kcal mol^{-1} (keto)). There is also the other enol structure (with the proton "pointing" toward C3, **6b'**) that is 2.1 kcal mol^{-1} less stable than the enol tautomer **6b**; its PA and acidity are comparable to the keto form. In the parent pyridone, the analogous enol form (H "pointing" toward C3, **1b'**) is not particularly stable. However, in this 3-chloro compound, a stabilizing interaction between the 3-Cl and the 2-OH exists (calculated Cl-H distance is 2.4 Å), stabilizing the tautomer.
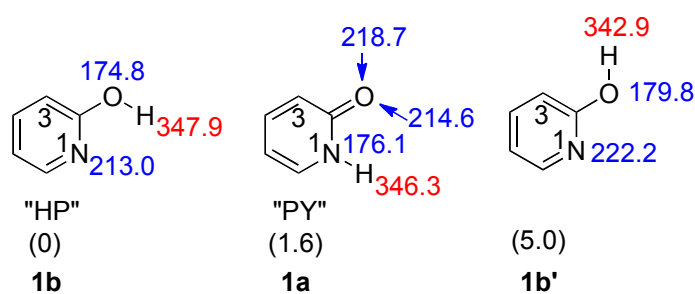

Figure 3.8. 3-Chloro-2-pyridone calculations at M06-2X/6-311+G(2df,2p). Values in parentheses are relative stabilities. Proton affinity values are in blue; acidity values are in red. All are ΔH_{298} values, in kcal mol^{-1}.

3.3.2.2. Experiments: 3-chloro-2-pyridone acidity.

We measured the acidity of 3-chloro-2-pyridone using both bracketing and the Cooks kinetic method. In the bracketing experiment (**Table 3.4**), the reaction between the deprotonated 3-chloro-2-pyridone and trifluoro-*m*-cresol ($\Delta H_{acid} = 339.3 \pm 2.1$ kcal mol^{-1})

proceeds, as does the reaction in the opposite direction (trifluoro-*m*-cresolate with neutral 3-chloro-2-pyridone), yielding a ΔH_{acid} value of 339 ± 3 kcal mol^{-1}.

Table 3.4. Summary of results for acidity bracketing of 3-chloro-2-pyridone (**6**).

Reference compound	$\Delta H_{acid}{}^a$	Proton transfer[b]	
		Ref. acid	Conj. base
formic acid	346.0 ± 0.5	−	+
2,4-pentanedione	343.8 ± 2.1	−	+
methyl cyanoacetic acid	340.8 ± 0.6	−	+
trifluoro-*m*-cresol	339.3 ± 2.1	+	+
2-chloropropionic acid	337.0 ± 2.1	+	−
malononitrile	335.8 ± 2.1	+	−
pyruvic acid	333.5 ± 2.9	+	−

$^a\Delta H_{acid}$ is in kcal mol^{-1}.[33,182] ^{b}A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

Using the Cooks kinetic method and reference acids anthranilic acid ($\Delta H_{acid} = 337.3 \pm 2.2$ kcal mol^{-1}), 2,6-dimethylbenzoic acid ($\Delta H_{acid} = 338.4 \pm 2.1$ kcal mol^{-1}), trifluoro-*m*-cresol ($\Delta H_{acid} = 339.3 \pm 2.1$ kcal mol^{-1}), benzoic acid ($\Delta H_{acid} = 340.1 \pm 2.2$ kcal mol^{-1}), and methoxyacetic acid ($\Delta H_{acid} = 341.9 \pm 2.1$ kcal mol^{-1}) gives a ΔH_{acid} of 340 ± 3 kcal mol^{-1}.

3.3.2.3. Experiments: 3-chloro-2-pyridone proton affinity.

3-Methylpyrazole (PA = 216.5 ± 2.0 kcal mol^{-1}) deprotonates protonated 3-chloro-2-pyridone, but 2-chloropyridine (PA = 215.3 ± 2.0 kcal mol^{-1}) does not (**Table 3.5**). In the opposite direction, 3-chloro-2-pyridone deprotonates protonated 2-chloropyridine, but not

protonated 3-methylpyrazole. Therefore the bracketed PA of 3-chloro-2-pyridone is 216 $\pm$ 3 kcal mol^{-1}.

Table 3.5. Summary of results for proton affinity bracketing of 3-chloro-2-pyridone (**6**).

Reference compound	PA^a	Proton transfer^b	
		Ref. base	Conj. acid
N-methylaniline	219.1 $\pm$ 2.0	+	–
N,N-dimethylacetamide	217.0 $\pm$ 2.0	+	–
3-methylpyrazole	216.5 $\pm$ 2.0	+	–
2-chloropyridine	215.3 $\pm$ 2.0	–	+
m-toluidine	214.7 $\pm$ 2.0	–	+
o-toluidine	212.9 $\pm$ 2.0	–	+
pyrrole	209.2 $\pm$ 2.0	–	+

^{a} PA is in kcal mol^{-1}.[33] ^{b} A "+" indicates the occurrence and a "–" indicates the absence of proton transfer

In the Cooks kinetic method experiment, reference bases 2-chloropyridine (PA = 215.3 $\pm$ 2.0 kcal mol^{-1}), anthranilic acid (PA = 215.5 $\pm$ 2.0 kcal mol^{-1}), 3-methyl pyrazole (PA = 216.5 $\pm$ 2.0 kcal mol^{-1}), 4-methylpyrazole (PA = 216.7 $\pm$ 2.0 kcal mol^{-1}), and N,N-dimethylacetamide (PA = 217.0 $\pm$ 2.0 kcal mol^{-1}) were used, yielding a PA of 216 $\pm$ 3 kcal mol^{-1}.

3.3.2.4. Tautomer composition: 3-chloro-2-pyridone.

The ΔH_{acid} of 3-chloro-2-pyridone, regardless of method used, is measured to be 339-340 kcal mol^{-1}. Since the calculated acidity of the enol and keto tautomers are in the same range (337-339 kcal mol^{-1} for the three different structures **6a**, **6b**, **6b'**), the acidity is not indicative of which tautomers may be present. The bracketed PA is 216 kcal mol^{-1}.

As with the parent pyridone, when compared to calculations, the measured PA does not correspond to the most stable enol tautomer **6b**. In this case, the bracketed PA is consistent with either the keto form **6a** or the less stable enol form **6b'**. The latter is calculated to be 2.1 kcal mol^{-1} less stable than the more stable enol **6b**; if this estimate is accurate, this particular form should constitute a relatively small portion of the tautomer mixture. In the following discussion, therefore, we will focus on the keto form **6a** and the enol form **6b**.

In the bracketing experiment, the neutral 3-chloro-2-pyridone is able to deprotonate conjugate acids of reference bases with PAs 215 kcal mol^{-1} and lower (rightmost column, **Table 3.5**). This is consistent with the PA of keto form (calculated PA = 215.9 kcal mol^{-1}), indicating the presence of the keto tautomer **6a**. At PAs less than 208 kcal mol^{-1}, the enol tautomer **6b** could also be reacting, but we would not be able to discern its contribution to the overall reactivity. In the opposite direction, reference bases below 216.5 kcal mol^{-1} cannot deprotonate protonated 3-chloro-2-pyridone. This would imply the presence of structure **7**, but not of structure **8**. As with the parent pyridone, we speculate that if **8** is present, it converts under our conditions to the more stable protonated form **7** (more stable than protonated **8** by 6.7 kcal mol^{-1}).

As with the parent pyridone, therefore, we can conclude that the keto (**6a**) tautomer is present, but cannot discount a mixture that possibly also includes enol tautomer (**6b**), since the nature of the experiment dictates that we bracket only the most basic tautomer. Given that calculations indicate a roughly 1 kcal mol^{-1} difference in stability for **6a** versus **6b**, a mixture is probable.
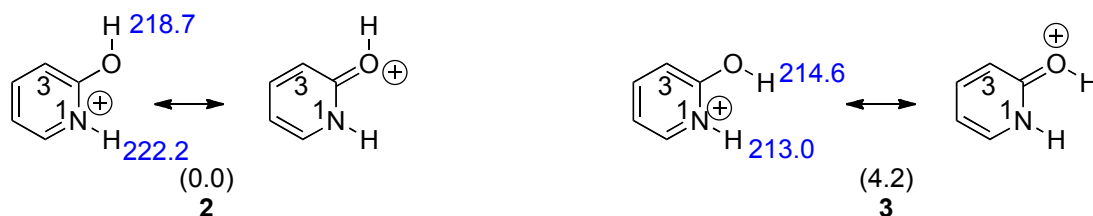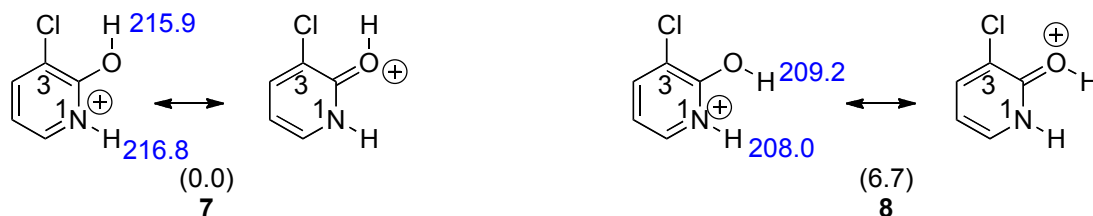
Figure 3.9. Possible structures for protonated 3-chloro-2-pyridone. Values in parentheses are relative stabilities. Enthalpy required to deprotonate protons are in blue. All are ΔH_{298} values, in kcal mol^{-1}, calculated at M06-2X/6-311+G(2df,2p).

The Cooks kinetic PA value is also 216 kcal mol^{-1}, which indicates that we are measuring the carbonyl O (on the "C3" side or face) of **6a**, implying the keto form dominates in solution.[40,185]

3.3.3. 3-Formyl pyridone.

3.3.3.1. Calculations: 3-formyl-2-pyridone tautomers, acidity, proton affinity

3-Formyl pyridone is somewhat more complicated in that several different structures are possible due to both the formyl and the enol moieties. The three lowest energy structures at M06-2X/6-311+G(2df,2p) are shown in **Figure 3.10**. The remaining structures are over 4 kcal mol^{-1} less stable than the most stable tautomer **9b'**; all are shown in **Figure 3.11**). As with the parent and 3-chloro-2-pyridone, an enol structure (**9b'**) is predicted to be most stable in the gas phase. However, in this case, the proton of the most stable enol is pointing "toward" the C3. This is in contrast to the parent and 3-chloro derivatives, where the analogous structures (**1b'** and **6b'** for the parent and 3-chloro, respectively) were the least stable. The high stability of this structure is due to the

internal hydrogen bond that exists between the enol H and the carbonyl O (calculated distance of 1.8 Å). The parent pyridone has no such hydrogen bond, so structure **1b'** is quite unstable, relative to **1b** (**Figure 3.5**). In the 3-chloro-2-pyridone, the analogous structure **6b'** is somewhat stabilized by a weak internal hydrogen bond (calculated Cl-H distance of 2.4 Å). Interestingly, with the formyl system, the stabilities are reversed and **9b'** becomes the most stable structure. The calculated acidities and proton affinities are also shown in **Figure 3.10** for the three tautomers.
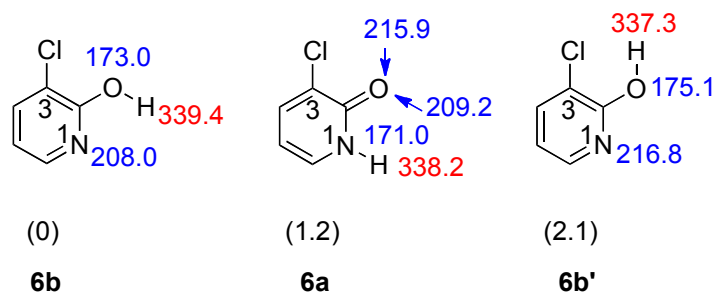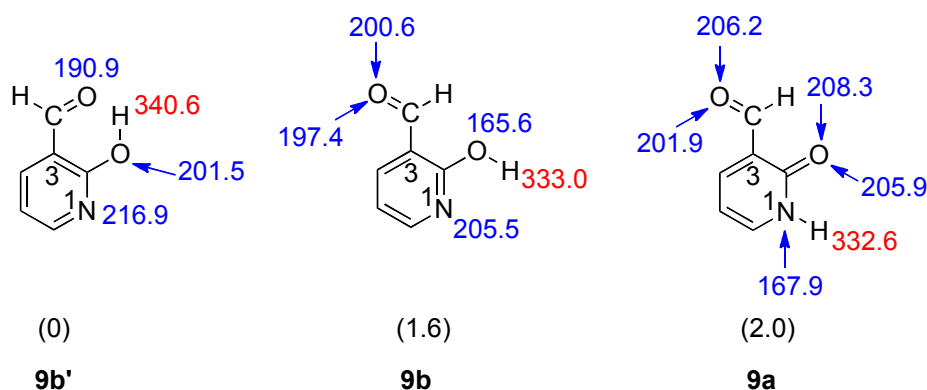

Figure 3.10. 3-Formyl-2-pyridone calculations at M06-2X/6-311+G(2df,2p). Values in parentheses are relative stabilities. Proton affinity values are in blue; acidity values are in red. All are ΔH_{298} values, in kcal mol^{-1}.
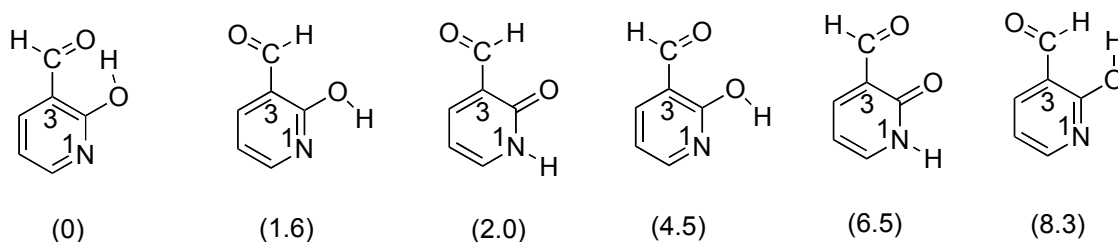
Figure 3.11. Calculated (M06-2X/6-311+G(2df,2p)) relative enthalpies (in kcal mol^{-1}) for 6 possible tautomers of 3-formyl-2-pyridone.

3.3.3.2. Experiments: 3-formyl-2-pyridone acidity

Using the bracketing method, we find that deprotonated 3-formyl pyridone does not deprotonate 2-chloropropionic acid (ΔH_{acid} = 337.0 ± 2.1 kcal mol^{-1}); the opposite reaction occurs (**Table 3.6**). Deprotonated 3-formyl pyridone does deprotonate malononitrile but deprotonated malononitrile does not deprotonate 3-formyl pyridone. We thus bracket 3-formyl pyridone to be ΔH_{acid} = 336 ± 3 kcal mol^{-1}.

Table 3.6. Summary of results for acidity bracketing of 3-formyl-2-pyridone (**9**).

Reference compound	ΔH_{acid}[a]	Proton transfer[b]	
		Ref. acid	Conj. base
methyl cyanoacetate	340.8 ± 0.6	−	+
trifluoro-*m*-cresol	339.3 ± 2.1	−	+
2-chloropropionic acid	337.0 ± 2.1	−	+
malononitrile	335.8 ± 2.1	+	−
pyruvic acid	333.5 ± 2.9	+	−
difluoroacetic acid	331.0 ± 2.2	+	−
1,1,1-trifluoro-2,4-pentanedione	328.3 ± 2.9	+	−

[a] ΔH_{acid} is in kcal mol^{-1}.[33] [b] A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

We also measured the acidity of 3-formyl-2-pyridone using the Cooks kinetic method. Five reference acids were used: 2-chloropropionic acid (ΔH_{acid} = 337.0 ± 2.1 kcal mol^{-1}), 2-bromopropionic acid (ΔH_{acid} = 336.8 ± 2.1 kcal mol^{-1}), *p*-hydroxybenzoic acid (ΔH_{acid} = 335.9 ± 2.1 kcal mol^{-1}), 2-chlorobenzoic acid (ΔH_{acid} = 335.1 ± 2.1 kcal mol^{-1}), pyruvic acid (ΔH_{acid} = 333.5 ± 2.9 kcal mol^{-1}). The experiments yield an acidity of 335 ± 3 kcal mol^{-1}.

3.3.3.3. Experiments: 3-formyl-2-pyridone PA

When bracketing the PA of 3-formyl-2-pyridone, we find that the reaction proceeds in both directions for 3-methylpyrazole (PA = 216.5 ± 2.0 kcal mol^{-1}) and *N,N*-dimethylacetamide, placing the PA = 217 ± 3 kcal mol^{-1} (**Table 3.7**).

Table 3.7. Summary of results for proton affinity bracketing of 3-formyl-2-pyridone (**9**).

Reference compound	PA[a]	Proton transfer[b]	
		Ref. base	Conj. acid
n-butylamine	220.2 ± 2.0	+	−
N-methylaniline	219.1 ± 2.0	+	−
N,N-dimethylacetamide	217.0 ± 2.0	+	+
3-methylpyrazole	216.5 ± 2.0	+	+
2-chloropyridine	215.3 ± 2.0	−	+
o-toluidine	212.9 ± 2.0	−	+
pyrimidine	211.7 ± 2.0	−	+
aniline	210.9 ± 2.0	−	+
pyrrole	209.2 ± 2.0	−	+
m-chloroaniline	207.5 ± 2.0	−	+

[a] PA is in kcal mol^{-1}.[33] [b] A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

Six reference bases were used to measure PA via Cooks kinetic method: *n*-butylamine (PA = 220.2 ± 2.0 kcal mol^{-1}), *N*-methylaniline (PA = 219.1 ± 2.0 kcal mol^{-1}), *N*-benzylamine (PA = 218.3 ± 2.0 kcal mol^{-1}), *N,N*-dimethylacetamide (PA = 217.0 ± 2.0

kcal mol^{-1}) 3-methylpyrazole (PA = 216.5 ± 2.0 kcal mol^{-1}), 2-chloropyridine (PA = 215.3 ± 2.0 kcal mol^{-1}). These yield a PA of 217 ± 3 kcal mol^{-1}.

3.3.3.4. Tautomer composition: 3-formyl-2-pyridone.

The measured acidity is 335-336 kcal mol^{-1}, with a ±3 kcal mol^{-1} error bar. Because this value is right "in between" and could correspond to any of the various acidities of the three low-energy structures (enol **9b'** has a calculated ΔH_{acid} of 340.6 kcal mol^{-1}; the other enol **9b** and the ketone **9a**, both around 333 kcal mol^{-1}), the acidity cannot be used to discriminate among the possible tautomers.

The measured proton affinity of 217 kcal mol^{-1} (by both bracketing and Cooks) corresponds to the calculated PA for the most stable enol tautomer **9b'**: in this case, the most basic site and the most stable tautomer are consistent. Again, the other enol **9b** could also be present (as could, to a lesser extent, the least stable keto **9a**) as a mixture, but we can conclude that we do have enol **9b'** present, whether bracketing or Cooks conditions are used.

