

Systematic Assessment of the Flexibility of Uracil Damaged DNA

Paul B. Orndorff and Arjan van der Vaart*

Department of Chemistry, University of South Florida, 4202 East Fowler Avenue CHE 205,
Tampa, Florida 33620, United States.

Abstract

Uracil is a common DNA lesion which is recognized and removed by uracil DNA-glycosylase (UDG) as a part of the base excision repair pathway. Excision proceeds by base flipping, and UDG efficiency is thought to depend on the ease of deformability of the bases neighboring the lesion. We used molecular dynamics simulations to assess the flexibility of a large library of dsDNA strands, containing all tetranucleotide motifs with U:A, U:G, T:A or C:G base pairs. Our study demonstrates that uracil damaged DNA largely follows trends in flexibility of undamaged DNA. Measured bending persistence lengths, groove widths, step parameters and base flipping propensities demonstrate that uracil increases the flexibility of DNA, and that U:G base paired strands are more flexible than U:A strands. Certain sequence contexts are more deformable than others, with a key role for the 3' base next to uracil. Flexibilities are large when this base is an A or G, and repressed for a C or T. A 5' T adjacent to the uracil strongly promotes flexibility, but other 5' bases are less influential. DNA bending is correlated to step deformations and base flipping, and bending aids flipping. Our study implies that the link between substrate flexibility and UDG efficiency is widely valid, helps explain why UDG prefers to bind U:G base paired strands, and suggests that the DNA bending angle of the UDG-substrate complex is optimal for base flipping.

Keywords: uracil, DNA damage, uracil DNA glycosylase, simulation, molecular dynamics

* Corresponding author. Email: avandervaart@usf.edu

Introduction

DNA undergoes thousands of daily spontaneous lesions that can lead to mutations in the genome (Hoeijmakers, 2009). A common lesion is uracil, which arises because of misincorporation or through spontaneous deamination of cytosine (Krokan, Drablos, & Slupphaug, 2002). The presence of uracil from misincorporation produces U:A base pairs, which can lead to strand breakage (Williams, Lujan, & Kunkel, 2016). Uracil resulting from cytosine deamination results in U:G mismatches, which, if left unrepaired, are 100% mutagenic (Visnes et al., 2009). Uracil is removed from the genome by uracil-DNA glycosylase (UDG), which uses a base flipping mechanism to reorient uracil extrahelical for excision (Brooks, Adhikary, Robinson, & Eichman, 2013; Sudip S. Parikh et al., 2000; James T. Stivers & Jiang, 2003). UDG binds U:G base paired strands better than U:A strands (Panayotou, Brown, Barlow, Pearl, & Savva, 1998) and has higher excision rates for U:G than U:A (S. S. Parikh et al., 1998), although the preference is not absolute and also depends on sequence (Bellamy & Baldwin, 2001; Nilsen, Yazdankhah, Eftedal, & Krokan, 1995; G. Slupphaug et al., 1995). While the initial steps of damage recognition remain unclear, the dynamics and deformability of the lesion site are thought to trigger UDG detection rather than its structure. Uracil is structurally similar to thymine; while it formed a wobble base pair with G, it did not appear to modify the DNA structure in NMR studies (Carbonnaux, Fazakerley, & Sowers, 1990). However, NMR experiments showed that U:A base pairs open faster than T:A base pairs (Parker & Stivers, 2011), and lower barriers for flipping of uracil relative to cytosine (Fuxreiter, Luo, Jedlovsky, Simon, & Osman, 2002) and thymine (Fadda & Pomès, 2011) were observed in computer simulations. The importance of uracil's flexibility was also shown in simulations of UDG-assisted base flipping (Franco, Sgrignani, Bussi, & Magistrato, 2013; Priyakumar & MacKerell, 2006).

UDG catalyzed excision of uracil requires hydrogen bond breakage, destabilization of base stacking interactions, and increased bending and twisting of the DNA, and it is thought that DNA is locally destabilized before base flipping (J. T. Stivers, Pankiewicz, & Watanabe, 1999). Since deformability of DNA is strongly dependent on sequence (Hagerman, 1988; Marin-Gonzalez, Vilhena, Perez, & Moreno-Herrero, 2021; Peters & Maher, 2010), a dependence of the efficiency of repair on the sequence around the lesion is to be expected. Rates were shown to be dependent on the sequence surrounding the lesion, with poor repair rates when the uracil is directly adjacent

to a 3' thymine or when the GC content is high near the lesion (Bellamy & Baldwin, 2001; Eftedal, Guddal, Slupphaug, Volden, & Krokan, 1993; Hölz, Pavlic, Lietard, & Somoza, 2019; Nilsen et al., 1995; G. Slupphaug et al., 1995). Rates for substrates with TUA motifs were markedly higher than for substrates with AUT motifs (Eftedal et al., 1993; Nilsen et al., 1995). A recent study confirmed this observation, by quantifying k_{cat}/K_m for various U:A base paired sequences with AUT, TUA, AUA and TUT motifs, and measuring flexibilities of the strands by a combination of time-resolved fluorescence, NMR imino proton exchange, and molecular dynamics (MD) simulations (Orndorff et al., 2023). It was shown that TUA strands are significantly more flexible in terms of DNA bending and base flipping than AUT strands. Moreover, it was demonstrated that for the tested sequences, substrate flexibility controlled repair efficiency, with faster excision for more flexible strands. More bendable sequences were shown to flip more (Orndorff et al., 2023); this coupling between bending and base flipping was also seen in simulations of UDG (Franco et al., 2013) and other systems (Ning Ma & Arjan van der Vaart, 2017; Ramstein & Lavery, 1988), and may help explain why DNA is bent by $33 \pm 8^\circ$ in crystal structures of UDG (Bianchet et al., 2003; Burmeister et al., 2015; Earl et al., 2018; H. Kosaka, Hoseki, Nakagawa, Kuramitsu, & Masui, 2007; H. Kosaka, Nakagawa, N., Masui, R., Kuramitsu, S., Hoseki, J., 2007; S. S. Parikh et al., 1998; Sudip S. Parikh et al., 2000; Parker et al., 2007; Pedersen, Johnson, McVey, Leiros, & Moe, 2015; Geir Slupphaug et al., 1996).

