Ground State Destabilization in Uracil DNA Deglycosylase (UDG): Let’s Not Forget “Tautomeric Strain” in Substrates

Ranjita Das,‡ Erik A. Vázquez-Montelongo,‡ G. Andrés Cisneros‡ and Judy I. Wu†,*

†Department of Chemistry, University of Houston, Houston, Texas 77204, United States
‡Department of Chemistry, University of North Texas, Denton, Texas 76201, United States

Supporting Information Placeholder

KEYWORDS: Uracil DNA deglycosylase, ground state destabilization, substrate strain, uracil, tautomers, aromaticity

ABSTRACT: Enzymes like uracil DNA deglycosylase (UDG) can achieve ground state destabilization, by polarizing substrates to mimic rare tautomers. Based on computed nucleus independent chemical shifts, NICS(1), and harmonic oscillator model of electron delocalization (HOMED) analyses, of QM and QM/MM models of the UDG active site, uracil is strongly polarized when bound to UDG, and resembles a tautomer > 12 kcal/mol higher in energy. Natural resonance theory (NRT) analyses identified a dominant O2 imidate resonance form for residue bound 1-methyl-uracil. This “tautomeric strain” raises the energy of uracil, making uracilate a better than expected leaving group. Computed gas-phase S2 reactions of free and hydrogen bonded 1-methyl-uracil demonstrate the relationship between degree of polarization in uracil and leaving group ability of uracilate.

In this communication, we present computational evidence showing that some enzymes can polarize substrates to mimic rare tautomers, and in this way, achieve ground state destabilization through substrate strain. In the active site of uracil DNA deglycosylase (UDG), uracil is polarized and the resulting structure bears resemblance to a tautomer form that is much higher in energy.

UDG is a base excision repair enzyme that removes mutagenic uracil from DNA by cleaving the glycosidic N1–C1´ bond of uridine. Among the family of base excision repair enzymes, UDG is the most efficient and well characterized, showing an enormous catalytic rate acceleration of 1012. Studies based on NMR, Raman, kinetic isotope effect measurements, and quantum mechanic/molecular mechanics (QM/MM) computations, established a remarkable stepwise dissociative mechanism, involving uracilate as the leaving group5–17 (Figure 1a).

But a puzzle regarding UDG is how it activates the removal of N1 deprotonated uracilate. In water, uracil N1–H has a pKa of 9.8 (i.e., only as acidic as ammonium) and N1 deprotonated uracilate is not considered to be a very good leaving group at all. Parikh and Tainer et al. suggested that in UDG, uridine adopts a strained conformation that weakens the glycosidic N1–C1´ bond through anomic effects.10 Drohat and Süvers showed, based on NMR experiments, that uracil displayed increased acidity (pKa = 6.4) in the active site of UDG, indicating the appearance of an O2 imidate form.11,12 Wetmore et al. showed that uracil exhibited increased acidity when hydrogen bonded at the C=O and N–H sites.18,19 Nevertheless, Lee et al. pointed out that increased (thermodynamic) acidity in uracil cannot be used to reason the (kinetic) leaving group ability of uracilate in UDG.20–22 In the gas-phase, uracil N1–H is as acidic as HCl, but the leaving group ability of N1 deprotonated uracilate is much poorer compared to Cl+.21

Figure 1. a) Proposed stepwise dissociative pathway for UDG. b) diketo (1a) and (1b-1f) imidate resonance forms of uracil.