3.3.4. S$_N$2 studies

As stated earlier, the initial motivation for this study was to examine the correlation between acidity and leaving group ability for resonance-stabilized versus non-resonance-stabilized anionic leaving groups. However, characterization of the model system - substituted pyridones - is of interest in its own right, as described in much of this Chapter. In this section, we wish to briefly report computational results comparing the acidity (ΔH_{acid}) and S$_N$2 barrier ($\Delta H^{\ddagger}$ for the S$_N$2 reaction using formate as a nucleophile) for a

series of 3-substituted pyridones (**Figure 3.3**, **Scheme 3.2**, **Table 3.8**). The parent 2-pyridone is X=H. The moieties C_2H_3, C_4H_5, and HCO were chosen as groups that can delocalize by resonance an anion at N1. Groups that do not allow for resonance delocalization stabilization are X= Cl and Br.[77] In Table 3.8, "$\Delta\Delta H_{acid}$" represents the difference in ΔH_{acid} between the parent 2-pyridone and a given substituted pyridone. "$\Delta\Delta H^{\ddagger}$" represents the difference in $\Delta H^{\ddagger}$ between the parent 2-pyridone and a given substituted pyridone. The ratio of $\Delta\Delta H^{\ddagger}/\Delta\Delta H_{acid}$ (last column of **Table 3.8**) indicates the relationship between the effect an X group has on acidity versus the effect of that same X group on the S_N2 enthalpic barrier. The "better" the correlation between acidity and the S_N2 barrier, the closer to 1 this value should be. We hypothesize that groups that stabilize the N1-anion by resonance delocalization will have a weaker correlation (smaller value) because that delocalization will enhance acidity more than it will lower the S_N2 barrier. The argument is that in the S_N2 transition state, the N1-anion is not fully formed so the full benefit of the resonance delocalization is not realized.[77] The trends in **Table 3.8** do appear to support the hypothesis: the resonance-delocalized groups C_2H_3 and C_4H_5 have a smaller $\Delta\Delta H^{\ddagger}/\Delta\Delta H_{acid}$ value (0.575 and 0.549) than do the halide substituents (0.654 and 0.620 for Cl and Br, respectively). The formyl group is an interesting data point as its correlation is quite high (0.642) for a resonance-delocalized group. We speculate that HCO may not be a good model since the oxygen is inductively electron withdrawing, making the HCO not strictly a resonance delocalization moiety. Hammett s values support this theory: the s_m values, which reflect inductive ability, are similar for Br, Cl and HCO (0.39, 0.37, and 0.35, respectively).[186-188] In contrast, the s_m value for C_2H_3 is very small: 0.05. The comparison of C_2H_3 and HCO is particularly

useful, as the groups differ by the "exchange" of a CH2 for an O ($H_2C=CH_2$ versus $H_2C=O$). Both provide resonance stabilization through the double bond, but HCO is also stabilizing via induction, which means it is not a strictly resonance-stabilizing group.

Scheme 3.2.
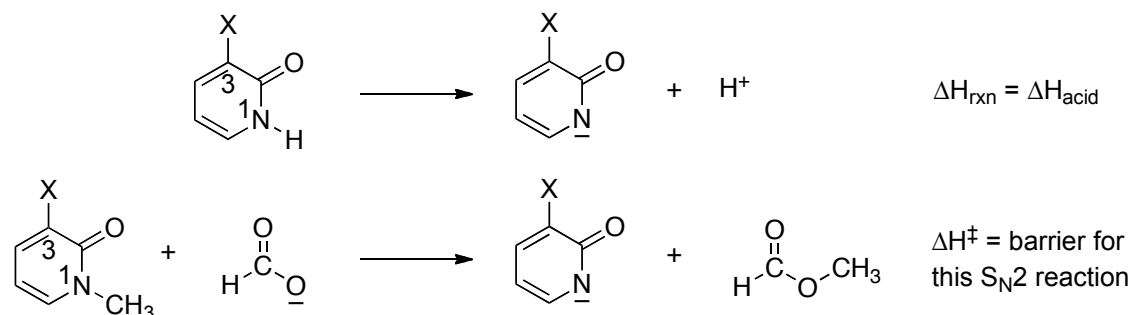
Table 3.8. M06-2X/6-311+G(2df,2p) calculations of acidity and S_N2 barrier for a series of 3-substituted-2-pyridonones (structure shown in **Figure 3.3**, R=H).

X	ΔH_{acid}, kcal mol^{-1} ^{a}	$\Delta H^{\ddagger}$, kcal mol^{-1} ^{b}	$\Delta\Delta H^{\ddagger}/\Delta\Delta H_{acid}$
H	346.3 (0)	18.4	–
C_2H_3	342.3 (4.0)	16.1(2.3)	0.575
C_4H_5	339.2 (7.1)	14.5(3.9)	0.549
HCO	332.6(13.7)	9.6 (8.8)	0.642
Cl	338.2 (8.1)	13.1(5.3)	0.654
Br	337.1(9.2)	12.7(5.7)	0.620

^{a} The values in parentheses are the differences in ΔH_{acid} ($\Delta\Delta H_{acid}$) for the various X substituents, relative to the parent pyridone (X=H) [ΔH_{acid}(parent) – ΔH_{acid}(substituted)].

^{b} The values in parentheses are the differences in $\Delta H^{\ddagger}$ ($\Delta\Delta H^{\ddagger}$) for the various X substituents, relative to the parent pyridone (X=H) [$\Delta H^{\ddagger}$(parent) – $\Delta H^{\ddagger}$(substituted)].

3.4. Conclusions.

In summary, we have characterized the acidity and proton affinities of 2-pyridone, 3-chloro-2-pyridone, and 3-formyl-2-pyridone (**Figure 3.12**). For 2-pyridone, we find that gas phase calculations at M06-2X/6-311+G(2df,2p) correctly indicate that the keto and enol forms (**1a** and **1b**) are close in energy, with the enol being slightly more stable. Comparison of calculated and measured PAs indicate that the keto form is present. Most likely the more stable enol form is also present: we do not bracket its PA as it is less basic. Interestingly, measurement of the PA using the Cooks kinetic method, which vaporizes the pyridone from aqueous solution, indicates the keto tautomer. This is consistent with the solution phase preference for the keto structure.

For the 3-chloro-2-pyridone, which has not heretofore been studied, calculations indicate that the keto **6a** and enol **6b** are close in energy, with the enol being slightly more stable (much like the parent pyridone). The PA measurements point to the keto structure, again because the bracketing experiment targets the more basic tautomer. The more stable enol tautomer is probably also present, in a mixture of keto and enol. The alternate enol structure **6b'** is somewhat stabilized by a weak internal hydrogen bond between the Cl and H, but is still the least stable structure and if present, will be a small component in the mixture. As with the parent 2-pyridone, the Cooks kinetic experiment indicates the keto tautomer, which is probably more stable in solution.

The 3-formyl derivative, which also has not been studied, has an enol conformation with an internal hydrogen bond (between the aldehyde O and the enol H) that renders **9b'**

as the most stable tautomer. The PA measurements confirm the presence of this enol form (both with bracketing and the Cooks kinetic method). This particular derivative is interesting as the presence of the formyl group reverses the relative stability of the two enol tautomers (compared to the parent and 3-chloro compounds). Different substitution can therefore allow one to "tune" for tautomer preference.
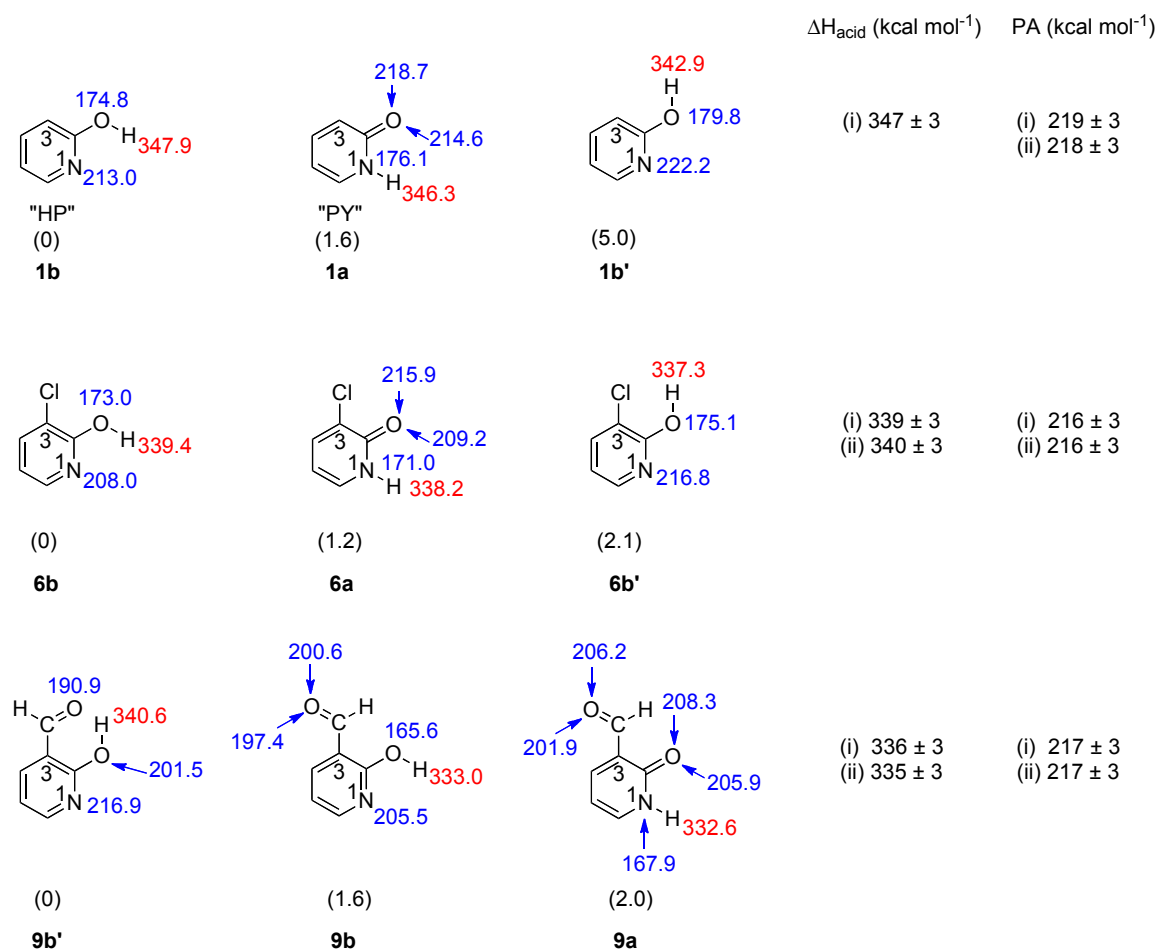
Figure 3.12. Summary of gas-phase computational (M06-2X/6-311+G(2df,2p)) and experimental data for the pyridones studied herein. Calculated relative stabilities are in parentheses; values in blue are calculated PAs and values in red are calculated acidities.

All are ΔH_{298} values. For the experimental data, (i) indicates use of the bracketing method; (ii) indicates Cooks kinetic method measurement.

In terms of the substituted pyridones as a model system for testing acidity-leaving group correlations, our calculations indicate that leaving groups that allow for resonance delocalization of the product anion (pyridones substituted with $X = C_2H_3$, C_4H_5 (**Figure 3.3**)) do show less correlation than non-resonance-stabilizing groups ($X = Cl$, Br). That is, anions that are stabilized by resonance may be stable conjugate bases (thus their conjugate acids are acidic), but may not be correspondingly good leaving groups since that stabilization is not fully felt in the S_N2 transition state.

Chapter 4. Gas Phase Studies of 3- and 7-Methylsubstituted Purine 3-Methyladenine DNA Glycosylase II (AlkA) Substrates

4.1. Introduction

Maintaining the integrity of DNA is essential for the health of living organisms. Unfortunately, DNA is constantly under assault; one of the most common modifications is alkylation, both by cellular metabolites as well as exogenous alkylating agents. Alkylation damage threatens proper cell function and compromises the correct propagation of the genetic code.[5,8] The base excision repair (BER) pathway is the primary means for excising damaged bases. In *Escherichia coli*, 3-methyladenine DNA glycosylase II (called AlkA, after the gene that encodes for it) is an enzyme that is up-regulated following exposure to DNA alkylating agents.[189-191] Because of its ability to cleave a wide range of substrates, AlkA is considered a particularly intriguing enzyme.[5,8] The other alkylation-specific enzyme found in *E. coli*, 3-methyladenine DNA glycosylase I (TAG), is quite specific, catalyzing the excision of only 3-alkyl substituted adenine and guanine (but not other alkylated nucleobases).[192,193] In contrast, AlkA has a very broad substrate range, catalyzing the excision of various N3 and N7-alkyl purines, O2-alkyl pyrimidines, and other lesions that are not the product of alkylation, such as hypoxanthine, xanthine, and 1,N^6-ethenoadenine.[5,194-197]

Because AlkA cleaves such a diverse set of damaged bases, the active site is thought to be indiscriminate, with the reactivity of the *N*-glycosidic bond of a given substrate dictating the rate of AlkA-catalyzed excision.[20] Excision is believed to occur via an S_N1-type mechanism, where the nucleobase leaves first (**Scheme 4.1**).[5]

Scheme 4.1.
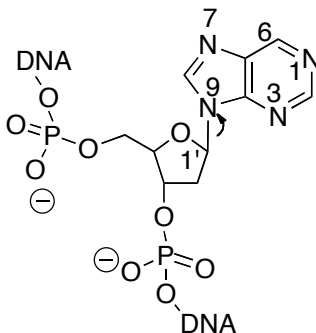
In prior work, we hypothesized that a related enzyme, alkyl adenine glycosylase (AAG), which catalyzes the excision of a wide range of damaged bases in mammalian cells, may provide a hydrophobic active site which aids in the discrimination of damaged from normal bases by enhancing the differences in their leaving group ability.[36,40,41,76,129] We hypothesize that AlkA may do the same.

The examination of properties in the gas phase, which provides the "ultimate" nonpolar environment, reveals intrinsic reactivity that can be correlated to activity in other media, such as hydrophobic active sites.[10,36,38,40,41,76,198] In this Chapter, we calculate and measure the gas phase acidities and proton affinities of a series of alkylated purine substrates (most of which have not been heretofore studied in vacuo), and compare the results to acities and proton affinities of other AlkA substrates in order to discuss the results in the context of the AlkA mechanism.

4.2. Experimental

All the purine substrates and reference compounds are commercially available and were used as received.

4.2.1. Bracketing method

Acidity and proton affinity bracketing measurements were conducted using a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FTMS) with a dual cell setup, which has been described previously, and in **Chapter 1** in detail.[10,13,36,38-41]

The typical protocol for bracketing experiments has been described previously and in **Chapter 1**.[10,13,36,38-41] Proton transfer reactions are conducted in both directions. The occurrence of proton transfer is regarded as *prima facie* evidence that the reaction is exothermic ("+" in the Tables).

4.2.2. Cooks kinetic method

We also used the Cooks kinetic method in a quadrupole ion trap (LCQ) mass spectrometer[84-87] to measure the acidities and proton affinities of substituted purines. The theory and methodology of Cooks kinetics method described in details in previous Chapters (**Chapter 1**, **Chapter 3**).

The proton-bound complex ions are generated by electrospray (ESI) of 100–500 μM solutions of purine and reference acid (or base, for PA measurement). Methanol or water–methanol (20%) solution are used as a solvent.[72] An electrospray needle voltage of ~4 kV and the flow rate of 25 μL/min is applied. The proton-bound complex ions are isolated and then dissociated by applying collision-induced dissociation (CID); the complexes are activated for about 30 ms. Finally, the dissociation product ions are

detected to give the ratio of the deprotonated (or protonated) analyte and deprotonated (or protonated) reference acid. A total of 40 scans are averaged for the product ions.

4.2.3. Calculations.

Calculations are conducted at B3LYP/6-31+G(d)[92-94] levels using Gaussian03[90] and Gaussian09[91]; the geometries are fully optimized and frequencies are calculated. No scaling factor is applied. All the values reported are ΔH at 298 K. Dielectric medium calculations were done using the conductor-like polarizable continuum solvent model (CPCM, full optimization; UAKS cavity) at B3LYP/6-31+G(d) as implemented in Gaussian03.[105,106,107] The "total free energy in solution" values are reported, and the solvation free energy of a proton in water or DMSO (-265.9 or -273.3 kcal mol^{-1} respectively) is accounted for.[199]

4.3. Results

4.3.1. 7-Methyladenine (7meA, 1)

4.3.1.1. Calculations: 7-methyladenine tautomers, acidity, proton affinity.

In our experience DFT methods generally yield accurate values for thermochemical properties of nucleobases, so we utilized B3LYP/6-31+G(d) to calculate the relative tautomeric stabilities, acidities (ΔH_{acid}), and proton affinities (PA) of 7-methyladenine.[10,13,38,39,41] 7MeA has five possible tautomeric structures (**Figure 4.1**). The most stable tautomer (amino 7MeA **1a**) is over 8 kcal mol^{-1} more stable than the next most stable species. The most acidic site of **1a** is predicted to be the exocyclic NH_2

(ΔH_{acid} = 342.2 kcal mol^{-1}). The most basic site of tautomer **1a** is the N3 (PA = 234.7 kcal mol^{-1}).
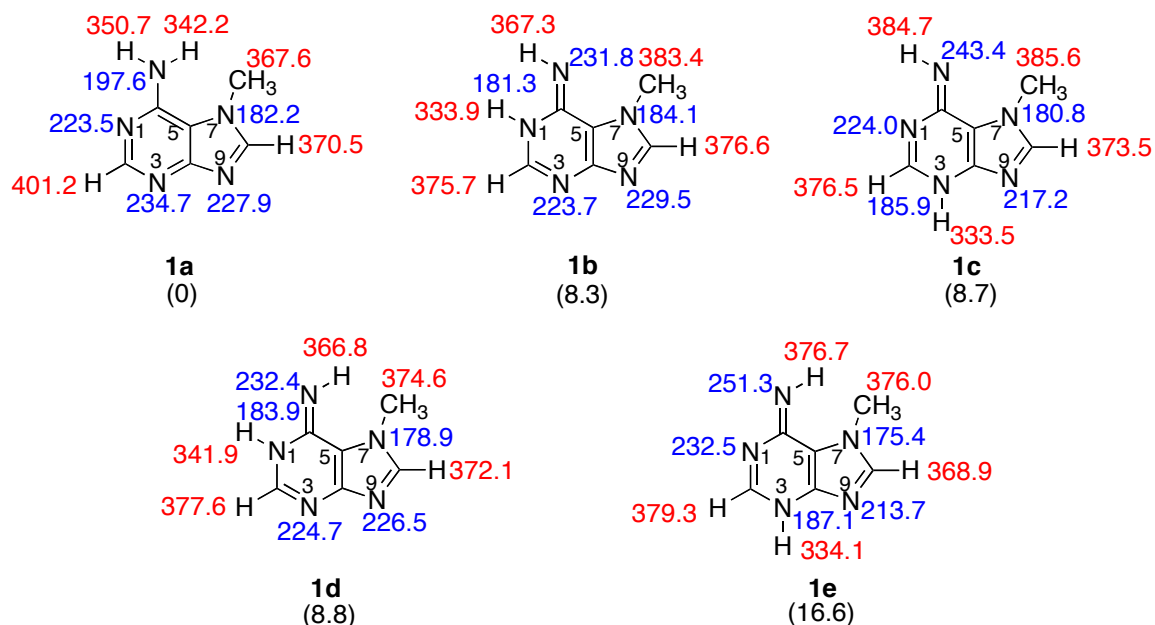
Figure 4.1. The five possible tautomeric structures of 7-methyladenine. Gas phase acidities are in red; gas phase proton affinities are in blue. Relative stabilities are in parentheses. Calculations were conducted at B3LYP/6-31+G(d); reported values are ΔH at 298 K.