Given the strong link between DNA flexibility and UDG repair efficiency, we set out to use MD to systematically assess the inherent flexibility of uracil containing DNA as a function of sequence. Since nearest neighbors contribute most to the flexibility of base steps in undamaged DNA (D. L. Beveridge et al., 2004; David L. Beveridge, Cheatham, & Mezei, 2012; Dixit et al., 2005; Lavery et al., 2010), we focused on uracil containing tetranucleotide motifs. Statistics on these motifs were increased by embedding three motifs per strand (D. L. Beveridge et al., 2004; David L. Beveridge et al., 2012; Dixit et al., 2005; Lavery et al., 2010). By simulating all 256 unique combinations, our study greatly extends previous work that focused on a few strands (Fadda & Pomès, 2011; Fuxreiter et al., 2002; Mardt et al., 2022; Orndorff et al., 2023; Peguero-Tejada & van der Vaart, 2017; Seibert, Ross, & Osman, 2003). Moreover, all U:A and U:G base paired sequences were studied, together with undamaged T:A and C:G controls. Sequence effects in the coupling between uracil base flipping and DNA bending were further quantified by free energy simulations for a

number of representative strands. Together, the simulations revealed clear trends in how the sequence around the lesion modulates the flexibility of DNA, and how the lesion modifies the sequence-dependent trends in the flexibility of undamaged DNA. When combining our data with findings from the literature, our study shows that the link between substrate flexibility and UDG efficiency can be generalized to much wider sequence contexts, explains why U:G base paired strands are better bound than U:A strands, and demonstrates that the UDG bending angle in the complex is optimal for flipping for all studied sequences.

Methods

The flexibility of uracil damaged DNA was systematically studied by explicit solvent MD simulations of 256 designed dsDNA 17mers with the 5'-GCGZ(UXYZ)₃G-3' sequence. Tetranucleotides are minimal constructs that incorporate the effect of neighboring base pairs on central dinucleotide steps; by embedding three uracil-containing tetranucleotide motifs per sequence, statistics on the motions these motifs was increased (D. L. Beveridge et al., 2004; David L. Beveridge et al., 2012; Dixit et al., 2005; Lavery et al., 2010). In these sequences uracil was either hydrogen bonded to adenine or to guanine; both systems were simulated. In addition, two sets of undamaged sequences with all U replaced by T and base paired to A, and all U replaced by C and base paired to G were simulated as well. Initial coordinates were constructed in the BII conformation using 3DNA (X.-J. Lu & Olson, 2008), and each sequence was solvated in 150 mM KCl of TIP3P (Jorgensen, Chandrasekhar, Madura, Impey, & Klein, 1983) water, using a rectangular box that surrounded the DNA by at least 15 Å of solvent in each dimension. Each system was energy minimized and subsequently heated and equilibrated. Heating from 100 to 300 K was completed over 2.5 ns with a 1 kcal/mol·Å² harmonic force constant placed on all atoms. In addition, flat bottom distance restraints with a force constant of 1 kcal/mol·Å² were placed on base paired hydrogen bonds. After heating, harmonic restraints were removed over 1.2 ns, while flat bottom restraints were still in effect; the flat bottom restraints were subsequently removed over an additional 3 ns. The unrestrained systems were equilibrated for 100 ns followed by 400 ns of production. Convergence of the production runs was assessed by ensuring the cumulative average of the DNA bending angle for each sequence had plateaued to constant value within the simulation time. Root mean square deviations (rmsds) were also monitored; representative plots are shown in

Fig. S1. The simulations used the AMBER OL15 DNA force field (Galindo-Murillo et al., 2016) in which the deoxyribose parameters for U were copied from T deoxyribose. Heating and equilibration were run with Langevin dynamics in the AMBER (Case et al., 2018) program, and production was run with Langevin dynamics using OPENMM (Eastman et al., 2017). Simulations were run under constant pressure using periodic boundary conditions. SHAKE (Ryckaert, Ciccotti, & Berendsen, 1977) was applied to all covalently bonded hydrogen atoms, and long-range electrostatics were treated by the particle-mesh Ewald method (Darden, York, & Pedersen, 1993).

DNA bending angles (ϕ) and bending persistence lengths (BPLs) were calculated as a measure of global flexibilities. Bending was monitored by calculating the angle between the center of mass (COM) between each terminal residue's phosphodiester backbone and the COM of the phosphodiester backbone of the central base pair. The angle between these three coordinates was evaluated every 10 ps of the simulation. BPLs were calculated from: $BPL = -L \frac{\partial \ln P(\phi)}{\partial (1 - \cos(\phi))}$, where L is the contour length, and $P(\phi)$ the probability of observing a particular bending angle (Mazur, 2007). Contour lengths were calculated from the sum of the helical rise steps (X.-J. Lu & Olson, 2008). To account for fraying, the terminal base steps were excluded from the ϕ and BPL analyses.

Base step parameters and major and minor groove widths were calculated with 3DNA (X.-J. Lu & Olson, 2008). Spontaneous base flipping of uracil was frequently observed in many of the sequences, and quantified by monitoring pseudodihedral angles. This flipping angle was taken as the dihedral angle between the COM of the U base ring, the COM of the U backbone, the COM of the backbone of the two bases 3' to the U, and the COM of the backbone of U's complementary base (Banavali & MacKerell, 2002). Uracil was considered flipped out if the pseudodihedral angle was greater than 40° or less than -40°. Fluctuations in these local properties were monitored and correlated with base step parameters to elucidate sequence trends.

Coupling between DNA bending and base flipping was studied by two-dimensional umbrella sampling simulations (Torrie & Valleau, 1977). In these simulations, the DNA bending angle was biased from 0° to 90° and the flipping angle was biased between 0° and 180° out of the major and minor grooves. Each of these angles was biased in 15° windows with a force constant of 65 kcal/mol•rad². Simulations were run under constant pressure Langevin dynamics using OPENMM

(Eastman et al., 2017). Given the cost of these simulations, only a small subset of representative sequences were studied. The sequences used for the 2D umbrella sampling were chosen to represent flexible (TUA) and more rigid (AUT) motifs at the same nucleotide content, plus two controls (AUA and TUT); the central uracil was base paired to either a guanine or an adenine (Table 1). Each window was run until at least 100 decorrelated snapshots of the flipping angle were obtained; this required at least 2 ns per window. Decorrelation and PMF computations were performed with pymbar (Shirts & Chodera, 2008).

Table 1. DNA Sequences for umbrella sampling simulations; central motifs are bolded. Uracil was base paired to either guanine or adenine; both possibilities were simulated.

Name	Motif
TUA1	5' CGT GCA TUA ATG ACG 3'
TUT1	5' CGT GCA TUT ATG ACG 3'
AUT1	5' CGT GCA AUT ATG ACG 3'
AUA1	5' CGT GCA AUA ATG ACG 3'
TUA2	5' CGT GAA TUA GTT ACG 3'
TUT2	5' CGT GAA TUT GTT ACG 3'
AUT2	5' CGT GAA AUT GTT ACG 3'
AUA2	5' CGT GAA AUA GTT ACG 3'