Here we show that when uracil is bound to UDG, it does not exist in the canonical diketo form (1), but instead transforms to a structure akin to a rare tautomer of uracil.23–28 In UDG, uracil forms short hydrogen bonds to several nearby residues (His268, Asn204, Gln144, and the backbone of Gln144). These hydrogen bonding interactions polarize the π-electrons of uracil, increase cyclic [4π+2] π-electron delocalization in the ring (see resonance forms 1b-1f, in Figure 1b), and the resulting polarized uracil mimics a higher energy keto-enol tautomer (2 and 3, see also dienol form, 4, Figure 2). Notably, all of these tautomers are > 12 kcal/mol higher in energy than the canonical diketo form (1) (see ΔE, Figure 2). We now show that this “tautomeric strain” raises the energy of uracil in UDG, bringing it closer to the transition state structure for N1–C1´ bond cleavage. In this way, ground state destabilization helps reduce the activation barrier to expel uracilate.
According to computed nucleus independent chemical shifts,29,30 NICS(1)zz uracil, 1, is essentially non-aromatic (−2.5 ppm), and solvation in water, based on either implicit (−3.9 ppm, employing the IEF-PCM approach) or explicit (−3.3 ppm, including three water molecules) models, increases its aromaticity negligibly. This implies that solvating naked uracil in aqueous solution has a minor effect on its tautomeric form. But when uracil is bound to the active site of UDG, it becomes markedly aromatic. NICS(1)zz values computed for uracil in truncated QM (−7.9 ppm) and QM/MM (−12.9 ppm) models of the UDG active site document significant aromaticity gain, suggesting close resemblance of the electronic structure of the substrate to the rare tautomers of uracil. Despite being higher in energy, 2 (−8.2 ppm), 3 (−7.9 ppm), 4 (−19.9 ppm) are all more aromatic than 1 (−2.5 ppm) (Figure 2). Differences in the estimated aromaticity gain of uracil based on QM and QM/MM models suggest that uracil is “aromatized” not only by nearby hydrogen bonding residues, but also by other long-range electrostatic interactions in the active site of UDG. See description of truncated QM and QM/MM models in Methods.

Harmonic oscillator model of electron delocalization (HOMED) values,31 a geometric index for aromaticity, indicate significant aromaticity gain for uracil 1 (0.702) when its geometry is considered in truncated QM (0.789) and QM/MM (0.829) models of the UDG active site (cf. HOMED values for 2: 0.774, 3: 0.800, 4: 0.994). HOMED values close to 1 indicate fully aromatic rings. This “aromatizing” effect brings the uracil substrate closer in geometry to the N1 deprotonated uracil (HOMED: 0.811). Natural resonance theory (NRT) analyses32–34 for 1-methyl-uracil reveal a dominant diketo resonance form 1a (31.3%), followed by smaller weights for the O2 imidate (16.5%) and O4 imidate (7.9%) forms. But when 1-methyl-uracil is hydrogen bonded to His, Asn, a fragment of protein backbone, and freely optimized, the resonance weight of 1a decreases (15.0%), while those of O2 imidate (24.1%) and O4 imidate (19.5%) increase. Changes in these resonance weights suggest that, in UDG, the electronic structure of polarized uracil resembles its keto-enol tautomers (2 and 3), which are 12-23 kcal/mol higher in energy. We note that the dominant O2 imidate form of hydrogen bonded 1-methyl-uracil agrees with Drohat and Stivers’ finding of increased acidity of uracil due to appearance of an O2 imidate.

Remarkably, the majority of aromaticity gain in uracil happens as soon as it is transferred from a non-enzymic environment (i.e., in water) into the UDG active site. From there, the geometry and electronic distribution of uracil already resembles a higher energy keto-enol tautomer, and the incipient uracilate is set up to be a better than expected leaving group. Computed NICS(1)zz for uracil and uracilate rings at stationary points along the computed reaction potential energy surface reveals that uracil continues to gain aromaticity throughout the stepwise dissociative reaction: RC (−7.9 ppm), TS1 (−9.4 ppm), INT (−10.8 ppm), TS2 (−11.3 ppm), PD (−11.2 ppm) (see Figure 1 and Figure S1 in the Supporting Information, SI). All stationary points were computed based on a constrained and truncated QM model of the UDG active site and the relative energies closely follow prior QM/MM studies.15

Computed gas-phase S0/2 reactions for free and hydrogen bonded 1-methyl-uracil (with formate as the nucleophile), suggest that increased aromaticity in 1-methyl-uracil enhances the leaving group ability of uracilate. In the absence of hydrogen bonding interactions, the activation free energy barrier for N–CH2 bond cleavage is (ΔG‡ = 37.8 kcal/mol), but the barrier reduces when 1-methyl-uracil is hydrogen bonded to three waters (35.7 kcal/mol) and even more so when 1-methyl-uracil is hydrogen bonded to two zwitterionic glycines (33.3 kcal/mol) (see Figure S2 in the SI). Computed NICS(1)zz values show that these changes correlate to increased aromaticity in 1-methyl-uracil in the reaction complex (free 1-methyl-uracil: −4.0 ppm, bound to three waters: −5.0 ppm, bound to two glycines: −5.8 ppm). As the uracil ring becomes structurally more like its rare keto-enol tautomers, uracilate becomes a better leaving group.