4.3.1.2. Experiments: 7-methyladenine acidity.

We measured the acidity of 7-methyladenine using acidity bracketing. The conjugate base of 7-methyladenine deprotonates 2,4-pentanedione; the reaction in the opposite direction (the conjugate base of 2,4-pentanedione with 7-methyladenine) also occurs (**Table 4.1**). We therefore bracket the ΔH_{acid} of 7meA as 344 $\pm$ 3 kcal mol^{-1}.

Table 4.1. Summary of results for acidity bracketing of 7-methyladenine (**1**).

Reference compound	ΔH_{acid}[a]	Proton transfer[b]	
		Ref. acid	Conj. base
4-(trifluoromethyl)aniline	353.3 ± 2.1	−	+
m-cresol	349.6 ± 2.1	−	+
acetic acid	347.4 ± 0.5	−	+
butyric acid	346.8 ± 2.0	−	+
2,4-pentanedione	343.8 ± 2.1	+	+
methylcyanoacetate	340.8 ± 0.6	+	−
α,α,α-trifluoro-*m*-cresol	339.2 ± 2.1	+	−
2-chloropropanoic acid	337.0 ± 2.1	+	−

[a] ΔH_{acid} is in kcal mol^{-1}.[33,182] [b] A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

We also measured the acidity of 7meA using the Cooks kinetic method. Seven reference acids were used: butyric acid (ΔH_{acid} = 346.8 ± 2.0 kcal mol^{-1}), valeric acid (ΔH_{acid} = 346.2 ± 2.1 kcal mol^{-1}), *iso*-valeric acid (ΔH_{acid} = 345.5 ± 2.1 kcal mol^{-1}), pivalic acid (ΔH_{acid} = 344.6 ± 2.1 kcal mol^{-1}), methacrylic acid (ΔH_{acid} = 344.1 ± 2.9 kcal mol^{-1}), 4-aminobenzoic acid (ΔH_{acid} = 343.4 ± 2.1 kcal mol^{-1}), and methoxyacetic acid (ΔH_{acid} = 341.9 ± 2.1 kcal mol^{-1}), yielding a ΔH_{acid} of 344 ± 3 kcal mol^{-1}.

4.3.1.3. Experiments: 7-methyladenine (7meA) proton affinity.

In bracketing the PA of 7meA, we find that di-*sec*-butylamine (PA = 234.4 ± 2.0 kcal mol^{-1}) deprotonates protonated 7-methyladenine; the opposite reaction (7-methyladenine deprotonating protonated di-*sec*-butylamine) also occurs (**Table 4.2**). We therefore bracket the PA of 7meA to be 234 ± 3 kcal mol^{-1}.

Table 4.2. Summary of results for proton affinity bracketing of 7-methyladenine (**1**).

Reference compound	PA^a	Proton transfer[b]	
		Ref. base	Conj. acid
2,2,6,6-tetramethylpiperidine	235.9 ± 2.0	+	−
N,N-dimethylcyclohexylamine	235.1 ± 2.0	+	−
triethylamine	234.7 ± 2.0	+	−
di-sec-butylamine	234.4 ± 2.0	+	+
1-methylpiperidine	232.1 ± 2.0	−	+
1-methylpyrrolidine	230.8 ± 2.0	−	+
pyrrolidine	226.6 ± 2.0	−	+

[a]PA is in kcal mol^{-1}.[33] [b]A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

Using the Cooks kinetic method with five reference bases (2,2,6,6-tetramethylpiperidine (PA = 235.9 ± 2.0 kcal mol^{-1}), N,N-dimethylcyclohexylamine (PA = 235.1 ± 2.0 kcal mol^{-1}), triethylamine (PA = 234.7 ± 2.0 kcal mol^{-1}), 1-methylpiperidine (PA = 232.1 ± 2.0 kcal mol^{-1}), and N,N-dimethylbenzylamine (PA = 231.5 ± 2.0 kcal mol^{-1})), we measure a PA of 234 ± 3 kcal mol^{-1}.

4.3.2. 7-Methylguanine (7meG, 2)

4.3.2.1. Calculations: 7-methylguanine tautomers, acidity, proton affinity.

There are ten possible tautomers for 7-methylguanine (**2**) (**Figure 4.2**); the six lowest (all below 15 kcal mol^{-1} relative to the most stable tautomer) are shown in **Figure 4.3**. The most stable form is the keto-amino **2a**; for this tautomer, the most acidic site is the N1-H (ΔH_{acid} = 335.6 kcal mol^{-1}) and the most basic site is the N9 (PA = 231.4 kcal mol^{-1}).
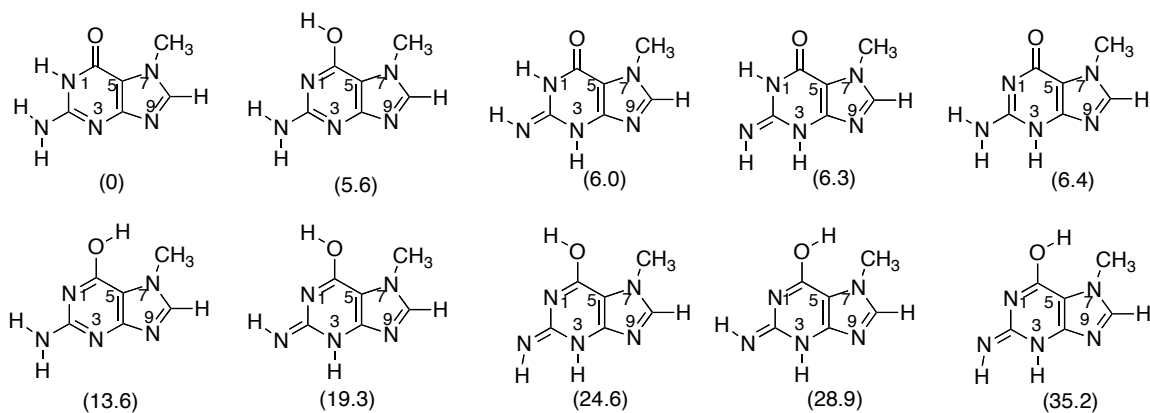
Figure 4.2. Possible tautomeric structures of 7-methylguanine, including higher energy tautomers. Relative stabilities (ΔH at 298 K) are listed in parentheses. Calculations were conducted at B3LYP/6-31+G(d).

Figure 4.3. The six low energy tautomeric structures of 7-methylguanine. Gas phase acidities are in red; gas phase proton affinities are in blue. Relative stabilities are in parentheses. Calculations were conducted at B3LYP/6-31+G(d); reported values are ΔH at 298 K.

4.3.2.2. Experiments: 7-methylguanine acidity.

We measured the acidity of 7-methylguanine using the bracketing method. The reaction of 2-chloropropionic acid (ΔH_{acid} = 337.0 ± 2.1 kcal mol^{-1}) and deprotonated 7meG proceeds, as does the reaction in the opposite direction (2-chloropropionate with 7meG), allowing us to bracket the ΔH_{acid} to be 337 ± 3 kcal mol^{-1} (**Table 4.3**).

Table 4.3. Summary of results for acidity bracketing of 7-methylguanine (**2**).

Reference compound	$\Delta H_{acid}{}^{a}$	Proton transfer[b]	
		Ref. acid	Conj. base
2,4-pentanedione	343.8 ± 2.1	−	+
methylcyanoacetate	340.8 ± 0.6	−	+
α,α,α-trifluoro-*m*-cresol	339.2 ± 2.1	−	+
2-chloropropionic acid	337.0 ± 2.1	+	+
malononitrile	335.8 ± 2.1	+	−
pyruvic acid	335.5 ± 2.9	+	−
difluoroacetic acid	331.0 ± 2.2	+	−

[a]ΔH_{acid} is in kcal mol^{-1}.[33] [b]A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

Five reference acids were used in the Cooks kinetic method measurement of 7meG acidity: 3-chloropropionic acid (ΔH_{acid} = 340.8 ± 2.7 kcal mol^{-1}), 2-chloropropionic acid (ΔH_{acid} = 337.0 ± 2.1 kcal mol^{-1}), 4-hydroxybenzoic acid (ΔH_{acid} = 335.9 ± 2.1 kcal mol^{-1}), 2-chlorobenzoic acid (ΔH_{acid} = 335.1 ± 2.1 kcal mol^{-1}), and pyruvic acid (ΔH_{acid} = 333.5 ± 2.9 kcal mol^{-1}). The ΔH_{acid} was found to be 337 ± 3 kcal mol^{-1}.

4.3.2.3. Experiments: 7-methylguanine proton affinity.

We also bracketed the PA of 7meG (**Table 4.4**). Protonated 1-methylpyrrolidine reacts with 7meG; likewise, protonated 7meG reacts with 1-methylpyrrolidine, placing the PA at 231 ± 3 kcal mol^{-1}.

Table 4.4. Summary of results for proton affinity bracketing of 7-methylguanine (**2**).

Reference compound	PA^a	Proton transfer^b	
		Ref. base	Conj. acid
2,2,6,6-tetramethylpiperidine	235.9 ± 2.0	+	−
di-*sec*-butylamine	234.4 ± 2.0	+	−
1-methylpiperidine	232.1 ± 2.0	+	−
1-methylpyrrolidine	230.8 ± 2.0	+	+
piperidine	228.0 ± 2.0	−	+
4-picoline	226.4 ± 2.0	−	+
3-picoline	225.5 ± 2.0	−	+

^{a}PA is in kcal mol^{-1}.[33] ^{b}A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

For the Cooks PA measurement, six reference bases were used: triethylamine (PA = 234.7 ± 2.0 kcal mol^{-1}), di-sec-butylamine (PA = 234.4 ± 2.0 kcal mol^{-1}), 1-methylpiperidine (PA = 232.1 ± 2.0 kcal mol^{-1}), *N,N*-dimethyl-iso-propylamine (PA = 232.0 ± 2.0 kcal mol^{-1}), *N,N*-dimethylbenzylamine (PA = 231.5 ± 2.0 kcal mol^{-1}), and *N*-methylpiperidine (PA = 230.8 ± 2.0 kcal mol^{-1}), yielding a PA of 232 ± 3 kcal mol^{-1}.

4.3.3. 3-Methyladenine (3meA, 3)

4.3.3.1. Calculations: 3-methyladenine tautomers, acidity, proton affinity.

The acidity and the relative stabilities of the possible tautomers of 3-methyladenine have been calculated by our group previously; these data plus new calculations of PA are shown in **Figure 4.4.**[36] 3-Methyladenine has five possible tautomers; the three lowest are shown. The most stable is the one with the exocyclic amino group (**3a**), for which the calculated acidity is 346.8 kcal mol^{-1} (for the proton on the amino group). The most basic site has a PA of 234.5 kcal mol^{-1}, at the N7.
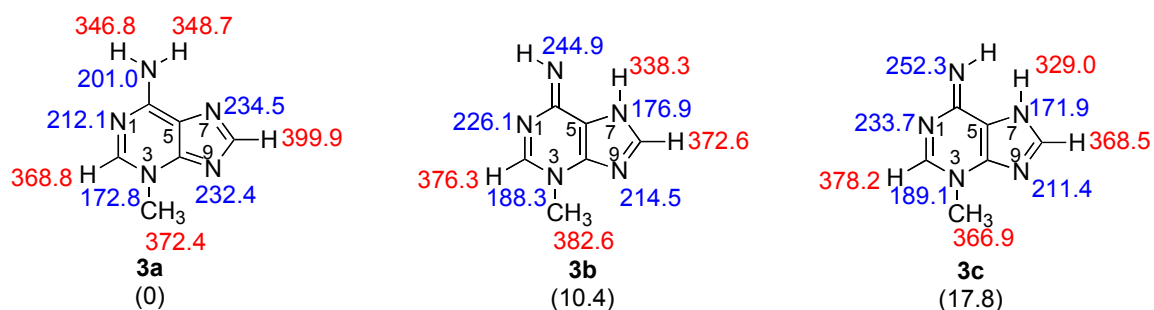
Figure 4.4. Tautomeric structures of 3-methyladenine. Gas phase acidities are in red; gas phase proton affinities are in blue. Relative stabilities are in parentheses. Calculations were conducted at B3LYP/6-31+G(d); reported values are ΔH at 298 K.

4.3.3.2. Experiments: 3-methyladenine acidity and proton affinity.

The acidity of 3-methyladenine was previously measured to be 347 ± 4 kcal mol^{-1}.[36] The PA bracketing results for 3meA are shown in **Table 4.5**. Di-*sec*-butylamine (PA = 234.4 ± 2.0 kcal mol^{-1}) can deprotonate protonated 3-methyladenine, but 1-methylpiperidine (232.1 ± 2.0 kcal mol^{-1}) cannot. In the reverse direction, 3-

methyladenine deprotonates protonated 1-methylpiperidine, but not protonated di-*sec*-butylamine. We therefore bracket the PA of 3-methyladenine to be 233 ± 3 kcal mol^{-1}.

Table 4.5. Summary of results for proton affinity bracketing of 3-methyladenine (**3**).

Reference compound	PA^a	Proton transfer^b	
		Ref. base	*Conj. acid*
2,2,6,6-tetramethylpiperidine	235.9 ± 2.0	+	−
trimethylamine	234.7 ± 2.0	+	−
di-*sec*-butylamine	234.4 ± 2.0	+	−
1-methylpiperidine	232.1 ± 2.0	−	+
2,4-lutidine	230.1 ± 2.0	−	+
3-picoline	225.5 ± 2.0	−	+

^{a}PA is in kcal mol^{-1}.[33] ^{b}A "+" indicates the occurrence and a "−" indicates the absence of proton transfer

4.3.4. 3-Methylguanine (3MeG, 4). Calculations: 3-methylguanine tautomers, acidity, proton affinity.

3-Methylguanine has fourteen possible tautomers (**Figure 4.5**); the seven structures within 15 kcal mol^{-1} of the most stable form are shown, along with calculated acidities and proton affinities, in **Figure 4.6.** The most acidic site of the most stable tautomer **4a** is the N7-H, with a calculated ΔH_{acid} of 328.6 kcal mol^{-1}. The most basic site is on the imino NH (PA = 231.8 kcal mol^{-1}).[200]
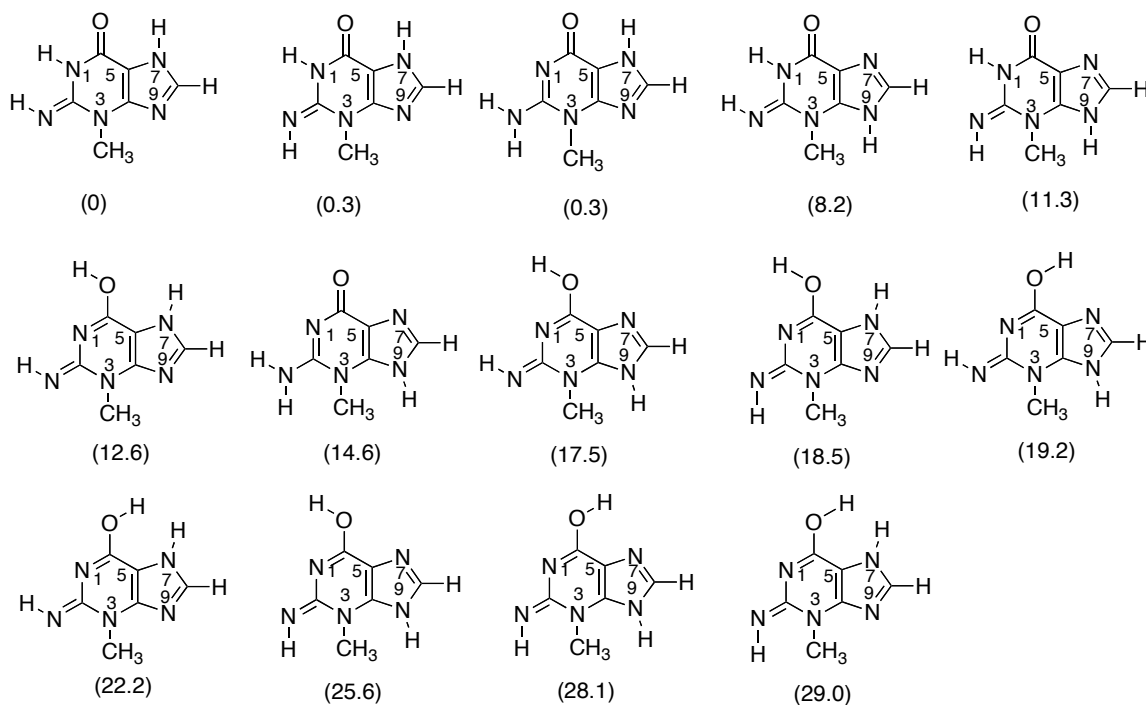
Figure 4.5. All possible tautomeric structures of 3-methylguanine. Relative stabilities (ΔH at 298 K; B3LYP/6-31+G(d)) are listed in parentheses.
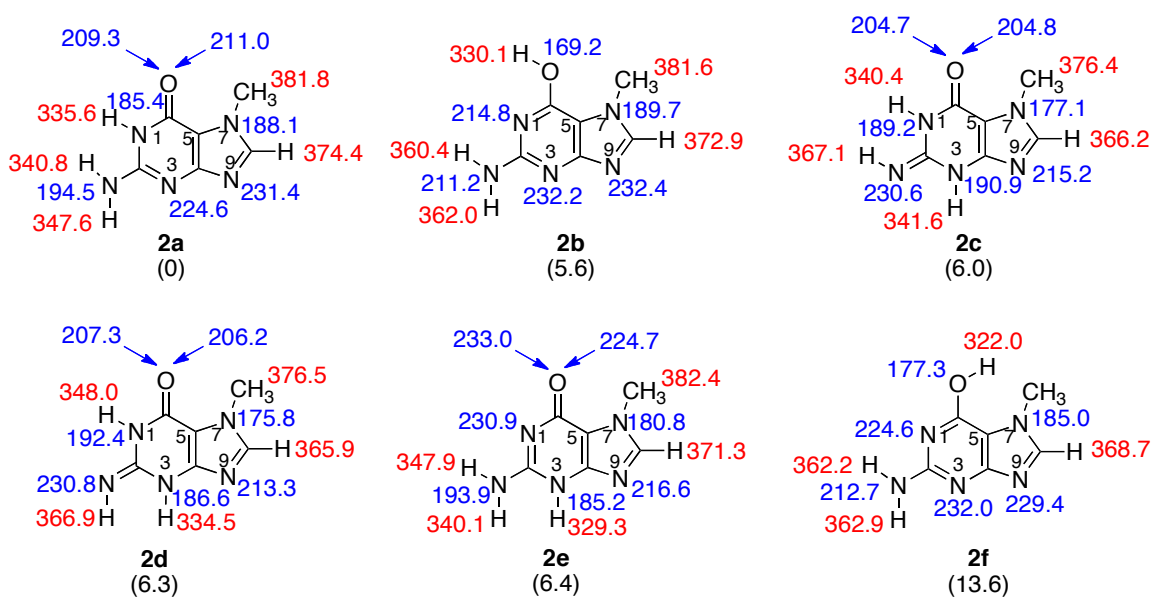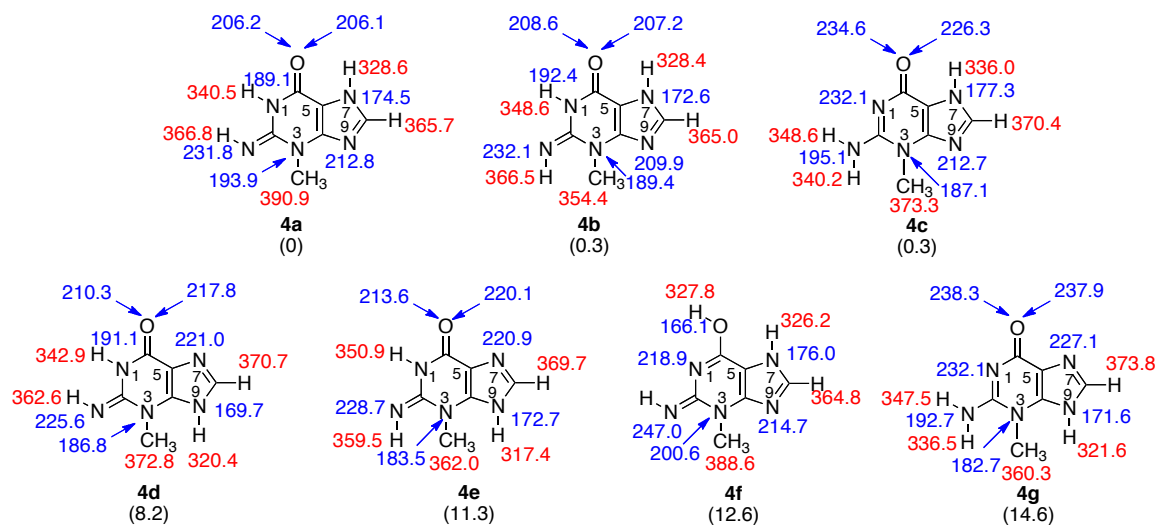

Figure 4.6. Lower energy tautomeric structures of 3-methylguanine. Gas phase acidities are in red; gas phase proton affinities are in blue. Relative stabilities are in

parentheses. Calculations were conducted at B3LYP/6-31+G(d); reported values are ΔH at 298 K.