Results

We performed MD simulations of all unique 5'-GCGZ(NXYZ)₃G-3' dsDNA 17-mers, where N represents either a U:A, U:G, T:A, or C:G base pair. Uracil-containing strands will be referred to as damaged DNA; strands without uracil as undamaged DNA. The uracils in each of the tetramer nucleotide motifs behaved more or less independently, as shown by the distribution of Pearson correlation coefficients in Fig. S2. Bending persistence lengths (BPLs) were calculated for each strand. In Fig. 1a these BPLs are averaged over strands with the ZN or NX motifs. Fig. 1b shows the differences of these averages between the undamaged T:A and damaged U:A strands, and the undamaged C:G and damaged U:G strands. BPLs for undamaged DNA are in line with literature values and known trends; for example, GC rich sequences displayed larger BPLs than AT rich sequences (Baumann, Smith, Bloomfield, & Bustamante, 1997; Brunet et al., 2015; Bustamante, Bryant, & Smith, 2003; Herrero-Galán et al., 2013; Lipfert, Kerssemakers, Jager, & Dekker, 2010; Y. Lu, Weers, & Stellwagen, 2002). Values are also consistent with BPLs of other sequences calculated using the OL15 force field (Chhetri, Sharma, Naskar, & Maiti, 2022; Velasco-Berrelleza et al., 2020). All undamaged sequences had significantly larger BPLs than the damaged strands, indicating that the uracil lesion increases the flexibility of DNA. The BPLs of U:G strands were significantly lower than the BPLs of U:A strands; moreover, differences between C:G and U:G strands were larger than differences between T:A and U:A strands. These findings indicate that the lesion has a significant effect on DNA stiffness, and that sequences containing U:G are more easily bent than those containing U:A.

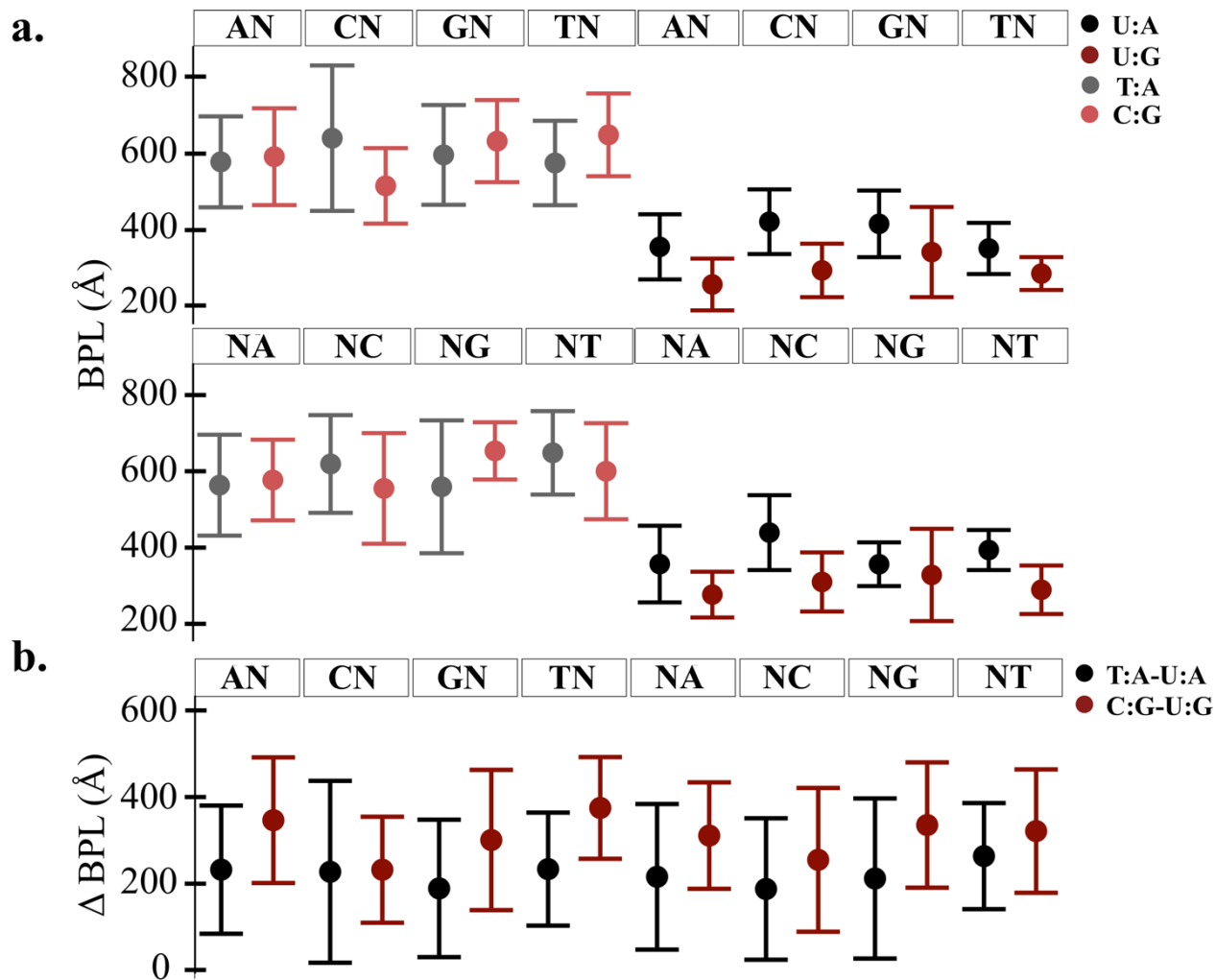


Figure 1: BPLs for damaged and undamaged DNA (a), and their difference in values (b). Values are averaged over all sequences with NX or ZN contexts, and error bars represent standard deviations.

For the U:A and T:A strands, sequences where uracil (for U:A) or thymine (for T:A) is followed by a 3' C or T display larger BPLs than those with a 3' A or G (NX motifs in Fig. 1a); moreover, a 5' neighboring T displays the lowest BPL. In U:G and C:G strands, the largest BPLs occur in sequences where uracil (U:G) or cytosine (C:G) neighbors a 3' G; for C:G strands sequences with CN steps are most flexible. No other trends were observed for the C:G and U:G sequences. Fig. 1b shows that differences in BPLs between the undamaged and damaged strands depend stronger on sequence context for the U:G damage than for the U:A damage. For U:A the average difference is 220.7 ± 25.0 Å for all contexts, except for the TN and NT contexts where the difference is 249.1

± 21.2 Å. For U:G, differences vary more over the contexts. Differences follow similar trends for its ZN and NX motifs, with the largest difference in BPLs for CN and NC motifs. Combined, these findings indicate that while uracil increases bendability of DNA, U:A damage largely follows the same trends in bending flexibility as undamaged DNA, while U:G damage deviates from these trends. Overall, these observed differences between U:A and U:G damaged strands likely stem from a difference in size and hydrogen bonding of the lesion. Whereas U readily substitutes for T in terms of hydrogen bonding and size, U:G represents a mismatch; explaining the increased flexibility and larger deviations from undamaged DNA for the U:G lesion.