Tautomeric strain is one way to raise the energy of substrates in enzymes, and recognizing this mode of ground state destabilization has important interpretive merit for understanding how enzymes work. Although examples of electronic strain in substrates have been reported for various enzymes, for example, substrates with polarized C=O, C≡C, and C–H bonds or red-shifted UV absorptions,46 the possibility and impact of substrate tautomerization has not been fully appreciated. Other enzymes that catalyze the glycosidic hydrolysis or biosynthesis of purine or pyrimidine substrates may take advantage of a similar mechanistic trick.

Computational Methods

All quantum mechanical (QM) geometries were optimized at wB97X-D/6-31+G(d) in the gas-phase employing Gaussi-
Vibrational frequency analyses verified the nature of the stationary points. Computed nucleus independent chemical shifts (NICS),
harmonic oscillator model of electron delocalization (HOMED), and natural resonance theory (NRT) data were performed at the same level. Computations in implicit solvation employed the IEF-PCM model.

A QM model of the UDG active site was computed based on a modified, constrained, and geometry optimized crystal structure of an UDG-inhibitor complex (PDBID: 1EMH). In the pdb file, the positions of atoms C1 and N5 in the deoxyscytidine substrate were exchanged to give a deoxurycidin substrate containing a N1–C1′ glycosidic bond. Residues and protein backbone fragments that formed hydrogen bonding (His268, Asn204, Gln144, and the backbone of Glu144) and π-stacking (Phe158) interactions to uracil, along with those relevant for activating (His148 and Asp145) the nucleophilic water molecule, were included to the model and truncated at the β carbon position. Prior QM/MM studies document the important roles of these residues in UDG. The 5′ phosphate group was included because of its recognized importance for facilitating the dissociation of uracilate. The positions of selected atoms were frozen (see green dots in Figure 3) during geometry optimization of the RC, TS1, INT, TS2, and PD. Following prior QM/MM studies, bond distances for N1–C1′ were fixed to 2.04 Å (for TS1), 2.74 Å (for INT), and 2.95 Å (for TS2), and distances between C1′ to the O atom of water were fixed to 3.04 Å (for TS1), 2.85 Å (for INT), and 1.95 Å (for TS2).

All QM/MM calculations were performed using the LICHEM software package, which employs Gaussian09 for the quantum region (QM) and TINKER8 for the classical region (MM). The crystal structure of an UDG-inhibitor complex was not solvated for QM/MM calculations. The QM region (UDG: Gln144, Asp145, Pro146, Tyr147, His148, Phe158, Asn204, His268, deoxycyuridine, and one water molecule) was calculated at 6-31+G(d). The QM/MM system involved the cleavage of covalent bonds and thus a pseudo-bond approach was employed. Two QM/MM simulation systems (see Figures S3 and S4 in the SI) were created based on the crystal structure of an UDG-inhibitor complex: one to performed QM/MM geometry optimization (126 QM atoms + 6 psuedobonds) and the other for QM/MM single-point calculation (154 QM atoms + 8 psuedobonds). The MM region was modeled using the AMOEBABIO18 force field.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author
* jiwu@central.uh.edu

Author Contributions
The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources
National Science Foundation CHE-1751370, MRI-1531814, CHE-1531468

ACKNOWLEDGMENT

J.I.W. thanks the National Science Foundation (NSF) for support (CHE-1751370), as well as computational resources provided by the uHPC cluster managed by the University of Houston and acquired through support from the NSF (MRI-1531814). G.A.C. thanks the UNT Department of Chemistry for the use of the HPC Cluster CRUNTCr3 and the NSF for support (CHE-1531468). E.A.V.M. wishes to acknowledge CONACyT for funding. We thank Dr. Chia-Hua Wu for helpful discussions.

REFERENCES

14. Dong, J.; Drohat, A. C.; Stivers, J. T.; Pankiewicz, K. W.; Carey, P. R. Raman spectroscopy of uracil DNA glycosylase-DNA complexes: insights into DNA damage recognition and catal-
36. Gaussian 09, Revision D. 01, Gaussian, Inc., Wallingford CT, 2013. See full citation in the SI.  