4.4. Discussion.

4.4.1. Calculated versus experimental values.

The calculated acidity and proton affinity values for all the substrates studied herein are summarized in **Table 4.6**. Generally, B3LYP/6-31+G(d) appears to provide fairly accurate predictions for the thermochemical values, and it demonstrated a good accuracy for 3- and 7-methylsubstituted purine calculations here too.[201,202]

Table 4.6. Calculated (B3LYP/6-31+G(d); 298 K) and experimental data for damaged bases.

Substrate	Calculated value	Experimental value[b]
$\Delta H_{acid}{}^{a}$		
7-methyladenine (**1**)	342.2	344 (344)
7-methylguanine (**2**)	335.6	337 (337)
3-methyladenine (**3**)	346.8	347
3-methylguanine (**4**)	328.6	N/A
PA^{a}		
7-methyladenine (**1**)	234.7	234 (234)
7-methylguanine (**2**)	231.4	231 (232)
3-methyladenine (**3**)	234.5	233
3-methylguanine (**4**)	231.8	N/A

[a]ΔH_{acid} and PA values are in kcal mol^{-1}; [b]First listed experimental value is bracketed; Cooks kinetic method value, if available, is in parentheses. Error is ±3-4 kcal mol^{-1}.

4.4.2. Biological implications.

AlkA is a glycosylase with a particularly broad substrate range, cleaving a wide variety of damaged bases from double-stranded DNA.[4,19-21,23,196,197,203-209] The exact mechanism

by which AlkA cleaves damaged bases with greater efficiency than the normal bases adenine and guanine is unknown. The main hypothesis is that cleavage is related to the intrinsic stability of the *N*-glycosidic bond and that the enzyme merely provides a non-base-specific active site.[5,8,20,205] Therefore, the better a leaving group of the nucleobase is, the more easily it is cleaved. Since acidity and leaving group ability are generally correlated, we would expect the damaged bases to be more acidic than the normal bases.

We further postulate, based on our previous studies of other glycosylases, that AlkA may provide a nonpolar active site that serves to *enhance* the differences in acidity between damaged and normal nucleobases, and in doing so, aids in the discrimination of normal from damaged bases.[10,36,40,41,129] Thus, not only do we expect the damaged bases to be more acidic than the normal bases, but those differences in acidity should be significantly greater in the gas phase than in aqueous solution.

We first sought to compare the acidities of damaged and normal substrates of AlkA to ascertain whether the damaged bases are more acidic. If rate of excision is based on the intrinsic stability of the *N*-glycosidic bond, then the acidity of the N9 position is relevant (**Figure 4.7**; note that the biologically relevant structure is not always the most stable structure in the gas phase). The substrates studied herein, as well as other known substrates of AlkA that we have previously studied (purine, hypoxanthine, 1,N^6-ethenoadenine, adenine, guanine) are shown.[36,38,40,41]
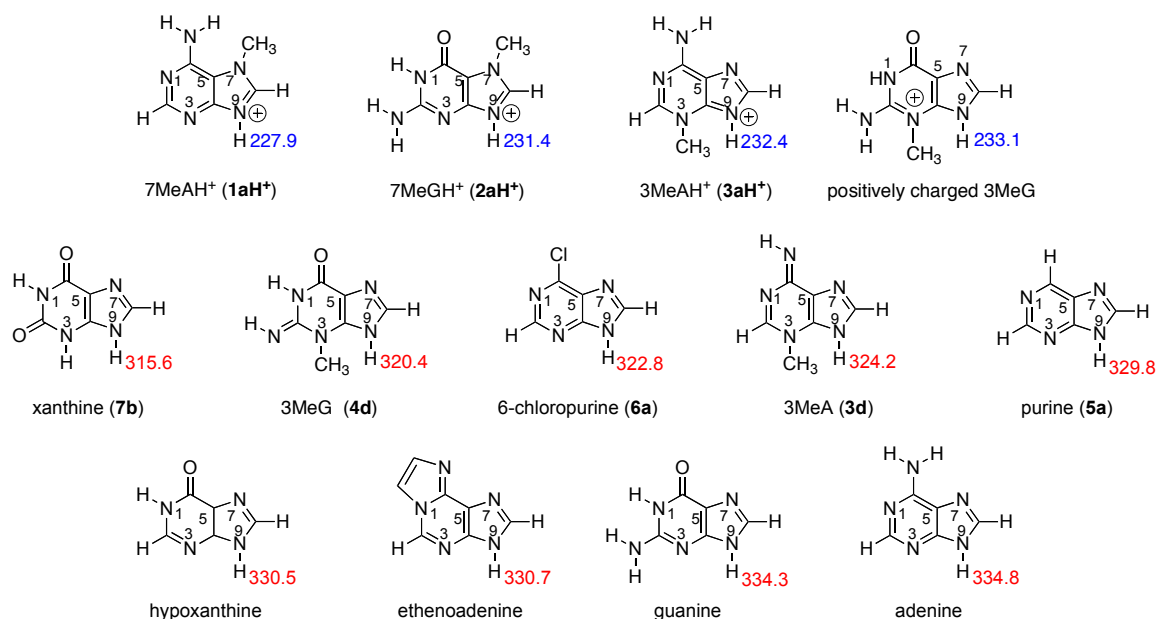
Figure 4.7. Gas phase acidity (ΔH_{298K}, calculated, B3LYP/6-31+G(d), in kcal mol^{-1}) of biologically relevant structures.[36,38,40,41]

One interesting feature of AlkA is that it cleaves both positively charged and neutral nucleobases. For example, when adenine and guanine are alkylated at N7 to form 7meA and 7meG, the result is a positively charged nucleobase (**1aH⁺** and **2aH⁺** in **Figure 4.7**). Cleavage of that positively charged nucleobase results in a neutral nucleobase leaving group (path A in **Scheme 4.2**, where **Scheme 4.2** shows possible cleavage paths for 3meA). Therefore, the relevant acidity to correlate to leaving group ability is the N9-H acidity value for the positively charged substrates, as shown for **1aH⁺**, **2aH⁺**, **3aH⁺**, and positively charged 3MeG in **Figure 4.7.** Those values are in blue because they are equivalent to the proton affinity values at those positions for the corresponding neutral substrates (**1a**, **2a**, **3a** and 3MeG).

Scheme 4.2.
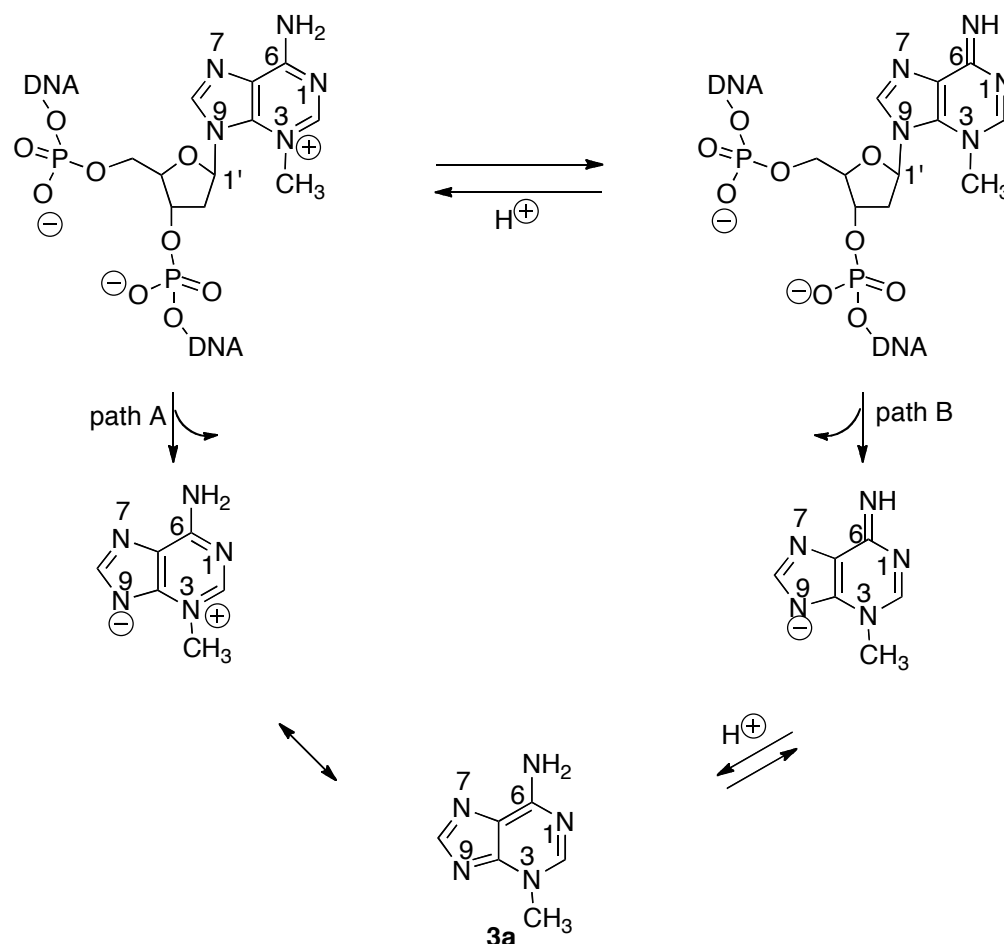
Other nucleobases, such as xanthine (**7**), 6-chloropurine (**6**), purine (**5**), hypoxanthine, ethenoadenine (eA), guanine and adenine (**Figure 4.5**) are neutral substrates for AlkA. In these cases, the leaving group ability would be related to the acidity of the neutral nucleobase at N9-H (values shown in red, **Figure 4.5**).[117,129,210]

3MeA (**3**) and 3meG (**4**) are intriguing because cleavage could either occur from the positively charged form (path A, **Scheme 4.2**) or from the neutral form (path B, **Scheme 4.2**). That is, cleavage of the nucleobase could occur via path A to yield the neutral nucleobase, or, a proton could be lost prior to cleavage, such that the deprotonated

nucleobase is the leaving group (path B). Therefore, these two substrates appear twice in **Figure 4.7**, in both positively charged and neutral forms.

The substrates in **Figure 4.7** are arranged in order of *decreasing* acidity. The substrate with the lowest acidity value (amino 7meAH$^+$ (**1aH$^+$**)) is expected to be the best leaving group. The question is, do these relative acidities correlate to known AlkA experimental excision rates? Known data for the rate constants by which AlkA cleaves various nucleobases are compiled in **Table 4.7**.[211] The nucleobases are listed in the order of *decreasing* rate constants. The data show a rough correlation between the rate constant for excision and the calculated ΔH_{acid} value. 7MeGH$^+$ is cleaved the most quickly, and the acidity value is the lowest (most acidic substrate). The next most easily cleaved nucleobase is 3meAH+, then 3meA. Thus, whether 3-methyladenine is excised as a neutral (path A, **Scheme 4.2**) or in the anionic deprotonated form (path B, **Scheme 4.2**), the acidity value correlates to the leaving group ability in a qualitative sense; that is, both acidity values are higher than that for 7meG, but less than that for eA (the next most easily cleaved nucleobase). For the next three nucleobases (ethenoadenine, purine, hypoxanthine), the gas phase acidities are similar (around 329-330 kcal mol^{-1}) and the rate constants for cleavage are in the same ballpark (10^{-2} min^{-1}).[211] For the normal nucleobases guanine and adenine, as expected for natural nucleobases, cleavage is slowest; the attendant acidities are also the highest in value (so least acidic).

Table 4.7. Rate constants for excision of various nucleobases by AlkA compared to gas phase acidity.

Substrate	k_{st} (min^{-1})^{a,b}	ΔH_{acid} (kcal mol^{-1})^{c}
7meGH$^+$ (**2H$^+$**)	300	231.4
3meAH$^+$(**3H$^+$**)/3meA (**3**)	0.5	232.4/324.2

ethenoadenine	7.5×10^{-2}	330.7^d
purine (**5**)	5.4×10^{-2}	329.8
hypoxanthine	2.9×10^{-2}	330.5^e
guanine	6.9×10^{-3}	334.3^f
adenine	5.2×10^{-3}	334.8^g

[a]Reference [20]; [b]k_{st} is single turnover rate constant with saturating AlkA; [c]Calculated ΔH_{acid} values at 298 K (B3LYP/6-31+G(d)); [d]Reference [40]; [e]Reference [41]; [f]Reference [38]; [g]References [36,38]

Our results showing correlation between excision rate constants and N9-H acidity values lend support to the argument that cleavage of the damaged base is dependent on the intrinsic stability of the *N*-glycosidic bond.

Based on the correlation we see herein, we would also predict that 6-chloropurine (as well as other halo-substituted purines) should (based on its high acidity) be easily cleaved by AlkA.

We also further postulate that the active site, by providing a hydrophobic environment, will aid in the discrimination of normal from damaged bases by *enhancing* the *relative* leaving group ability of the damaged bases. To examine that hypothesis, we compare the gas and solution phase N9-H acidities for those damaged bases whose pK_a values are known (**Table 4.8**). In an effort to "draw a line" from solution to the gas phase, we conducted dielectric medium calculations on the acidities in DMSO (ε=48) and water (ε=78) as well, to ascertain how acidities change with medium dielectric. The most acidic neutral substrate, regardless of medium, is the damaged base xanthine, which has a gas phase ΔH_{acid} of 315.6 kcal mol^{-1}, and a pK_a of 7.3. The least acidic substrate in the gas phase is the normal base adenine, with a gas phase acidity of 334.8. In the gas phase, xanthine is more acidic than adenine by 19 kcal mol^{-1}. In a dielectric of 48, that difference drops to 7.7; in a dielectric of water, it is even smaller (5.9 kcal mol^{-1}). When

adenine and xanthine are actually fully solvated in water (experimental pK_a column), that acidity difference is only 3.4 kcal mol^{-1}. Overall, in comparing the pK_as and ΔH_{acid} values of the damaged bases (xanthine, 6-chloropurine, purine, hypoxanthine, ethenoadenine) versus the normal bases (adenine and guanine), the same trend is seen: the difference in acidity of the damaged versus normal bases is greatest in the gas phase, and least in water. In fact, the solution phase pK_a values are so close that adenine is actually more acidic than its damaged counterpart, ethenoadenine. The intrinsically higher acidity of ethenoadenine is only evident in the gas phase values. The nonpolar active site in AlkA could thus contribute to specificity by enhancing the differences in acidity among adenine, guanine and damaged bases.

Table 4.8. Summary of calculated N9-H acidity values (in the gas phase (ε=1), DMSO (ε =48) and water (ε =78) and experimental pK_a values (in water) for AlkA substrates (structures shown in **Figure 4.7**).[a,b,c]

Substrate	Acidity (kcal mol^{-1})[a,b]			pK_a[c]
	ε=1	ε =48	ε =78	
7meGH$^+$(**2H$^+$**)	231.4	4.7	18.7	–
3meAH+(**3H$^+$**)/3meA (**3**)	232.4/324.2	36.3/–	20.1/6.6	–
xanthine (**7b**)	315.6	16.6	24.5	7.3[h]
6-chloropurine (**6a**)	322.8	18.6	26.3	7.7-7.8[i]
purine (**5a**)	329.8	21.0	28.2	8.9[j]
hypoxanthine	330.5[d]	21.7	28.2	8.9[k]
ethenoadenine	330.7[e]	22.7	29.7	9.9[e]
guanine	334.3[f]	24.7	30.2	10.0[l]
adenine	334.8[g]	24.3	30.4	9.8[m]

[a] Calculated N9-H acidities, in kcal mol^{-1}. [b]Calculated using B3LYP/6-31+G(d); [c]pK_a values in water are all experimental, except for xanthine, which is calculated; [c]Reference

[41]; [e]Reference [40]; [f]Reference [38]; [g]References [36,38]; [h]Reference [212]; [i]References [213,214]; [j]References [213-215]; [k]References [214,216]; [l]Reference [217]; [m]References [214,218]

4.5. Conclusions

We have calculated the tautomeric energies, and calculated and measured the acidic and basic properties for a 3- and 7-methylsubstituted purines not heretofore studied. The results indicate that the damaged purines are all more acidic than the normal nucleobases adenine and guanine, and would therefore be expected to be more easily cleaved (that is, their conjugate bases are better leaving groups). Furthermore, the gas phase acidity trends track with the AlkA excision rates (**Table 4.7**). This is consistent with the proposal that AlkA provides a nonspecific active site and that the ease of nucleobase excision depends on the intrinsic stability of the *N*-glycosidic bond.

Our data also support our hypothesis that AlkA provides a hydrophobic site that enhances the discrimination of damaged from normal bases; in a nonpolar environment, the damaged bases are acidic by a greater amount over the normal bases than in aqueous solution.

Chapter 5. Gas Phase Studies of Adenine Analogs: Implications for Adenine Removal by MutY

5.1. Introduction

Cellular DNA is inevitably damaged by both exogenous and endogenous agents, resulting in a variety of chemical modifications that are associated with mutagenesis, carcinogenesis and aging.[25,26,219,220] Oxidative damage is extremely prevalent, and one of the most common species formed by reactive oxygen species is 7,8-dihydro-8-oxo-guanine (OG).[27,221,222] During DNA replication, adenine (A) is usually inserted opposite OG to form a relatively stable OG:A mismatch.[28] Because undamaged guanine (G) prefers to pair with cytosine (C), not adenine, the oxidation, if not repaired, can result in deleterious DNA mutations (a permanent "G:C to T:A" transversion mutation).