Fig. 2 displays the major and minor groove widths of the NXYZ motif for damaged and undamaged sequences; values were averaged over the three NXYZ motifs of each strand. Like Fig. 1, all measurements are shown as averages over all sequences with NX or ZN contexts. Major and minor groove widths were the same for palindromic steps of undamaged DNA (Table S1), indicating convergence of the simulations. Bending toward the major groove is observed specifically around the lesion. As the major groove compresses upon bending, the minor groove expands, and this pinching/bending behavior is enhanced when U is followed by a 3' A or G (Fig. 2b) or preceded by a 5' T (Fig. 2a). The relation between groove pinching and DNA bending is also clear from the fact that for U:A, these sequences (i.e. with UA, UG, and TG motifs) have the lowest BPLs (Fig. 1) and therefore the largest bending flexibility. In these sequence contexts, the XY, YZ and ZN major groove widths gradually increase up to 2 steps away from the lesion. As the major groove width increases beyond the NX step, the minor groove decreases. The latter indicates a bend back to the minor groove up to 2 steps away from the lesion. The compression and expansion of groove widths was muted for strands in which U is neighbored by a 3' C or T or a 5' A or G, and widths for the neighboring steps did not exhibit this pinching or bending behavior around the lesion in these sequences. In the undamaged sequences, similar trends were observed, with pinching or bending behavior towards the major groove for NA, NG and TN contexts. These correspond to the U:A sequences with the lowest BPLs (Fig. 1). However, for undamaged DNA, the changes in groove width upon pinching and bending were less than for the damaged strands.

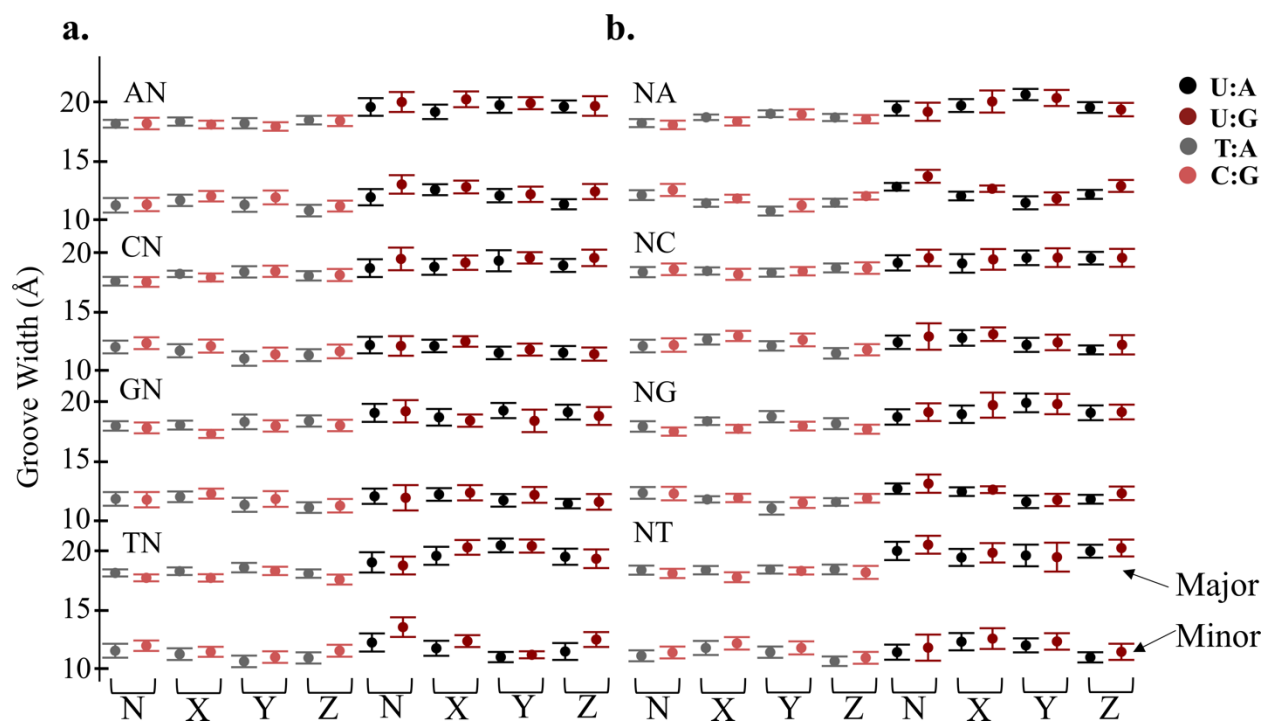


Figure 2: Groove widths averaged over NXYZ motifs and averaged for all sequences with ZN (a) and NX (b) contexts; bars indicate standard deviations.

More detail on the effect of the lesion on groove widths is captured in Fig. S3, which shows the difference in major and minor groove widths between undamaged and damaged DNA. As was done in Fig. 2, these differences were averaged over NXYZ motifs and averaged by sequence context. Uracil widened both major and minor grooves, but differences between damaged and undamaged DNA were larger for the major groove (Fig. S3). This was seen for both U:A and U:G damaged DNA, but the effect was larger in U:G strands than in U:A strands. The increased groove widths are likely due to a combination of the increased polarity of U over T and increased DNA bending. The lack of the methyl group opens the major groove for increased solvent accessibility, thus increasing the size and fluctuations in bending around the grooves. Increased dynamics around the lesion exacerbates the effect, leading to larger fluctuations in bending and enhanced motions in major and minor groove widths. U:G lesions contribute to even more motion, resulting in further increase in groove widths. Widened grooves and increased fluctuations for uracil containing sequences suggest that damaged DNA is more likely to bend back and forth whereas undamaged DNA is stiffer and more compactly coiled. These trends are most visible in TN, NA and NG steps and consistent with the trends observed in the BPL analysis of U:A base pairs. Even

though the measured BPLs for C:G and U:G did not follow these trends, the pinching behavior near the lesion was consistent in all sequences with TN, NA or NG contexts.

The effect of the lesion on local DNA structure was probed by analysis of DNA step parameters. Fig. 3 displays the average and standard deviation of the ZN and NX step parameters, while Fig. S4 shows the difference in these step parameters between undamaged and damaged DNA. Reported standard deviations are over these sequence averages; these standard deviations followed the same trend as standard deviations over all snapshots (Fig. S5) and therefore provide a measure of flexibility. A few trends are apparent for the translational shift, slide, and rise parameters. First, fluctuations in these parameters were generally larger for damaged DNA, indicating more flexibility, and largest for the U:G sequences. Second, slide was consistently lower for damaged DNA. Third, shift was lower in U:A damaged DNA than undamaged DNA, while it was larger in U:G damaged DNA, and particularly for its ZN step. Lastly, for all NX steps, rise was lower in undamaged than damaged DNA, particularly for the U:G sequences; but for the ZN step, no consistent pattern was observed. Fig. 3a and S4a also show that the difference in slide between the C:G and U:G sequences was larger for CN and TN steps. These enhanced translational parameters for U:G sequences resulted in a more open and solvent accessible state than U:A sequences.