In the face of the constant assault to DNA, organisms have developed elaborate DNA repair pathways. In *Escherichia coli*, oxidative damage is repaired by a "GO" repair pathway that utilizes three enzymes: MutT, Fpg, and MutY.[24,223,224] MutT hydrolyzes the OG deoxynucleoside triphosphate (to yield the OG deoxynucleoside monophosphate and pyrophosphate), preventing its incorporation into replicating DNA.[225] Fpg (also called MutM) cleaves OG from OG:C base pairs, and also catalyzes the cleavage of phosphodiester bonds associated with the resultant abasic site.[26,226] MutY is a somewhat unusual glycosylase enzyme; rather than targeting a damaged base, MutY cleaves adenine (at the N9-C1' bond) when it is base paired to OG (**Figure 5.1**).[26,225,226] Thus, MutY cleaves a normal base - adenine - when it is mispaired. Adenine in A:T pairs remains untouched.
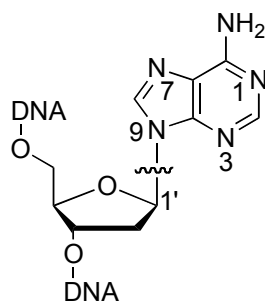
Figure 5.1. The bond cleaved by MutY when adenine is excised.

Because of the importance of base repair to genome integrity, the mechanisms of repair enzymes are of great interest. MutY crystal structures, in particular a 2009 *Bacillus stearothermophilus* structure with a fluorinated 2'-deoxyadenosine, show multiple hydrogen bonding contacts as well as hydrophobic interactions between substrate and enzyme. Kinetic isotope effect studies imply an S_N1-type reaction where the nucleobase leaves (possibly protonated at N7) to yield an oxacarbenium ion which is then attacked by water.[30]

In an effort to lend further insight into the MutY mechanism, our colaborators have, over the past several years, examined the behavior of synthetically derived DNA analogs that are designed to target the elucidation of the interactions responsible for substrate recognition and catalysis (**Figure 5.2**, where "A" is adenine).[227-230] 1-deazaadenine (Z1), 3-deazaadenine (Z3) and 7-deazaadenine (Z) are missing nitrogen at the N1, N3 and N7 positions, respectively (as compared to the parent adenine "A"), and were designed to test the importance of the nitrogen at those positions.[228-231] Substrates 4-methylbenzimidazole (B), 9-methyl-1*H*-imidazo[4,5,b]pyridine (Q) and 4-methylindole (M) are nonpolar isosteres of adenine, designed to test the importance of hydrogen-bonding interactions.[227,229,230]
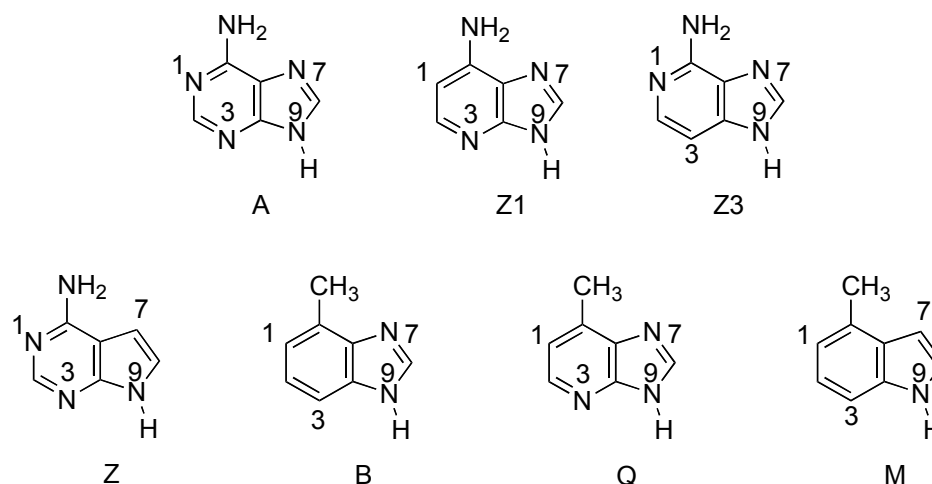
Figure 5.2. Adenine and analogs studied herein.

Although the MutY affinity and excision rates of these analogs have been examined, the fundamental properties have not.[227-229,231] In studying other glycosylases, we found that insight into enzyme mechanism could be gained by characterizing the intrinsic thermochemical properties of substrates.[13,36,38,40,41,76] The acidity and proton affinity of the various nitrogens are related to how easily those sites will hydrogen bond, or be protonated, or be cleaved. We have also found that measuring and calculating those thermochemical values in the gas phase, in the absence of solvent, can be particularly useful for extrapolation to the nonpolar environment of the enzyme active site.[13,36,38,40,41,76] In **Chapter 5**, we focus on the properties of the analogs of adenine, to gain insight into the MutY cleavage mechanism.

5.2. Experimental

All the nucleobase analogs and reference acids and bases are commercially available and were used as received.

The bracketing method (described in details in **Chapter 1**) was used to measure the gas

phase acidity and proton affinity values. A Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR) with dual cell setup (described previously) was used.[10,35-37,40]

The typical protocol for bracketing experiments has been described previously.[10,13,35,39,40] Proton transfer reactions were conducted in both directions. For example, for Z3 acidity bracketing, hydroxide is used to deprotonate neutral Z3. Deprotonated Z3 is transferred into the adjoining cell where it is allowed to react with the neutral reference acid AH with known gas phase acidity. In the opposite direction, the deprotonated reference acid A^- is generated and transferred into the adjoining cell where it is allowed to react with neutral Z3. The occurrence of proton transfer is regarded as evidence that the reaction is exothermic (denoted as "+" in the tables). Bracketing experiments are run under pseudo-first-order conditions with the neutral reactant in excess, relative to the reactant ions. Reading the pressure of the neutral compounds from the ion gauges is not always accurate; therefore, we "back out" the neutral substrate pressure from fast control reactions (described previously).[38-41,79,80]

The gas phase calculations were conducted at the B3LYP/6-31+G(d) level using Gaussian03 and Gaussian09.[90-94] All the structures were fully optimized in the gas phase, and frequencies calculated (no imaginary frequencies were found). Acidity and proton affinity values are reported as ΔH at 298 K.

Dielectric medium calculations were done using the conductor-like polarizable continuum solvent model (CPCM, single point calculations on B3LYP/6-31+G(d) gas phase optimized structures; UAKS cavity) at B3LYP/6-31+G(d) as implemented in Gaussian03.[105,106,107] The "total free energy in solution" values are reported, and the

solvation free energy of a proton (-264.0 kcal mol^{-1}) is accounted for.[28,232,233]

5.3. Results

Our first goal was to characterize these various nucleobase analogs, using theory and experiment. Because cleavage at the N9 position is involved (**Figure 5.1**), the acidity of that site is important: presumably more acidic substrates will have a greater propensity for cleavage, since the resultant conjugate base should be more stable (so a better leaving group). Because the substrate may be protonated before cleavage, the proton affinities of the various heteroatoms and the N9-H acidity for the protonated substrates are relevant values.

The scission of the C1'-N9 bond for these nucleobase analogs has been studied both in the enzyme active site, and in aqueous solution.[227-229,231] Because previous studies from our lab show that gas phase values can be relevant to understanding activity in hydrophobic active sites, we calculate and measure thermochemical properties *in vacuo*. We also calculate relevant values in a water dielectric, to lend insight into how the properties of the nucleobase analogs change in a more polar medium.

5.3.1. Nitrogen proton affinities

The gas phase proton affinities (PAs) for the nucleobase analogs are shown in **Figure 5.3**.[38] For adenine, the most basic site is the N1; this is true for Z3 and Z as well. Since Z1 has no N1, its most basic site is the N3. For B and Q, the N7 is the most basic. M has only one N, at N9.
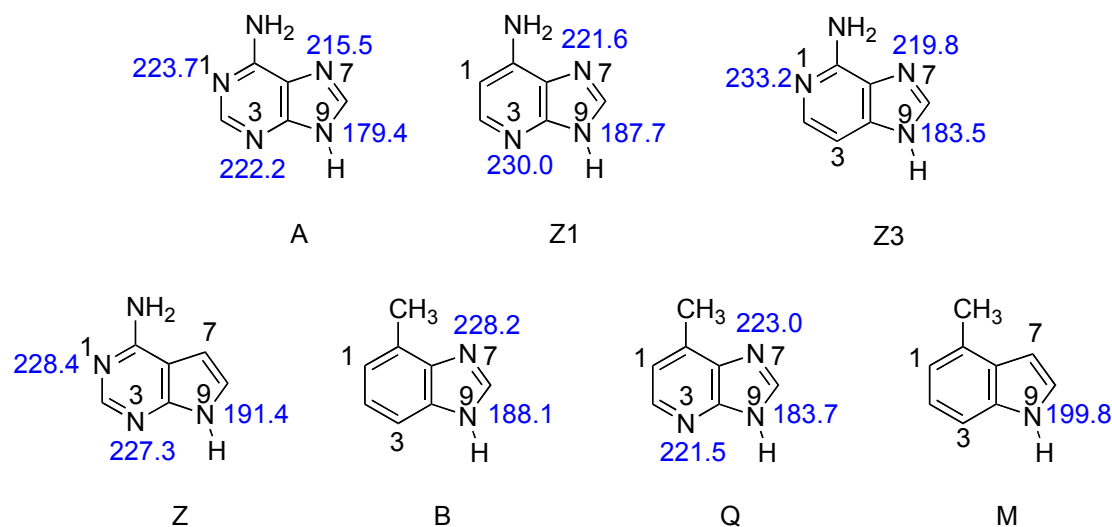
Figure 5.3. Calculated (B3LYP/6-31+G(d)) gas phase proton affinities (kcal mol^{-1}) of the various nitrogens of nucleobase analogs.

5.3.2. N9-H acidity of neutral nucleobase analogs

The calculated (B3LYP/6-31+G(d)) values for the N9-H acidity for the neutral nucleobase analogs are shown in **Figure 5.4.** Lower values are more acidic, so adenine is the most acidic substrate (ΔH_{acid} of 334.8 kcal mol^{-1}).[36,38,234-240] The next most acidic substrate is Z3 (ΔH_{acid} = 335.3 kcal mol^{-1}). The trend from most to least acidic in the gas phase is: adenine > Z3 > Q > B > Z1 > Z > M (where A is most acidic).
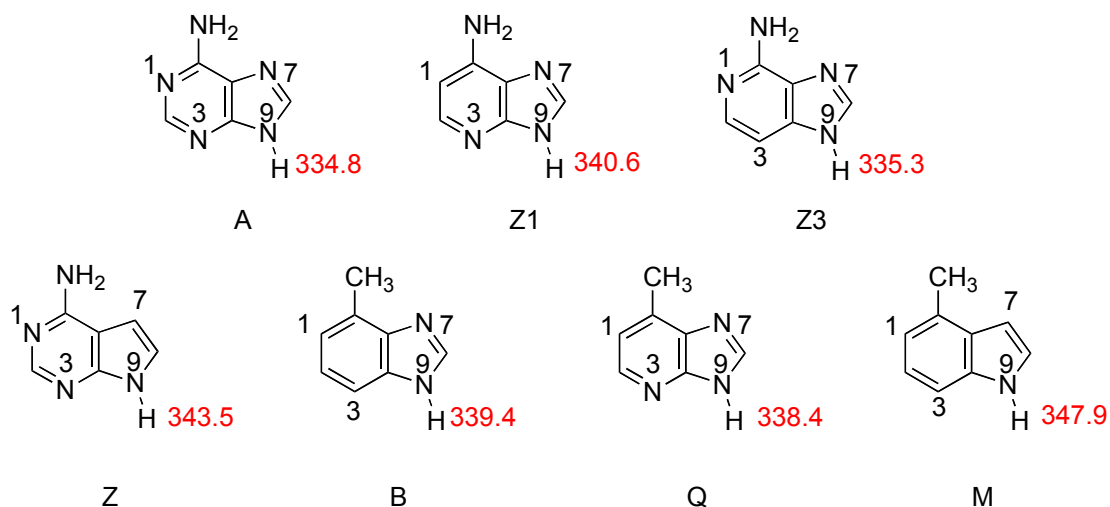
Figure 5.4. Calculated (B3LYP/6-31+G(d)) gas phase acidities (kcal mol^{-1}) of the N9-H for neutral nucleobase analogs.

5.3.3. N9-H acidity of protonated substrates

To assess the effect of protonation on acidity, we calculated the N9-H acidity of the various protonated substrates.

5.3.3.1. Acidity: N1-protonated substrates

The N9-H acidity values for the N1-protonated nucleobase analogs are shown in **Figure 5.5.** Only three of the analogs have an N1 that can be protonated. The most acidic substrate is A, followed by Z3, then Z (A > Z3 > Z). Protonation of the N1 decreases the gas phase acidity values overall by about 90 kcal mol^{-1}.
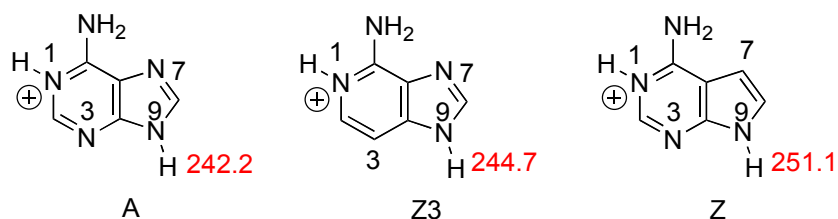
Figure 5.5. Calculated (B3LYP/6-31+G(d)) gas phase acidities (kcal mol^{-1}) of the N9-H for N1-protonated nucleobase analogs.

5.3.3.2. Acidity: N3-protonated substrates

The N9-H acidity values for the N3-protonated nucleobase analogs are shown in **Figure 5.6.** Four substrates have an N3 that can be protonated. As with protonation at N1, protonation at N3 greatly enhances the acidity, decreasing the values in the gas phase by more than 100 kcal mol^{-1} (and about 10 kcal mol^{-1} in water). The trend in terms of acidity in the both media is A ~ Q > Z1 > Z.[241]
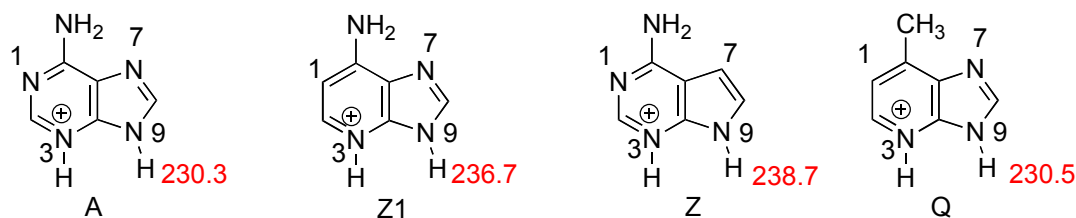
Figure 5.6. Calculated (B3LYP/6-31+G(d)) gas phase acidities (kcal mol^{-1}) of the N9-H for N3-protonated nucleobase analogs.

5.3.3.3. Acidity: N7-protonated substrates

The N9-H acidity values for the N7-protonated nucleobase analogs are shown in **Figure 5.7.** N7-Protonation greatly increases the N9-H acidity (by more than 100 kcal mol^{-1}). The acidity trend is: A ~Z3 > Q > B > Z1.
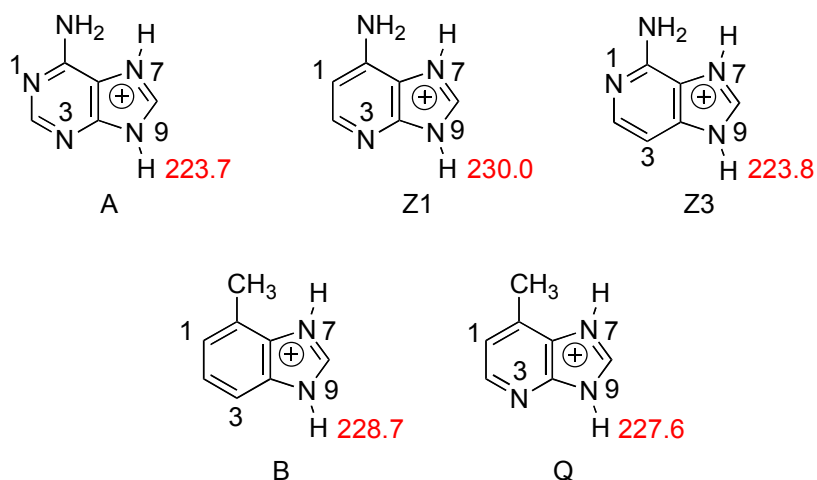
Figure 5.7. Calculated (B3LYP/6-31+G(d)) gas phase acidities (kcal mol^{-1}) of the N9-H for N7-protonated nucleobase analogs.

5.3.4. Gas phase measurements

We also used mass spectrometry to measure the thermochemical properties of some of the nucleobase analogs. These experiments provide new thermochemical data and also allow us to benchmark the calculations. Of the substrates studied herein, only the acidity and proton affinity of adenine have been previously measured ($\Delta H_{acid} = 333 \pm 2$ kcal mol^{-1} and PA = 224 ± 3 kcal mol^{-1}).[36-38,242]

5.3.4.1. Measurements: 3-deazaadenine (Z3).

Previous studies of other nucleobases have taught us that the "canonical" structure is not always the most stable tautomer in the gas phase. We therefore calculated the possible tautomers of Z3 to ascertain which structure is likely to be the most stable *in vacuo* (**Figure 5.8**). The results indicate that the canonical structure should be the most stable, by 4 kcal mol^{-1}. The acidities and proton affinities for the two most stable tautomers are also shown in **Figure 5.8.**
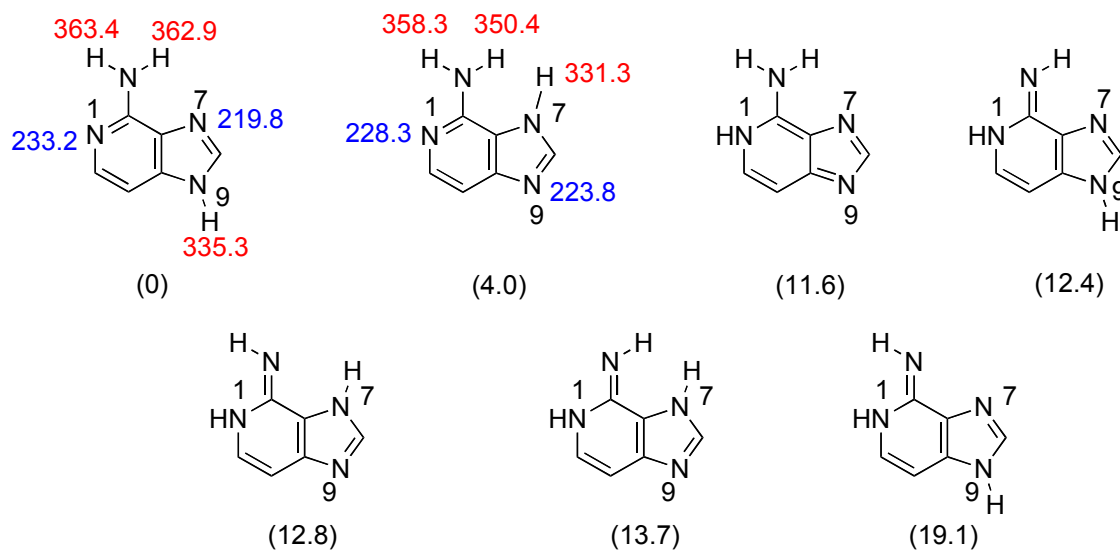
Figure 5.8. Possible tautomers of Z3. Relative stabilities are the parenthetical values. Acidities are in red and proton affinities are in blue. All are calculated values at B3LYP/6-31+G(d) (ΔH at 298 K, kcal mol^{-1}).

The acidity of Z3 was measured using the bracketing method (see Experimental for details; data in **Table 5.1**). We find that deprotonated Z3 is able to deprotonate 2-chloropropanoic acid ($\Delta H_{acid} = 337.0 \pm 2.1$ kcal mol^{-1}). The opposite reaction also occurs; that is, 2-chloropropanoate is able to deprotonate neutral Z3. We therefore bracket the acidity of Z3 to be 337 ± 3 kcal mol^{-1}. This value is consistent with (though about 2 kcal mol^{-1} higher than) the calculated N9-H acidity of the most stable Z3 tautomer (**Figure 5.8**).

Table 5.1. Summary of results for acidity bracketing of Z3

Reference compound	$\Delta H_{acid}{}^a$	Proton transfer^b	
		Ref.	Conj.

		acid	base
methyl cyanoacetate	340.80 ± 0.60	–	+
trifluoro-*m*-cresol	339.3 ± 2.1	–	+
2-chloropropanoic acid	337.0 ± 2.1	+	+
malononitrile	335.8 ± 2.1	+	–
pyruvic acid	333.5 ± 2.9	+	–
difluoroacetic acid	331.0 ± 2.2	+	–

[a] Acidities are in kcal mol^{-1}.[33] [b]A "+" indicates the occurrence and a "–" indicates the absence of proton transfer.

We also bracketed the proton affinity of Z3 (**Table 5.2**). We find that di-*sec*-butylamine (PA = 234.4 ± 2.0 kcal mol^{-1}) is able to deprotonate protonated Z3, but that the opposite reaction (protonated di-sec-butylamine with Z3) does not occur. 1-Methylpiperidine (PA = 224.7 ± 2.0 kcal mol^{-1}) can *not* deprotonate protonated Z3, but Z3 can deprotonate protonated 1-methylpiperidine. We therefore bracket the PA of Z3 to be 233 ± 3 kcal mol^{-1}. The measured proton affinity is consistent with the N1 of Z3 (**Figure 5.8**).

Table 5.2. Summary of results for proton affinity bracketing of Z3.

Reference compound	PA[a]	Proton transfer[b]	
		Ref. base	Conj. acid
2,2,6,6-tetramethylpiperidine	235.9 ± 2.0	+	–
N,N-dimethylcyclohexylamine	235.1 ± 2.0	+	–
triethylamine	234.7 ± 2.0	+	–
di-sec-butylamine	234.4 ± 2.0	+	–
1-methylpiperidine	232.1 ± 2.0	–	+
N,N-dimethylisopropylamine	232.0 ± 2.0	–	+
1-methylpyrrolidine	230.8 ± 2.0	–	+

piperidine	228.0 ± 2.0	–	+

[a] PAs are in kcal mol^{-1}.[33] [b]A "+" indicates the occurrence and a "–" indicates the absence of proton transfer

5.3.4.2. Measurements: 7-deazaadenine (Z).

The possible tautomers for Z are shown in **Figure 5.9.** The canonical tautomer is calculated to be the most stable by a large amount (over 10 kcal mol^{-1}). The acidities and proton affinities for the most stable tautomer are also shown in **Figure 5.9.**
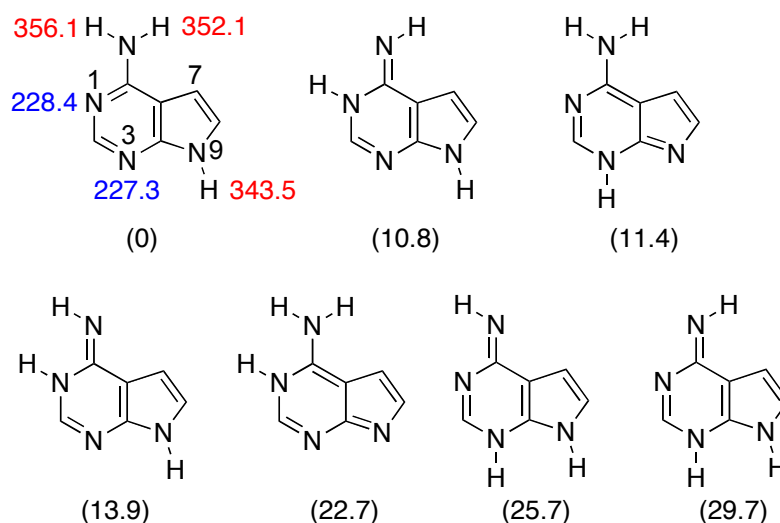
Figure 5.9. Possible tautomers of Z. Relative stabilities are the parenthetical values. Acidities are in red and proton affinities are in blue. All are calculated values at B3LYP/6-31+G(d) (ΔH at 298 K, kcal mol^{-1}).

The bracketing results for the acidity of Z are shown in **Table 5.3.** (Bracketing experiments on compound Z were performed by Ms. Yuan Tian). Butyrate deprotonates Z and deprotonated Z also reacts with butyric acid, placing the ΔH_{acid} at 347 ± 3 kcal mol^{-1}. The experimental value is slightly higher than the calculated N9-H value.

Table 5.3. Summary of results for acidity bracketing of Z.

Reference compound	ΔH_{acid}[a]	Proton transfer[b]	
		Ref. acid	Conj. base
m-cresol	349.5 ± 2.1	–	+
acetic acid	347.4 ± 0.5	–	+
butyric acid	346.8 ± 2.0	+	+
formic acid	346.0 ± 0.5	+	–
methacrylic acid	344.1 ± 2.9	+	–
methyl cyanoacetate	340.80 ± 0.60	+	–

[a] Acidities are in kcal mol^{-1}.[33,182] [b] A "+" indicates the occurrence and a "–" indicates the absence of proton transfer.

The results for proton affinity measurement are shown in **Table 5.4.** Z deprotonates protonated piperidine, and piperidine deprotonates protonated Z, placing the PA at 228 ± 3 kcal mol^{-1} (**Table 5.4**). The measured proton affinity is consistent with both the N1 and the N3 of Z .

Table 5.4. Summary of results for PA bracketing of Z

Reference compound	PA[a]	Proton transfer[b]	
		Ref. base	Conj. acid
1-methylpiperidine	232.1±2.0	+	–
1-methylpyrrolidine	230.8±2.0	+	–
piperidine	228.0±2.0	+	+
pyrrolidine	226.6±2.0	–	+
3-picoline	225.5±2.0	–	+

[a] PAs are in kcal mol^{-1}.[33] [b] A "+" indicates the occurrence and a "–" indicates the absence of proton transfer

5.4. Discussion.

5.4.1. Gas phase data.

A comparison of calculated and measured thermochemical values for Z3 and Z indicate that calculations (both acidity and PA) at B3LYP/6-31+G(d) are reasonably accurate (**Table 5.5**). These newly measured values therefore allow us to benchmark our calculations on these heretofore uncharacterized adenine analogs.

Table 5.5. Comparison of calculated and measured thermochemical values.

Substrate	Acidity or proton affinity [a]	Calculated (B3LYP/6-31+G(d)) [a]	Experimental [a]
adenine[b]	ΔH_{acid}	334.8	333 ± 2
	PA	223.7	224 ± 3
Z3	ΔH_{acid}	335.3	337 ± 3
	PA	233.2	233 ± 3
Z	ΔH_{acid}	343.5	347 ± 3
	PA	228.4	228 ± 3

[a]Values are in kcal mol^{-1}; [b]References[36-38,237,242]

5.4.2. Possible enzyme mechanisms.

Our interest is to relate our results to the possible mechanisms by which MutY excises adenine from DNA. The rates of excision by MutY of the nucleobase analogs are summarized in **Table 5.6**.[243]

Table 5.6. Relative MutY excision rates of nucleobase analogs.

Nucleobase	Excision rate, relative to A
A	
Z3	100x less than A[a]
Z1	40x less than A[b]
B	6000x less than A[c]
Q	10x less than A[c]
M	None[d]
Z	None[e]

[a]Reference[228] [b] Reference[230] [c]Reference[227] [d]Reference[229] [e]Reference[231]

Various mechanisms of cleavage by the enzyme are possible; in any scenario, the nucleobase substrate, in some form, is a leaving group (**Scheme 5.1**). The primary possible mechanisms are either simple cleavage of the deprotonated anionic nucleobase (Mechanism A, **Scheme 5.1**) or pre-protonation of the nucleobase, which should enhance excision (Mechanisms B, C, D, **Scheme 5.1**, showing protonation of N1, N3, and N7, respectively). The ease by which a given substrate is cleaved should be related to its leaving group ability. Because typically better leaving groups have stronger conjugate acids, we calculated the gas-phase acidities at the cleavage site (N9) for the various nucleobase analogs to assess whether a correlation between acidity and MutY excision rates could be found. We focus on gas phase values because these reveal intrinsic reactivity, and may be related as well to reactivity in the nonpolar active site.[13,36,38,40,41,76]

Scheme 5.1.
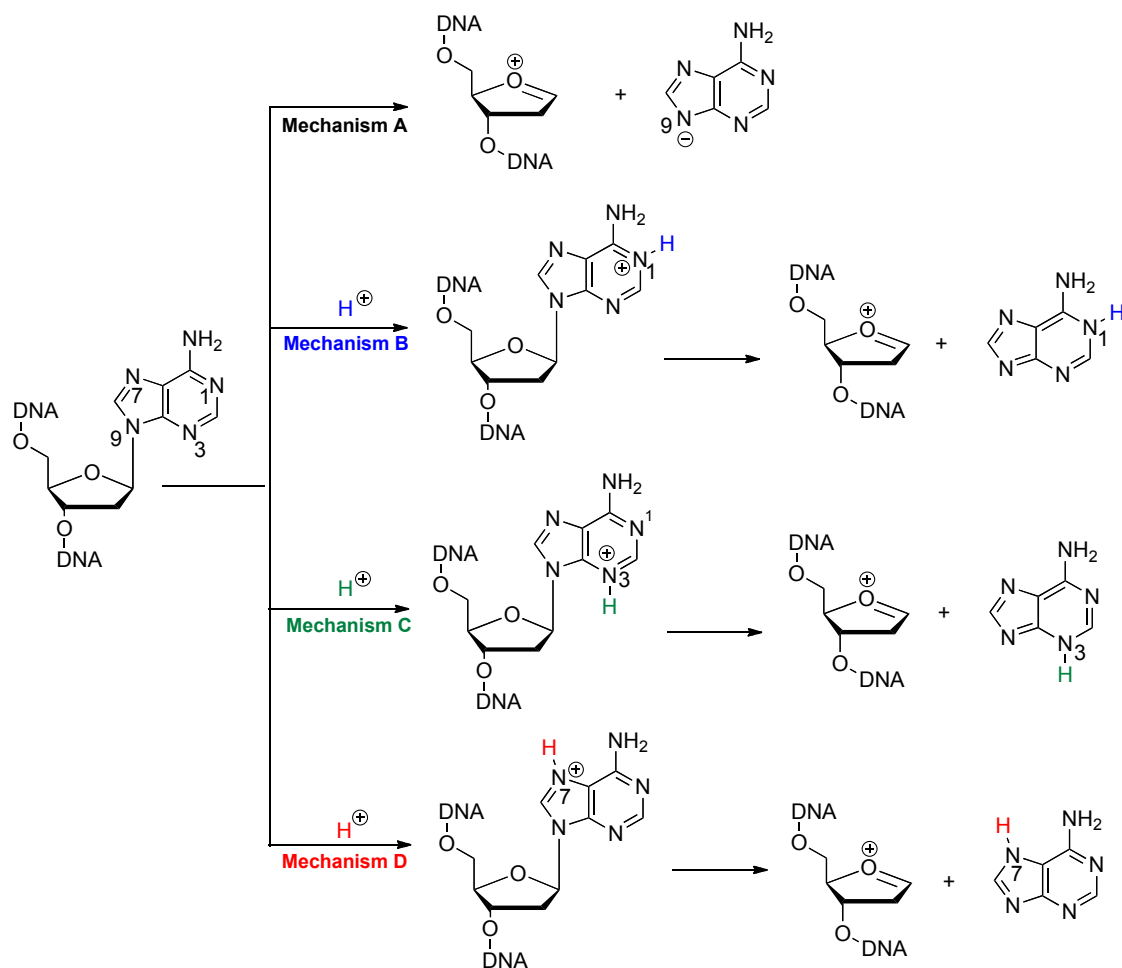
5.4.2.1. Deprotonated adenine as the leaving group.

The simplest mechanism would be the excision of the C1'-N9 bond wherein the nucleobase leaves as the deprotonated N9 anion (**Scheme 5.1**, Mechanism A). The relevant thermochemical value for this mechanism is the acidity of the N9-H proton for each analog: the more acidic that position, the more stable the conjugate base anion is. The more stable that conjugate base anion is, the better a leaving group it will be. Therefore, one would expect a correlation between N9-H acidity and the rates of excision by MutY shown in **Table 5.6.**

Our calculations indicate that the trend from most to least acidic in the gas phase is: A > Z3 > Q > B > Z1 > Z > M (where A is most acidic, **Figure 5.4**). **Table 5.6** indicates that the trend in terms of MutY excision (from highest to lowest excision rate) is A > Q > Z1 > Z3 > B > M ~ Z. Therefore, the N9-H gas phase acidity does not follow the same trend as the MutY excision rates, implying that the cleavage of these damaged bases does not occur by simple cleavage of the N9$^-$ anion.

5.4.2.2. N1 pre-protonation.

Another possibility is that the N1 is protonated prior to excision (**Scheme 5.1**, Mechanism B). The acidity trend for those substrates which have a nitrogen at the 1 position that can be protonated is A > Z3 > Z (where A is most acidic, **Figure 5.5**). The trend for the excision of the nucleobase analogs by MutY is: A > Q > Z1 > Z3 > B > M ~ Z. Therefore, the gas phase acidity of the three substrates does correlate to their MutY excision rates, though the comparison only involves a limited set of substrates (since only three of the analogs have an N1). Still, we can conclude that the gas phase acidity and the known MutY excision rates do correlate, such that pre-protonation of N1 could be involved in cleavage. However, because other analogs that are cleaved quickly by MutY, such as Q and Z1, do not have a nitrogen at the 1-position, it seems unlikely that protonation of that site would be the sole mechanism by which the enzyme enhances excision.

5.4.2.3. N3 pre-protonation.

We therefore also consider N3 protonation prior to excision (Mechanism C, **Scheme 5.1**). The calculated N9-H acidity values for those substrates which have a nitrogen at the

3 position that can be protonated are shown in **Figure 5.6.** The trend in terms of acidity for these substrates is A ~ Q > Z1 > Z (where A is most acidic). Again, since the MutY excision rate is A > Q > Z1 > Z3 > B > M ~ Z, the acidity of N3-protonated substrates is consistent with the relative MutY cleavage rates. However, because substrates such as B, which lacks an N3, are still cleaved more quickly than Z, it seems unlikely that N3 protonation is the *only* advantage provided by the enzyme.

5.4.2.4. N7 pre-protonation.

The remaining ring nitrogen that could be protonated to enhance excision is the N7 (**Scheme 5.1**, Mechanism D). The calculated acidity values for those substrates which have a nitrogen at the 7 position that can be protonated are shown in **Figure 5.7.** The trend in terms of acidity for these substrates is A ~ Z3 > Q > B > Z1 (where A is most acidic). The MutY catalyzed excision rate trend is A > Q > Z1 > Z3 > B > M ~ Z. There does not appear to be a correlation between the two trends -- for example, while the calculations indicate that N7 protonation should make Z3 as acidic as A (and therefore as cleavable), the enzyme actually cleaves Z3 more slowly than it does Q and Z1.

However, our other calculations do show that N1 and N3 protonation also enhance N9-H acidity. Therefore, an excision mechanism could involve protonation of one nitrogen and hydrogen bonding at other nitrogen(s) to further enhance acidity (and "cleavability"). Because M and Z, which are missing N7, are cleaved most slowly by MutY, we reasoned that N7 is key to cleavage, probably as a protonation site. Both N1 and N3 could then also be involved in hydrogen bonds in the active site to further encourage cleavage.

First we consider protonation at N7 with hydrogen bonding to N1. Nucleobase analogs that can protonate at N7 and bind at N1 are A and Z3. Because A and Z3 are already the two most acidic N7-protonated substrates, enhancement of those acidities via hydrogen bonding at N1 would not change the gas phase acidity trend.

What about hydrogen bonding to N3? For the nucleobase analogs with protonation at N7, the acidity trend is: A ~ Z3 > Q > B > Z1 (where A is most acidic). N7-Protonated substrates that have an N3 to which a hydrogen bond could be formed are A, Q, and Z1. This hydrogen bonding could alter the aforementioned trend, to make Q and Z1 more acidic at the N9-H, comparable to or even more acidic than Z3. The acidity trend of N3-hydrogen bonded, N7-protonated substrates could therefore shift to A > Q> Z1 > Z3 > B, which is comparable to the MutY catalyzed excision rate trend of A > Q > Z1 > Z3 > B > M ~ Z.

In summary, our measurements and calculations of the thermochemical properties of the various MutY nucleobase analogs are consistent with a mechanism where hydrogen bonding to N3 and protonation of N7 lead to more facile cleavage at N9. Hydrogen bonding at N1 could also be involved. Other experimental data provide support for this mechanism as well. The most recent crystal structure, from *Bacillus stearothermophilus,* is a complex of MutY with a fluorinated 2'-deoxyadenosine, and shows multiple hydrogen bonding contacts as well as hydrophobic interactions between substrate and enzyme.[30,244,245] Glu-43 and Tyr-126 coordinately contact the N7; the position of the glutamate indicates that it is probably protonated (the carboxylic acid as opposed to the carboxylate). Glu43 is expected to be quite acidic, allowing partial or full bonding of its proton to N7. A hydrogen bond from Arg-26 to water to N3 is observed, as well as a

hydrogen bond from Arg-31 to N1. Various salt bridges exclude water and create a hydrophobic environ. The crystal structure is therefore consistent with protonation at N7 and hydrogen bonding to N1 and N3, which correlates to the thermochemical properties of the nucleobase analogs. Furthermore, recent kinetic isotope effect studies indicate an inverse ^{15}N7 KIE (in *E. coli* MutY), consistent with protonation at N7.[29]

5.4.3. Possible aqueous mechanisms.

We have also studied the excision of some of these nucleobase analogs in acidic aqueous solution and found that B, Q and Z1 are depurinated more quickly than A while Z3 is depurinated more slowly (**Table 5.7**).[227,230] Because these cleavage reactions occur in water under acidic conditions, relevant calculations should involve pre-protonation and a polar environment. To that end, we calculated the N9-H acidity for the nucleobase analogs when the most basic site is protonated. To mimic water, we conducted the calculations in a water dielectric (**Figure 5.10**). The acidity values are much lower than those in the gas phase, as would be expected. These calculations indicate an N9-H acidity trend of: Q > B > Z1 > A > Z3. We would therefore expect Q, B and Z1 to be cleaved more quickly than adenine, and Z3 to be cleaved more slowly, which is consistent with the experimental results. The calculations, which do not include specific solvation, are not perfect; experiments indicate that Q and Z1 are cleaved with equal facility in acidic water, but the calculations predict that Q is more acidic at N9-H than Z1, and therefore should be easier to excise. However, the overall trend of which bases should be cleaved more quickly than A and which less quickly is consistent between calculations and experiment.