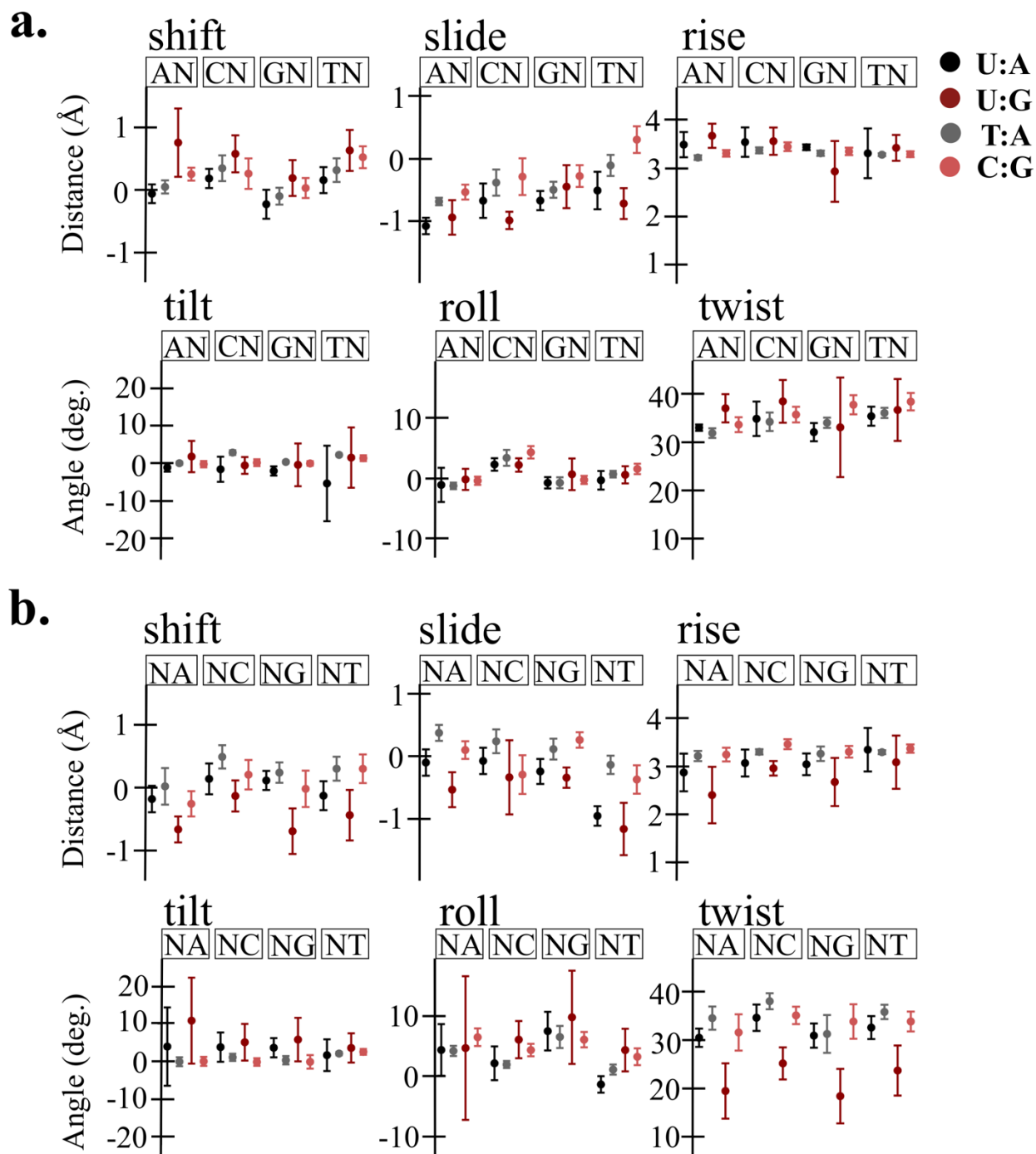


Figure 3: Step parameters averaged over all sequences with a ZN (a) or NX (b) context.

In terms of the rotational step parameters, the most important observation was a significant under twisting of the NX step of damaged DNA, particularly for the U:G sequences. Under twisting was also observed for the GN and TN steps, but the AN and CN steps were overtwisted. NC and NT steps had larger twist than NA and NG steps in both damaged and undamaged sequences, and all

damaged twist angles were lower than in undamaged stands. Fluctuations of the NX roll angle were much larger in damaged DNA than in undamaged DNA; for the ZN roll angle this difference was subdued. Roll angles for the NA and NG steps were larger than those of the NC and NT steps in both damaged and undamaged sequences. The ZN roll was largest for CN followed by TN for both damaged and undamaged DNA; the roll angles were smaller for damaged DNA in these steps. All NX tilt angles were larger in damaged sequences; ZN tilts were generally lower for damaged DNA, and the largest changes in tilt from undamaged DNA were observed for CN and TN.

Trends in step parameters mimic trends in DNA bending (Fig. 1, 3). Fluctuations in step parameters increased from undamaged DNA, to U:A, to U:G damaged DNA, and BPLs decreased in the same order. Even though all NX steps were under twisted in the damaged strands, under twisting was largest at NA and NG steps, for both damaged and undamaged DNA. Moreover, increased roll angles were observed at these steps (Fig. 3). The BPL analysis showed that sequences with NA and NG are more flexible than NC and NT contexts for both damaged and undamaged DNA in U:A and U:T sequences. Although this trend was not observed in BPLs of U:G and C:G sequences, the groove width analysis showed increased bending toward the major groove for NA and NG steps for both U:A and U:G damaged DNA and the undamaged sequences. In both damaged and undamaged DNA, little correlation was found between the ZN step parameters and BPLs. The TN step influenced groove widths, resulting in increased bending at the lesion (Fig. 2). Even though CN and TN step parameters followed the same trends, the 5' thymine appeared to have a greater effect on groove widths.

These observations suggest that the influence of the neighboring 3' base contributes to larger deformability around a lesion site than the neighboring 5' base. In fact, the rotational parameters of the NX step were more affected by the lesion than the parameters of the ZN step: averaged changes were generally higher in magnitude, and fluctuations were also larger (Fig. S4). Fig. 4 shows that BPLs were most strongly correlated to the tilt, roll and twist of the NX step, while correlations with the ZN step were weak. These findings confirm that the 3' base adjacent to U is more important in modulating local flexibilities than the 5' base.