Table 5.7. Relative excision rates of nucleobase analogs in acidic aqueous solution.

Nucleobase	Acid-catalyzed depurination rate, relative to A
A	–
Z3	3x less than A[a]
Z1	7x more than A[a]
B	2x more than A[b]
Q	7x more than A[b]

[a]Reference[230]. [b]Reference[227]
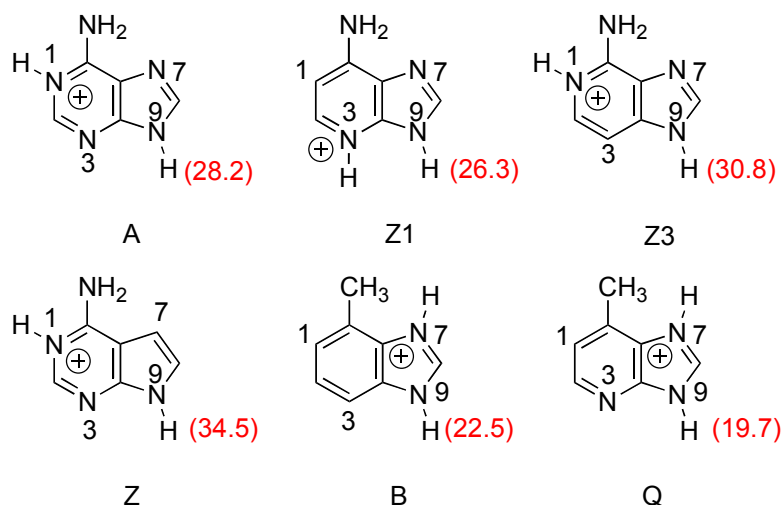
Figure 5.10. Aqueous N9-H acidities of nucleobase analogs with the most basic site protonated, in kcal mol^{-1}.

5.4.4. 1,3-Deazaadenine (Z13) Prediction.

One additional substrate we studied computationally is Z13 (**Figure 5.11**). This is a logical extension of the various analogs already studied (**Figure 5.2**); this particular derivative is missing an "N" at both the 1 and 3 positions. Z13 could be a substrate for

MutY, although the only nitrogen available for hydrogen bonding and/or protonation is the N7. Compared to the other nucleobase analogs, the acidity at N9-H when that N7 is protonated is fairly poor (ΔH_{acid} = 230.1 kcal mol^{-1}), comparable to Z1 (**Figure 5.11** versus **Figure 5.7**). We would therefore expect Z13 to be cleaved by MutY slowly, comparable to Z1. Protonation of the N7 and calculation of the N9-H acidity in a water dielectric drops the ΔH_{acid} to 23.6 kcal mol^{-1} (**Figure 5.11**). Comparing this value to those for the other nucleobase analogs (**Figure 5.10**), we would predict that Z13 would be depurinated in acidic water quite quickly, certainly with more ease than adenine.
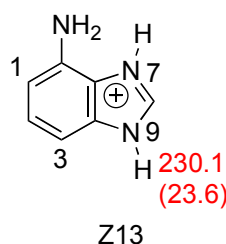
Figure 5.11. Calculated (B3LYP/6-31+G(d)) gas phase and (aqueous) acidities (kcal mol^{-1}) of the N9-H for N7-protonated nucleobase analog Z13.

5.5. Conclusions.

The heretofore unknown thermochemical properties of adenine and six adenine analogs have been calculated and measured herein. Gas phase measurements benchmark our calculations. Comparison of the stability of the N9-H bond (in terms of acidity) when various nitrogens are protonated versus known MutY excision rates point to a MutY-catalyzed mechanism involving protonation at N7, and hydrogen bonding at N3 (and possibly N1). This conclusion is consistent with other MutY mechanistic studies (crystal

structures, kinetic isotope effects). We also find that our calculations for the N9-H acidity when the most basic site is protonated are consistent with experimental data for acid-catalyzed depurination in water. Our work shows that fundamental studies of biological species are valuable for lending insight into mechanisms for which these species are substrates.

Reference

(1)   Kanvah, S.; Joshy, J.; Schuster, G. B.; Barnett, R. N.; Cleveland, C. L.; Landman, U. *Acc. Chem. Res.* **2010**, *43*, 280-287.

(2)   Watson, J. D. *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*; Norton Critical Editions, 1968.

(3)   Stone, M. P.; Huang, H.; Brown, K. L.; Shanmugam, G. *Chemistry & Biodiversity* **2011**, *8*, 1571-1615.

(4)   *DNA Damage and Repair, Vol. 1: DNA Repair in Procaryotes and Lower Eucaryotes*; Nickoloff, J. A.; Hoekstra, M. F., Eds.; Humana Press, Inc.: Totowa, NJ, 1998.

(5)   Berti, P. J.; McCann, J. A. B. *Chem. Rev.* **2006**, *106*, 506-555.

(6)   Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A.; Ellenberger, T. *DNA Repair and Mutagenesis*; 2nd ed.; ASM Press: Washington, DC, 2006.

(7)   Milanowska, K.; Rother, K.; Bujnicki, J. M. *Mol. Biol. Internat.* **2011**, 1-9.

(8)   Stivers, J. T.; Jiang, Y. L. *Chem. Rev.* **2003**, *103*, 2729-2759.

(9)   Bellamy, S. R. W.; Krusong, K.; Baldwin, G. S. *Nucleic Acids Research* **2007**, *35*, 1478-1487.

(10)   Kurinovich, M. A.; Lee, J. K. *J. Am. Chem. Soc.* **2000**, *122*, 6258-6262.

(11)   Shroyer, M. J. N.; Bennett, S. E.; Putnam, C. D.; Tainer, J. A.; Mosbaugh, D. W. *Biochemistry* **1999**, *38*, 4834-4845.

(12)   Kimura, E.; Kitamura, H.; Koike, T.; Shiro, M. *J. Am. Chem. Soc.* **1997**, *119*, 10909-10919.

(13)   Kurinovich, M. A.; Lee, J. K. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 985-995.

(14)   Parikh, S. S.; Walcher, G.; Jones, G. D.; Slupphaug, G.; Krokan, H. E.; Blackburn, G. M.; Tainer, J. A. *Proc. Natl Acad. Sci.* **2000**, *97*, 5083-5088.

(15)   Panayotou, G.; Brown, T.; Barlow, T.; Pearl, L. H.; Savva, R. *J. Biol. Chem.* **1998**, *273*, 45-50.

(16)   Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A. *Nature* **1996**, *384*, 87-92.

(17) Savva, R.; McAuley-Hecht, K.; Brown, T.; Pearl, L. *Nature* **1995**, *373*, 487-493.

(18) Mol, C. D.; Arvai, A. S.; Slupphaug, G.; Kavli, B.; Alseth, I.; Krokan, H. E.; Tainer, J. A. *Cell* **1995**, *80*, 869-878.

(19) Thomas, L.; Yang, C.-H.; Goldthwait, D. A. *Biochemistry* **1982**, *21*, 1162-1169.

(20) O'Brien, P. J.; Ellenberger, T. *J. Biol. Chem.* **2004**, *279*, 26876-26884.

(21) Bowman, B. R.; Lee, S.; Wang, S.; Verdine, G. L. *J. Biol. Chem.* **2010**, *285*, 35783-35791.

(22) Yamagata, Y.; Kato, S.; Odawara, K.; Tokuno, Y.; Nakashima, Y.; Matsushima, N.; Yasumura, K.; Tomita, K.; Ihara, K.; Fujii, Y.; Nakabeppu, Y.; Sekiguchi, M.; Fujii, S. *Cell* **1996**, *86*, 311-319.

(23) Zhao, B.; O'Brien, P. J. *Biochemistry* **2011**, *50*, 4350-4359.

(24) David, S. S.; O'Shea, V. L.; Kundu, S. *Nature* **2007**, *447*, 941-950.

(25) Neeley, W. L.; Essigmann, J. M. *Chem. Res. Toxicol.* **2006**, *19*, 491-505.

(26) David, S. S.; Williams, S. D. *Chem. Rev.* **1998**, *98*, 1221-1261.

(27) Burrows, C. J.; Muller, J. G. *Chem. Rev.* **1998**, *98*, 1109-1152.

(28) McAuley-Hecht, K. E.; Leonard, G. A.; Gibson, N. J.; Thomson, J. B.; Watson, W. P.; Hunter, W. N.; Brown, T. *Biochemistry* **1994**, *33*, 10266-10270.

(29) McCann, J. A. B.; Berti, P. J. *J. Am. Chem. Soc.* **2008**, *130*, 5789-5797.

(30) Lee, S.; Verdine, G. L. *Proc. Natl. Acad. Sci.* **2009**, *106*, 18497-18502.

(31) Simonson, T.; Brooks III, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452-8458.

(32) Gilson, M. K.; Honig, B. H. *Biopolymers* 1986, *25*, 2097-2119.

(33) NIST Chemistry WebBook, NIST Standard Reference Database Number 69; retrieved in 2011. Linstrom, P. J.; Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD 20899, http://webbook.nist.gov.

(34) Hunter, E. P.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413-656.

(35) Kurinovich, M. A.; Lee, J. K. *Chem. Commun.* **2002**, 2354-2355.

(36) Sharma, S.; Lee, J. K. *J. Org. Chem.* **2002**, *67*, 8360-8365.

(37)   Sharma, S.; Lee, J. K. *J. Org. Chem.* **2004**, *69*, 7018-7025.

(38)   Zhachkina, A.; Liu, M.; Sun, X.; Amegayibor, S.; Lee, J. K. *J. Org. Chem.* **2009**, *74*, 7429-7440.

(39)   Liu, M.; Li, T.; Amegayibor, S.; Cardoso, D. S.; Fu, Y.; Lee, J. K. *J. Org. Chem.* **2008**, *73*, 9283-9291.

(40)   Liu, M.; Xu, M.; Lee, J. K. *J. Org. Chem.* **2008**, *73*, 5907-5914.

(41)   Sun, X.; Lee, J. K. *J. Org. Chem.* **2007**, *72*, 6548-6555.

(42)   Kabarle, P. *Annu. Rev. Phys. Chem.* **1977**, *28*, 445-476.

(43)   Ervin, K. M. *Chem. Rev.* **2001**, *101*, 391-444.

(44)   Farneth, W. E.; Brauman, J. I. *J. Am. Chem. Soc.* **1976**, *98*, 7891-7898.

(45)   Grabowski, J. J.; DePuy, C. H.; Bierbaum, V. M. *J. Am. Chem. Soc.* **1983**, *105*, 2565-2571.

(46)   Gronert, S. *Chem. Rev.* **2001**, *101*, 329-360.

(47)   A solution phase S_N2 PES is "bell" shaped mainly due to the significant stabilization of both product and reactant ions by solvation, and smaller stabilization of the highly delocalized TS. Initial decrease in potential energy is washed away in solution by ions' desolvation. energy.

(48)   Riveros, J. M.; Jose, S. M.; Takashima, K. *Adv. Phys. Org. Chem.* **1985**, *21*, 197-240.

(49)   Olmstead, W. N.; Brauman, J. I. *J. Am. Chem. Soc.* **1977**, *99*, 4219-4228.

(50)   DePuy, C. H.; Gronert, S.; Mullin, A.; Bierbaum, V. M. *J. Am. Chem. Soc.* **1990**, *112*, 8650-8655.

(51)   Bohme, D. K.; Young, L. B. *J. Am. Chem. Soc.* **1970**, *92*, 7354-7358.

(52)   Tanaka, K.; Mackay, G. I.; Payzant, J. D.; Bohme, D. K. *Can. J. Chem.* **1976**, *54*, 1643-1659.

(53)   Pellerite, M. J.; Brauman, J. I. *J. Am. Chem. Soc.* **1983**, *105*, 2672-2680.

(54)   Wang, H.; Peslherbe, G. H.; Hase, W. L. *J. Am. Chem. Soc.* **1994**, *116*, 9644-9651.

(55)   Ingemann, S.; Nibbering, N. M. M. *Can. J. Chem.* **1984**, *62*, 2273-2281.

(56)   De Hoffmann, E.; Charette, J.; Stroobant, V. *Mass Spectrometry Principles and Applications*; John Wiley & Sons: Chichester, 1999.

(57)   Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.

(58)   Lawrence, E. O.; Edlefsen, N. E. *Science* **1930**, *72*, 376-377.

(59)   Sommer, H.; Thomas, H. A.; Hipple, J. A. *Phys. Rev.* **1949**, *76*, 1877-1878.

(60)   Comisarow, M. B.; Marshall, A. G. *J. Phys. Lett.* **1974**, *25*, 282-283.

(61)   Marshall, A. G.; Grosshans, P. B. *Anal. Chem.* **1991**, *63*, 215-229.

(62)   March, R. E. *Quadrupole Ion Trap Mass Spectrometer* Chichester, 2000.

(63)   Gronert, S. *Mass Spectrom. Rev.* **2005**, *24*, 100-120.

(64)   Paul, W.; Steinwedel, H. S. *US Patent* **1960**, *2*, 939-952.

(65)   Paul, W. *Angew. Chem.* **1990**, *102*, 780-789.

(66)   Stafford, G. C., Jr.; Kelley, P. E.; Syka, J. E. P.; Reynolds, W. E.; Todd, J. F. J. *Int. J. Mass Spectrom. and Ion Processes* **1984**, *60*, 85-98.

(67)   Glish, G. L.; Vachet, R. W. *Nature* **2003**, *2*, 140-150.

(68)   The scheme is from Thermo Scientific (former Finnigan) manual.

(69)   Zeleny, J. *Phys. Rev.* **1914**, *3*, 69-91.

(70)   Cole, R. B. *Electrospray Ionization Mass Spectrometry*; Willey: Chichester, 1997.

(71)   Bruins, A. P. *Mass Spectrom. Rev.* **1991**, *10*, 53-77.

(72)   Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.

(73)   Gomez, A.; Tang, K. *Phys. Fluids* **1994**, *6*, 404-414.

(74)   Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804-817.

(75)   Taylor, G. *Proc. Royal Soc. London Ser. A* **1964**, *280*, 383-397.

(76)   Lee, J. K. *Int. J. Mass Spectrom.* **2005**, *240*, 261-272.

(77)   Zhachkina, A.; Lee, J. K. *J. Am. Chem. Soc.* **2009**, *131*, 18376-18385.

(78)   Michelson, A. Z.; Petronico, A.; Lee, J. K. *J. Org. Chem.* **2012**, *77*, 1623-1631.

(79)   Chesnavich, W. J.; Su, T.; Bowers, M. T. *J. Chem. Phys.* **1980**, *72*, 2641-2655.

(80)   Su, T.; Chesnavich, W. J. *J. Chem. Phys.* **1982**, *76*, 5183-5185.

(81)   Miller, K. J.; Savchik, J. A. *J. Am. Chem. Soc.* **1979**, *101*, 7206-7213.

(82)   Bartmess, J. E.; Georgiadis, R. M. *Vacuum* **1983**, *33*, 149-153.

(83)   Su, T.; Bowers, M. T. *J. Am. Chem. Soc.* **1973**, *95*, 1370-1373.

(84)   Cooks, R. G.; Kruger, T. L. *J. Am. Chem. Soc.* **1977**, *99*, 1279-1281.

(85)   McLuckey, S. A.; Cameron, D.; Cooks, R. G. *J. Am. Chem. Soc.* **1981**, *103*, 1313-1317.

(86)   McLuckey, S. A.; Cooks, R. G.; Fulford, J. E. *Int. J. Mass Spectrom. and Ion Physics* **1983**, *52*, 165-174.

(87)   Green-Church, K. B.; Limbach, P. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 24-32.

(88)   Drahos, L.; Vékey, K. *J. Mass Spectrom.* **1999**, *34*, 79-84.

(89)   Gronert, S.; Feng, W. Y.; Chew, F.; Wu, W. *Int. J. Mass Spectrom.* **2000**, *195/196*, 251-258.

(90)   Frisch, M. J.; Trucks, G.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; C., B. J.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Peterson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nikai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G.; Salvador, P.; Dannenberg, J. J.; Zakrewski, V. G.; Dapprich, S.; Daniels, A.; Strain, M.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavarchari, K.; Foresman, J.; Ortiz, J.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M.; Johnson, B.; Chen, W.; Wong, M.; Gonzalez, C.; Pople, J. A. Gaussian03; Gaussian, Inc., Wallingford CT, 2004. .

(91)   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.;

Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O .; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision A.02; Gaussian, Inc., Wallingford CT, 2009.

(92)   Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.

(93)   Kohn, W.; Becke, A. D.; Parr, R. G. *J. Chem. Phys.* **1996**, *100*, 12974-12980.

(94)   Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

(95)   Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

(96)   Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157-167.

(97)   Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503-506.

(98)   Saebø, S.; Almlöf, J. *Chem. Phys. Lett.* **1989**, *154* 83-89.

(99)   Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618-622.

(100)   Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275-280.

(101)   Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 281-289.

(102)   Head-Gordon, M.; Head-Gordon, T. *Chem. Phys. Lett.* **1994**, *220*, 122-128.

(103)   Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822-2827.

(104)   Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **2000**, *112*, 6532-6542.

(105)   Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995-2001.

(106)   Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comp. Chem.* **2003**, *24*, 669-681.

(107)   Takano, Y.; Houk, K. N. *J. Chem. Theory Comput.* **2005**, *1*, 70-77.

(108)   Kavli, B.; Otterlei, M.; Slupphaug, G.; Krokan, H. E. *DNA Repair* **2007**, *6*, 505-516.

(109)   Connolly, B. A.; Fogg, M. J.; Shuttleworth, G.; Wilson, B. T. *Biochem. Soc. Transl.* **2003**, *31*, 699-702.

(110)   Parker, J. B.; Bianchet, M. A.; Krosky, D. J.; Friedman, J. I.; Amzel, L. M.; Stivers, J. T. *Nature* **2007**, *449*, 433-438.

(111)   Seiple, L.; Jaruga, P.; Dizdaroglu, M.; Stivers, J. T. *Nucleic Acids Res.* **2006**, *34*, 140-151.

(112)   Stivers, J. T. *Chem. Eur. J.* **2008**, *14*, 786-793.

(113)   Porecha, R. H.; Stivers, J. T. *Proc. Natl Acad. Sci.* **2008**, *105*, 10791-10796.

(114)   Krosky, D. J.; Schwarz, F. P.; Stivers, J. T. *Biochemistry* **2004**, *43*, 4188-4195.

(115)   Dong, J.; Drohat, A. C.; Stivers, J. T.; Pankiewicz, K. W.; Carey, P. R. *Biochemistry* **2000**, *39*, 13241-13250.

(116)   Drohat, A. C.; Stivers, J. T. *Biochemistry* **2000**, *39*, 11865-11875.

(117)   Drohat, A. C.; Stivers, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 1840-1841.

(118)   Jiang, Y. L.; Drohat, A. C.; Ichikawa, Y.; Stivers, J. T. *J. Biol. Chem.* **2002**, *277*, 15385-15392.

(119)   Drohat, A. C.; Jagadeesh, J.; Ferguson, E.; Stivers, J. T. *Biochemistry* **1999**, *38*, 11866-11875.

(120)   Drohat, A. C.; Xiao, G.; Tordova, M.; Jagadeesh, J.; Pankiewicz, K. W.; Watanabe, K. A.; Gilliland, G. L.; Stivers, J. T. *Biochemistry* **1999**, *38*, 11876-11886.

(121)   Werner, R. M.; Stivers, J. T. *Biochemistry* **2000**, *39*, 14054-14064.

(122)   Stivers, J. T.; Drohat, A. C. *Arch. Biochem. Biophys.* **2001**, *396*, 1-9.

(123)   Guthrie, R. D.; Jencks, W. P. *Acc. Chem. Res.* **1989**, *22*, 343-349.

(124)   Dinner, A. R.; Blackburn, G. M.; Karplus, M. *Nature* **2001**, *413*, 752-755.

(125)   Parker, J. B.; Stivers, J. T. *Biochemistry* **2008**, *47*, 8614-8622.

(126)   DePuy, C. H. *Int. J. Mass Spectrom.* **2000**, *200*, 79-96.

(127)   DePuy, C. H.; Bierbaum, V. M. *Acc. Chem. Res.* **1981**, *14*, 146-153.

(128)   Cao, C.; Jiang, Y. L.; Krosky, D. J.; Stivers, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 13034-13035.

(129)   Bennett, M. T.; Rodgers, M. T.; Hebert, A. S.; Ruslander, L. E.; Eisele, L.; Drohat, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 12510-12519.

(130)   Newkome, G. R.; Nayak, A.; Otemaa, J.; Van, D. A.; Benton, W. H. *J. Org. Chem.* **1978**, *43*, 3362-3367.

(131)   Milligan, J. R.; Ward, J. F. *Radiation Research* **1994**, *137*, 295-299.

(132)   Grabowski, J. J.; Zhang, L. *J. Am. Chem. Soc.* **1989**, *111*, 1193-1203.

(133)   Lee, J. K.; Grabowski, J. J. *J. Org. Chem.* **1996**, *61*, 9422-9429.

(134)   Schulze, S. M.; Santella, N.; Grabowski, J. J.; Lee, J. K. *J. Org. Chem.* **2001**, *66*, 7247-7253.

(135)   Brodbelt-Lustig, J. S.; Cooks, R. G. *Talanta* **1989**, *36*, 255-260.

(136)   Swart, M.; Sola, M.; Bickelhaupt, F. M. *J. Comput. Chem.* **2007**, *28*, 1551-1560.

(137)   Westaway, K. C.; Fang, Y.; MacMillar, S.; Matsson, O.; Poirier, R. A.; Islam, S. M. *J. Phys. Chem. A* **2007**, *111*, 8110-8120.

(138)   Bento, A. P.; Sola, M.; Bickelhaupt, F. M. *J. Comput. Chem.* **2005**, *26*, 1497-1504.

(139)   Vayner, G.; Houk, K. N.; Jorgensen, W. L.; Brauman, J. I. *J. Am. Chem. Soc.* **2004**, *126*, 9054-9058.

(140)   Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361-373.

(141)   Gronert, S.; Pratt, L. M.; Mogali, S. *J. Am. Chem. Soc.* **2001**, *123*, 3081-3091.

(142)   Gronert, S. *Acc. Chem. Res.* **2003**, *36*, 848-857.

(143)   Gronert, S.; Fong, L.-M. *Int. J. Mass Spectrom.* **1999**, *192*, 185-190.

(144)   Gronert, S.; Flores, A. E. *J. Am. Chem. Soc.* **1999**, *121*, 2627-2628.

(145)   Gronert, S. *J. Mass Spectrom.* **1999**, *34*, 787-796.

(146)   Gronert, S.; Fagin, A. E.; Okamoto, K.; Mogali, S.; Pratt, L. M. *J. Am. Chem. Soc.* **2004**, *126*, 12977-12983.

(147)   Gronert, S.; Fagin, A. E.; Wong, L. *J. Am. Chem. Soc.* **2007**, *129*, 5330-5331.

(148)   Wheeler, S. E.; Moran, A.; Pieniazek, S. N.; Houk, K. N. *J. Phys. Chem. A* **2009**, *113*, 10376-10384.

(149)   Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1996**, *104*, 2598-2619.

(150)   van Bochove, M. A.; Bickelhaupt, F. M. *Eur. J. Org. Chem.* **2008**, 649-654.

(151)   Gronert, S. *J. Am. Chem. Soc.* **1993**, *115*, 10258-10266.

(152)   Gronert, S. *Organometallics* **1993**, *12*, 3805-3807.

(153)   Gronert, S.; Glaser, R.; Streitwieser, A. *J. Am. Chem. Soc.* **1989**, *111*, 3111-3117.

(154)   Schreiner, P. R.; Schleyer, P. v. R.; Schaefer III, H. F. *J. Org. Chem.* **1997**, *62*, 4216-4228.

(155)   Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334-338.

(156)   Evans, M. G.; Planyi, M. *Trans. Faraday Soc.* **1938**, *34*, 11-24.

(157)   Preliminary experimental results for S_N2 reactions of deuterated 1,3-dMU with nucleophiles are inconclusive.

(158)   3-Methyluracil rather than uracil was studied to avoid complicating deprotonation at the N3-H site; see reference 77 for details.

(159)   Walsh, C. *Enzymatic Reaction Mechanisms*; Freeman: San Francisco, 1979.

(160)   Pullman, B.; Pullman, A. *Adv. Heterocycl. Chem.* **1971**, *13*, 77-159.

(161)   Kwiatkowski, J. S.; Pullman, B. *Adv. Heterocycl. Chem.* **1975**, *18*, 199-335.

(162)   Sonnenberg, J. L.; Wong, K. F.; Voth, G. A.; Schlegel, B. *J. Chem. Theory Comput.* **2009**, *5*, 949-961 and references therein.

(163)   Beak, P. *Acc. Chem. Res.* **1977**, *10*, 186-192.

(164)   Sato, H.; Hirata, F.; Sakaki, S. *J. Phys. Chem. A* **2004**, *108*, 2097-2102.

(165)   Hatherley, L. D.; Brown, R. D.; Godfrey, P. D.; Pierlot, A. P.; Caminati, W.; Damiani, D.; Melandri, S.; Favero, L. B. *J. Phys. Chem.* **1993**, *97*, 46-51.

(166)   Nowak, M. J.; Lapinski, L.; Fulara, J.; Les, A.; Adamowicz, L. *J. Phys. Chem.* **1992**, *96*, 1562-1569.

(167)   De Kowalewski, D. G.; Contreras, R. H.; Diez, E.; Esteban, A. *Molec. Physics* **2004**, *102*, 2607-2615.

(168)   Furlong, J. J. P.; Mercedes Schiavoni, M.; Castro, E. A.; Allegretti, P. E. *Russ. J. Org. Chem.* **2008**, *44*, 1725-1736.

(169)   Cook, M. J.; Katritzky, A. R. *Tetrahedron Lett.* **1976**, *31*, 2685-2688.

(170)   Beak, P.; Fry, F. S. *J. Am. Chem. Soc.* **1973**, *95*, 1700-1702 and references therein.

(171)   Piacenza, M.; Grimme, S. *J. Comput. Chem.* **2004**, *25*, 83-99.

(172)   Fu, A.; Li, H.; Du, D.; Zhou, Z. *J. Phys. Chem. A* **2005**, *109*, 1468-1477.

(173)   Maris, A.; Ottaviani, P.; Caminati, W. *Chem. Phys. Lett.* **2002**, *360*, 155-160.

(174)   Hazra, M. K.; Chakraborty, T. *J. Phys. Chem. A* **2006**, *110*, 9130-9136.

(175)   Barone, V.; Adamo, C. *J. Phys. Chem.* **1995**, *99*, 15062-15068.

(176)   Dkhissi, A.; Houben, L.; Smets, J.; Adamowicz, L.; Maes, G. *J. Mol. Struct.* **1999**, *484*, 215-227.

(177)   Tsuchida, N.; Yamabe, S. *J. Phys. Chem. A* **2005**, *109*, 1974-1980.

(178)   Guimon, C.; Garrabe, G.; Pfister-Guillouzo, G. *Tetrahedron Lett.* **1979**, *28*, 2585-2588.

(179)   Brown, R. S.; Tse, A.; Vederas, J. C. *J. Am. Chem. Soc.* **1980**, *102*, 1174-1176.

(180)   If more than one value is listed for an atom, the arrows show the site of protonation (for example, the O2 of 2-pyridone can be protonated on the N1 side or the C3 side).

(181)   We also attempted to measure the acidity using the Cooks kinetic method but the signal corresponding to the protonated dimer of 2-pyridone and a series of reference bases was neither strong nor stable enough to conduct the experiment.

(182)   Eyet, N.; Villano, S. M.; Bierbaum, V. M. *Int. J. Mass Spectrom.* **2009**, *283*, 26-29.

(183)   Based on calculations, enol **1b'** is 5 kcal mol^{-1} less stable than enol **1b**, and is thus unlikely to be present in any significant quantity.

(184)   There is evidence that some structures when electrosprayed can isomerize in the desolvation process, though this has not been shown for pyridones: Tian, Z.; Kass, S. R. J. Am. Chem. Soc. 2008, 130, 10842-10843 and references therein. .

(185)   It is also possible that the protonated dimer is a mixture of structures, although because the PA measured by Cooks is comparable to that obtained by bracketing, the likelihood is that the major, if not exclusive structure is proton-bound on the C3 side of the keto.

(186)   Ritchie, C. D.; Sager, W. F. *Prog. Phys. Org. Chem.* **1964**, *2*, 323-400.

(187)   Hansch, C.; Leo, A.; Unger, S.; Kim, K. H.; Nikaitani, D. *J. Med. Chem.* **1973**, *16*, 1207-1216.

(188)   Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley-Interscience: New York, 1979.

(189)   Samson, L.; Cairns, J. *Nature* **1977**, *267*, 281-283.

(190)   Evensen, G.; Seeberg, E. *Nature* **1982**, *296*, 773-775.

(191)   Nakabeppu, Y.; Miyata, T.; Kondo, H.; Iwanaga, S.; Sekiguchi, M. *J. Biol. Chem.* **1984**, *269*, 13730-13736.

(192)   Lindahl, T. *Nature* **1993**, *362*, 709-715.

(193)   Bjelland, S.; Bjoras, M.; Seeberg, E. *Nucleic Acids Res.* **1993**, *21*, 2045-2049.

(194)   Saparbaev, M.; Laval, J. *Proc. Natl Acad. Sci.* **1994**, *91*, 5873-5877.

(195)   Saparbaev, M.; Kleibl, K.; Laval, J. *Nucleic Acids Res.* **1995**, *23*, 3750-3755.

(196)   Masaoka, A.; Terato, H.; Kobayashi, M.; Honsho, A.; Ohyama, Y.; Ide, H. *J. Biol. Chem.* **1999**, *274*, 25136-25143.

(197)   Terato, H.; Masaoka, A.; Asagoshi, K.; Honsho, A.; Ohyama, Y.; Suzuki, T.; Yamada, M.; Makino, K.; Yamamoto, K.; Ide, H. *Nucleic Acids Res.* **2002**, *30*, 4975-4984.

(198)   Sun, X.; Lee, J. K. *J. Org. Chem.* **2010**, *75*, 1848-1854.

(199)   Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2007**, *111*, 408-422.

(200)   We did not measure the properties of 3-methylguanine; it is neither readily available commercially nor simple to synthesize.

(201)   Huang, Y.; Liu, L.; Liu, S. *Chem. Phys. Lett.* **2012**, *527*, 73-78.

(202)   B3LYP/6-31+G(d) method underestimates PA of another purine damaged base, xanthine, by 6-7 kcal mol^{-1}.

(203)   Guliaev, A. B.; Singer, B.; Hang, B. *DNA Repair* **2004**, *3*, 1311-1321.

(204)   Bjelland, S.; Birkeland, N.-K.; Benneche, T.; Volden, G.; Seeberg, E. *J. Biol. Chem.* **1994**, *269*, 30489-30495.

(205)   Hollis, T.; Lau, A.; Ellenberger, T. *Mutation Research* **2000**, *460*, 201-210.

(206)   Berdal, K. G.; Johanson, R. F.; Seeberg, E. *EMBO J.* **1998**, *17*, 363-367.

(207)   Habraken, Y.; Ludlum, D. B. *Carcinogenesis* **1989**, *10*, 489-492.

(208)   Habraken, Y.; Carter, C. A.; Kirk, M. C.; Ludlum, D. B. *Cancer Research* **1991**, *51*, 499-503.

(209)   McCarthy, T. V.; Karran, P.; Lindahl, T. *EMBO J.* **1984**, *3*, 545-550.

(210)   It is also possible that the neutral nucleobases are protonated before cleavage; for AlkA, the mechanism is not known but for other glycosylases (uracil DNA glycosylase, thymine DNA glycosylase), the leaving group is the anionic deprotonated nucleobase.

(211)   The rate constant for excision can also be affected by what nucleobase is base paired with the base being cleaved. For Table 4.7, all the nucleobases are paired with mismatches except for hypoxanthine.

(212)   Rogstad, K. N.; Jang, Y. H.; Sowers, L. C.; Goddard, W. A. *Chem. Res.Toxicol.* **2003**, *16*, 1455-1462.

(213)   Bendich, A.; Russell, P. J.; Fox, J. J. *J. Am. Chem. Soc.* **1954**, *76*, 6073-6077.

(214)   Albert, A.; Brown, D. J. *J. Chem. Soc.* **1954**, 2060-2071.

(215)   Milletti, F.; Storchi, L.; Goracci, L.; Bendels, S.; Wagner, B.; Kansy, M.; Cruciani, G. *Eur. J. Med. Chem.* **2010**, *45*, 4270-4279.

(216)   Langman, S. R.; Shohoji, M. C. B. L.; Telo, J. P.; Vieira, A. J. S. C.; Novais, H. M. *J. Chem. Soc., Perkin Trans.* **1996**, *2*, 1461-1465.

(217)   Jang, Y. H.; Goddard III, W. A.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, S.; Chung, D. S. *J. Phys. Chem. B* **2003**, *107*, 344-357.

(218)   Taylor, H. F. W. *J. Chem. Soc.* **1948**, 765-766.

(219)   Lindahl, T. *Nature* **1993**, *362*, 709-715.

(220)   Klaunig, J. E.; Kamendulis, L. M. *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 239-267.

(221)   Shigenaga, M. K.; Park, J.-W.; Cundy, K. C.; Gimeno, C. J.; Ames, B. N. *Methods Enzymol.* **1990**, *186,*, 521-530.

(222)   Dizdaroglu, M. *Biochemistry* **1985**, *24*, 4476-4481.

(223)   Cunningham, R. P. *Mutat. Res.* **1997**, *383*, 189-196.

(224)   Michaels, M. L.; Tchou, J.; Grollman, A. P.; Miller, J. H. *Biochemistry* **1992**, *31*, 10964-10968.

(225)   Michaels, M. L.; Miller, J. H. *J. Bacteriol.* **1992**, *174*, 6321-6325.

(226)   Michaels, M. L.; Cruz, C.; Grollman, A. P.; Miller, J. H. *Proc. Natl. Acad. Sci.* **1992**, *89,* 7022-7025.

(227)   Francis, A. W.; Helquist, S. A.; Kool, E. T.; David, S. S. *J. Am. Chem. Soc.* **2003**, *125*, 16235-16242.

(228)   Livingston, A. L.; O'Shea, V. L.; Kim, T.; Kool, E. T.; David, S. S. *Nature Chem. Biol.* **2008**, *4*, 51-58.

(229)   Chepanoske, C. L.; Langelier, C. R.; Chmiel, N. H.; David, S. S. *Org. Lett.* **2000**, *2*, 1341-1344.

(230)   David, S. S. *Personal Communication* **2009**.

(231)   Porello, S. L.; Williams, S. D.; Kuhn, H.; Michaels, M. L.; David, S. S. *J. Am. Chem. Soc.* **1996**, *118*, 10684-10692.

(232)   Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. *J. Phys. Chem. A* **1998**, *102*, 7787-7794.

(233)   Chipman, D. M. *J. Phys. Chem. A* **2002**, *106*, 7413-7422.

(234)   Podolyan, Y.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7346-7352 and references therein.

(235)   Hanus, M.; Kabelac, M.; Rejnek, J.; Ryjacek, F.; Hobza, P. *J. Phys. Chem. B* **2004**, *208*, 2087-2097.

(236)   Russo, N.; Toscano, M.; Grand, A.; Jolibois, F. *J. Comput. Chem.* **1998**, *19*, 989-1000.

(237)   Huang, Y.; Kenttämaa, H. *J. Phys. Chem. A* **2004**, *108*, 4485-4490.

(238)   Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811-6821 and references therein.

(239)   Chandra, A. K.; Nguyen, M. T.; Uchimaru, T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1999**, *103*, 8853-8860.

(240)   Del Bene, J. E. *J. Phys. Chem.* **1983**, *87*, 367-371.

(241)   The tilde (~) is used when two values are within 0.5 kcal mol^{-1} of each other.

(242)   Chen, E. C. M.; Herder, C.; Chen, E. S. *J. Mol. Struct.* **2006**, *798*, 126-133.

(243)   In terms of affinity for MutY, adenine, Z1, Z3, and M all have comparable affinity to MutY. B and Q appear to be roughly nine-fold less in terms of binding affinity versus adenine; Z is roughly twenty-fold less (refs 227-231).

(244)   Fromme, J. C.; Banerjee, A.; Huang, S. J.; Verdine, G. L. *Nature* **2004**, *427*, 652-656.

(245)   Guan, Y.; Manuel, R. C.; Arvai, A. S.; Parikh, S. S.; Mol, C. D.; Miller, J. H.; Lloyd, R. S.; Tainer, J. A. *Nat. Str. Biol.* **1998**, *5*, 1058-1064.

Curriculum Vitae

Anna Michelson

September 1998 – July 2003	Moscow State University, Moscow, Russia Subject: Chemistry Degree Earned: B.S., M.S.
September 2005 – October 2012	Rutgers, the State University of New Jersey New Brunswick, NJ Subject: Chemistry Degree Earned: Ph.D.

Publications

Michelson, A. Z., Chen, M., Wang, K., Lee, J. K. *J. Am. Chem. Soc.* **2012,** *134*, 9622–9633. "Gas Phase Studies of Purine 3-Methyladenine DNA Glycosylase II (AlkA) Substrates."

Michelson, A. Z., Petronico, A., Lee, J. K. *J. Org. Chem.* **2012,** *77,* 1623-1631. ("Featured Article" and also FEATURED ON THE COVER). "2-Pyridone and Derivatives: Gas-Phase Acidity, Proton Affinity, Tautomer Preference, and Leaving Group Ability."

Zhachkina, A., Lee, J. K. *J. Am. Chem. Soc.* **2009,** *131*, 18376–18385. "Uracil and Thymine Reactivity in the Gas Phase: The S_N2 Reaction and Implications for Electron Delocalization in Leaving Groups."

Zhachkina, A., Liu, M., Sun, X., Amegayibor, F. S., Lee, J. K. *J. Org. Chem.* **2009,** *74*, 7429– 7440. "Gas-Phase Thermochemical Properties of the Damaged Base O6-Methylguanine versus Adenine and Guanine."

Selina, A. A., **Zhachkina, A.,** Karlov, S. S., Churakov, A. V., Zaitseva, G. S. *Heteroatom Chem.* **2004,** 15, 169–174. "Iodochlorination of Silyl- and Germylphenylacetylenes."