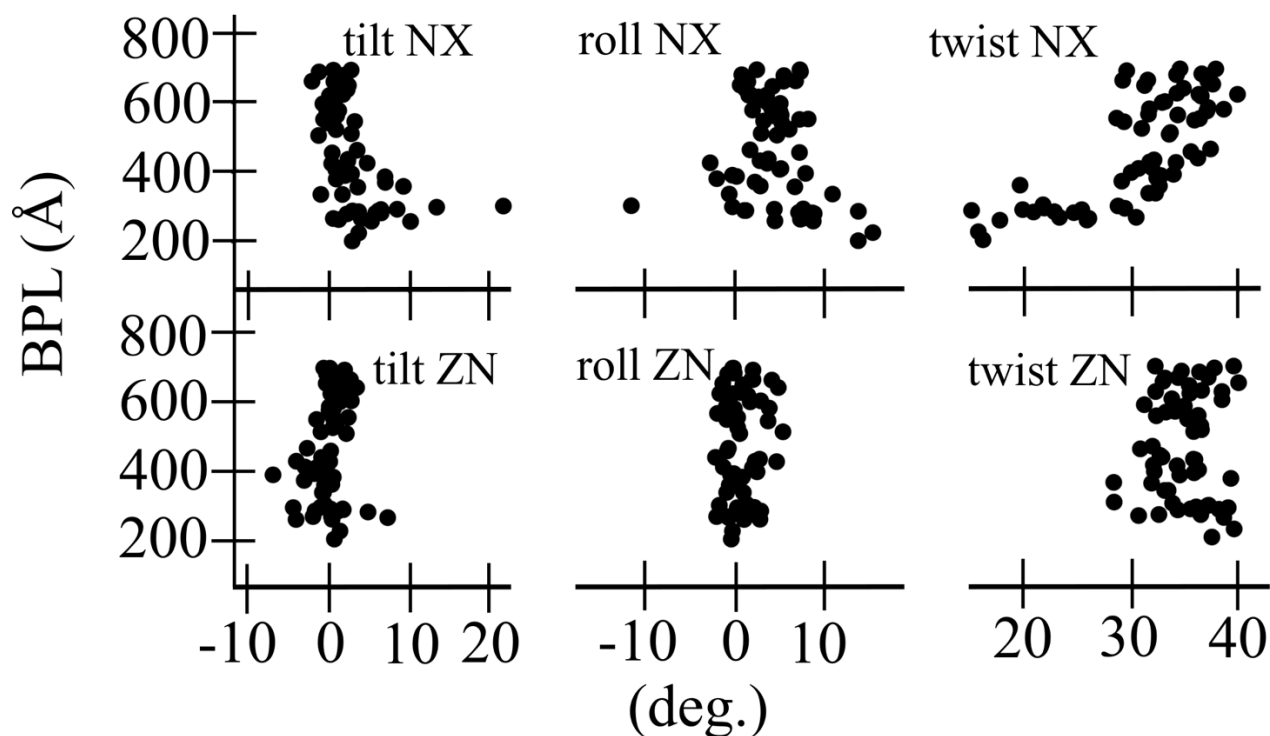


Figure 4: Correlation of BPLs with NX and ZN step parameters.

Significant base flipping of the uracil lesion was observed for the damaged strands in all sequence contexts (Table 2); most flipping occurred in the U:G sequences. Notably, sporadic spontaneous flipping of a T or C base was observed in some of the undamaged sequences. Several trends that link the propensity of base flipping to the identity of the 5' and 3' bases adjacent to the lesion emerge from Table 2. With the exception of CNG, increased flipping in undamaged DNA occurred when the central T or C was followed by a 3' purine and preceded by a 5' pyrimidine. This trend was also seen for the central U in damaged DNA, but certain sequence contexts did not fit this pattern. For example, significant base flipping was observed in ANG contexts for U:G sequences, and ANA contexts increased flipping for both U:G and U:A sequences. Apart from these exceptions, lesions with 3' purine and 5' pyrimidine neighboring bases experienced more base flipping than other contexts; furthermore, 5' T led to more flipping than 5' C.

Table 2. Percentage of time central N base is intrahelical in sequences with ZNX motifs.

Motif	U:A	T:A	U:G	C:G
ANA	93.6±6.4	100.0±0.0	85.9±2.8	100.0±0.0
ANC	99.6±0.4	100.0±0.0	92±3.8	100.0±0.0
ANG	98.5±4.2	100.0±0.0	55.6±9.1	100.0±0.0
ANT	99.1±1.1	100.0±0.0	91.4±3.6	100.0±0.0
CNA	96±5.7	99.8±0.2	95.5±1.8	99.9±0.0
CNC	96.3±6.7	100.0±0.0	96.6±2.2	100.0±0.0
CNG	98.1±2.6	99.9±0.2	87.2±3.9	100.0±0.0
CNT	99±1.1	100.0±0.0	98±1.8	100.0±0.0
GNA	99.6±0.5	100.0±0.0	93.5±4.6	100.0±0.0
GNC	99.8±0.4	100.0±0.0	98.5±1.1	100.0±0.0
GNG	98.5±3.3	100.0±0.0	90.1±4.2	100.0±0.0
GNT	99.4±1	100.0±0.0	98±0.9	100.0±0.0
TNA	81.8±21.7	99.9±0.1	85±20.4	99.9±0.0
TNC	89.8±9.9	100±0	91.9±3.5	100.0±0.1
TNG	89.1±8.6	99.9±0.1	75.2±4.5	99.9±0.1
TNT	100.0±0.0	100.0±0.0	92.8±2.3	100.0±0.0

This behavior is echoed in Fig. 5, where the fluctuations of the flipping angle and the average flipping angle are shown for ZNX contexts. Fluctuations for uracil damaged DNA were significantly greater than those for undamaged DNA (Fig. 5a). Particularly large averages and standard deviations were seen for the 5' T contexts that experienced frequent flipping (Table 2). Fig. 5b shows that the average flipping angle for 3' pyrimidines was generally larger than 3' purines. These trends were seen for both damaged and undamaged sequences, and indicate that 3' pyrimidines were better stacked and less solvent exposed. Lower flipping angles correspond to a more solvent exposed, open state, and for flipping angles below -40° to a state that is flipped toward the major groove. U:G sequences exhibited the lowest flipping angles; the average flipping angle of U:G sequences was -27° and significantly lower than for U:A sequences (-18°). In fact, throughout all U:G simulations uracil existed as a wobble base pair, in an open, more solvent

accessible state instead of being well-stacked; this opened state helped facilitate base flipping in U:G sequences. Flipping towards the major groove was more prevalent than flipping towards the minor groove; but when a base flipped completely extra helically from the major groove side, it could start interacting with the minor groove. However, all flipping started from the major groove. Little to no base flipping was observed in undamaged sequences; these sequences exhibited low flipping angle fluctuations ($\sim 5^\circ$) and large flipping angles (Fig. 5). Overall, these trends in flipping and flipping motions largely matched trends in BPLs, groove widths and step parameters, with more frequent flipping for more flexible sequences, particularly for sequences with a 3' purine and 5' pyrimidine adjacent to the flipped base. The link between flipping and flexibility is further supported by Fig. 6, which shows that the flipping fluctuations and flipping angle (Fig. 5) are strongly correlated to the BPLs (Fig. 1).

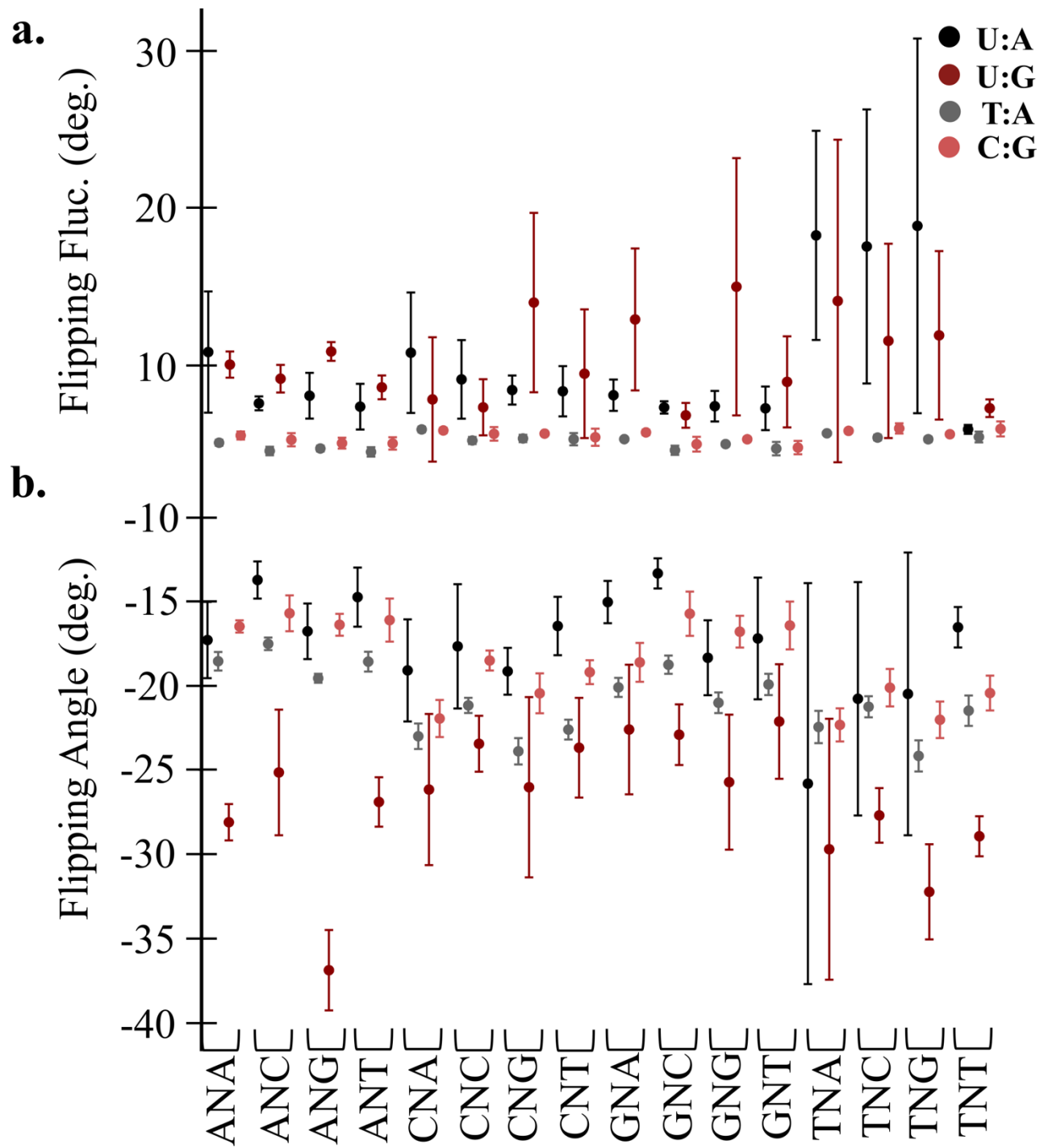


Figure 5: Base flipping motions of central N base, averaged over all sequences with a ZNX context. a) Fluctuations in the flipping angle, b) average flipping angle.

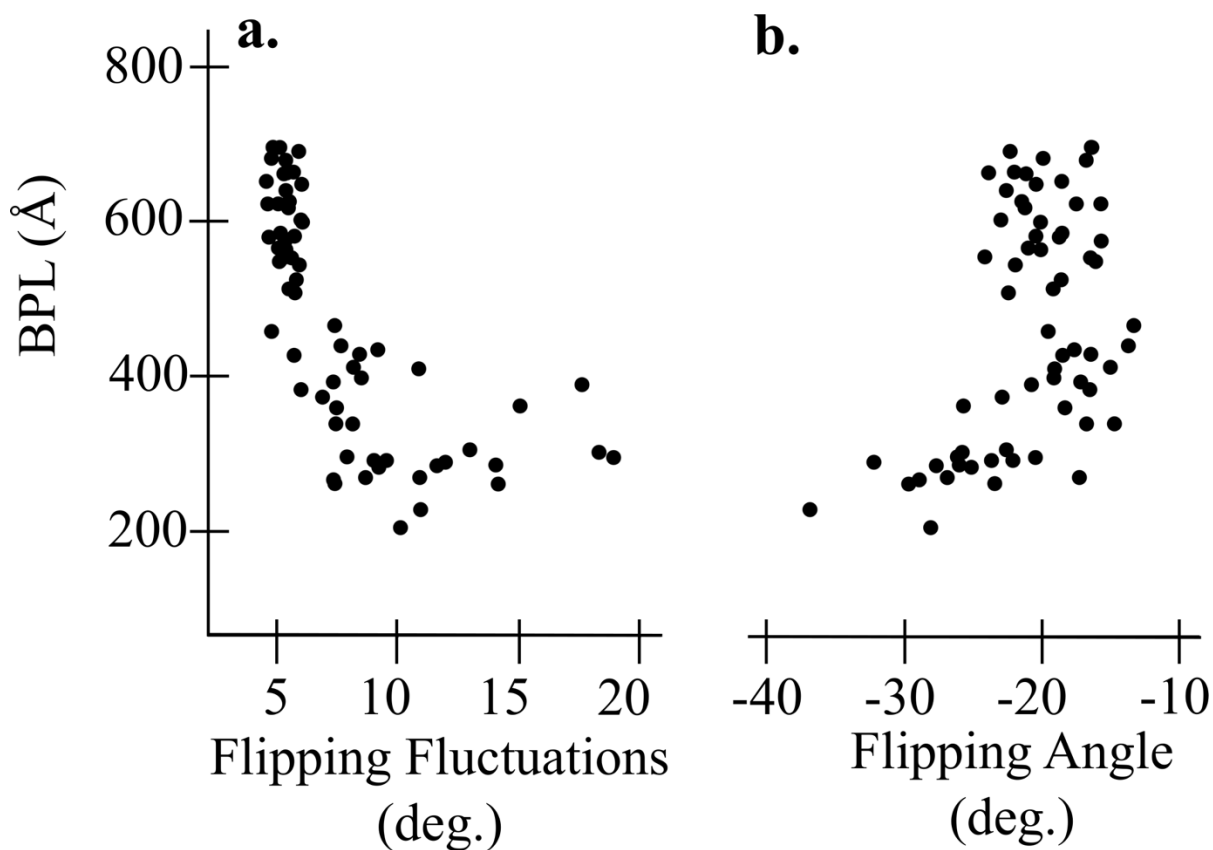


Figure 6: Correlation of uracil base flipping fluctuations (a) and flipping angle values (b) with BPLs. Each dot represents a sequence in a ZNX motif for damaged and undamaged DNA.

DNA bending and base flipping were favored toward the major groove in all sequences studied. To further quantify the coupling between bending and flipping, PMFs as a function of the bending and flipping angle were calculated from umbrella sampling simulations (Fig. 7). Due to the cost, these calculations were done for a few representative sequences (Table 1): U:A and U:G base paired sequences with the flexible TUA and rigid AUT motifs and two controls. Only bending toward the major groove was considered: this was the only direction observed in the unbiased simulations, and bending toward the minor groove is known to be significantly harder (Ma & van der Vaart, 2016; N. Ma & A. van der Vaart, 2017). The PMFs show that flipping is eased when DNA is moderately bent. Large barriers for flipping were observed when the DNA was either unbent or heavily bent (Fig. 7). For unbent DNA, flipping out of the minor groove pushed the DNA to bend toward the major groove to make enough space for the flipped base, indicating that DNA bends to facilitate flipping (Fig. S6). Back bending toward the minor groove was also

observed for flipping when DNA was heavily bent; consequently, flipping barriers were large when DNA was heavily bent. In either case, this behavior was due to the limited amount of space in the grooves for flipping. When DNA was unbent the minor groove was compressed, while the major groove was compressed when it was highly bent. This compression of the grooves at very high or low bending made it more difficult for the base to flip out.

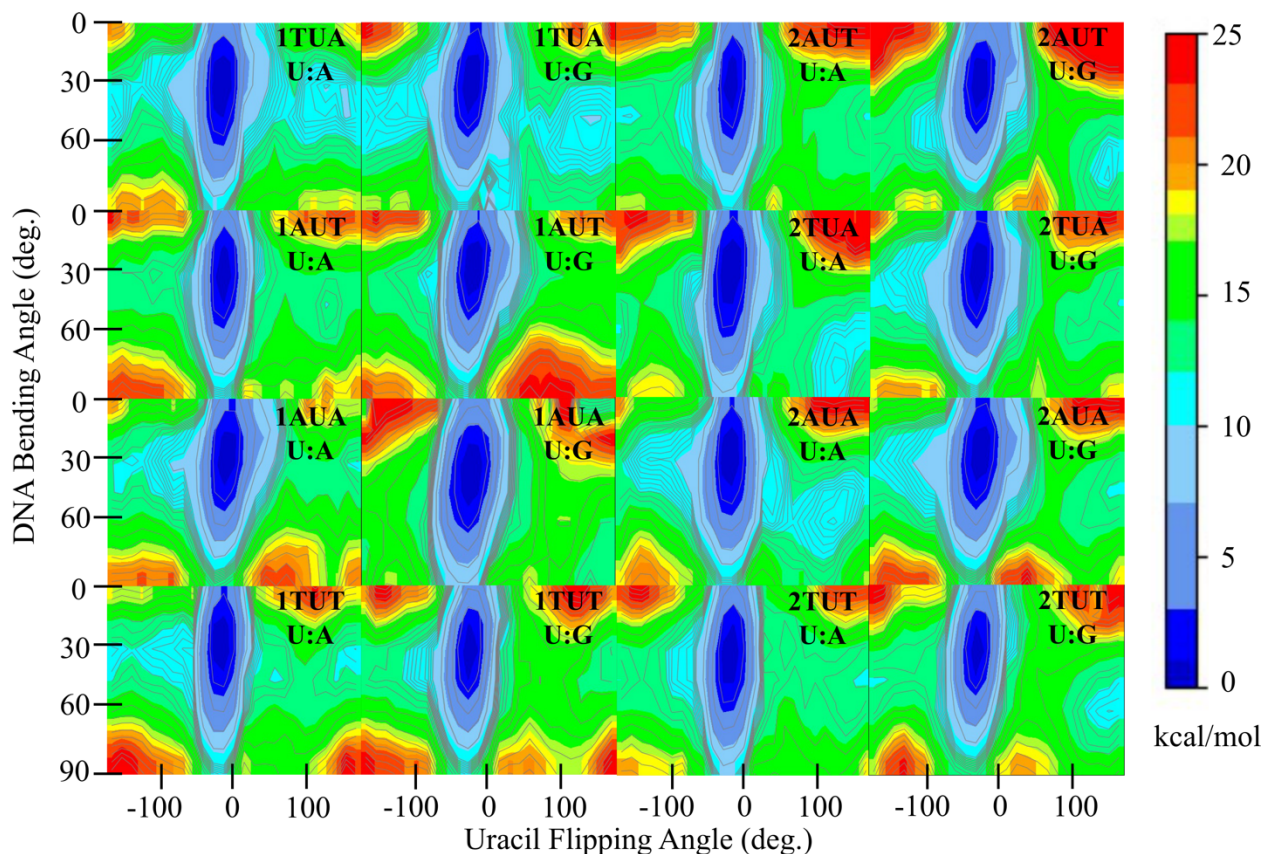


Figure 7 Potential of mean force as a function the DNA bending angle and uracil flipping angle (in kcal/mol).

Barriers to flipping were lowest when DNA was bent between 30° and 50°. Bending angles in this range did not display any back bending towards the minor groove since the groove widths were wide enough for uracil to be flipped in either direction. In nearly all cases, the main basin was oriented diagonally, indicating that bending primes flipping toward major groove. Flipping toward the major groove was indeed favored for all sequences which was consistent with unbiased observations as well. Many sequences displayed a high energy basin for bent DNA with a flipping angle of 160°-170°. In this basin, the extrahelical uracil interacted favorably with the minor groove.

This behavior was also common in many of the unbiased simulations. The free energy for major groove flipping was minimal at a DNA bending angle of $32 \pm 6^\circ$ for all sequences; for minor groove flipping it ranged between 15° and 45° . The minimum free energy was located at a flipping angle of $-28 \pm 7^\circ$ for U:G sequences and $-15 \pm 7^\circ$ for U:A sequences. The error bars correspond to the grid size for the PMF calculation; minima were found at the same grid location for all simulations. For U:G sequences this was at the same value as the average flipping angle in the unbiased simulations (-28°), while for U:A sequences it was slightly shifted (-19°). This consistency reaffirms that the pathway to flipping is lower in free energy for U:G than U:A.

Fig. 8 shows the free energy barriers for uracil flipping. These were obtained from the potential of mean force of Fig. 7 at a flipping angle of -40° . For U:A strands, major groove flipping barriers were largest for AUT sequences, followed by AUA and TUT sequences; barriers were lowest for TUA sequences. Similar trends were observed in U:G strands, except that 1AUA had the lowest barrier. Flipping toward the major groove was particularly facile for U:G strands, with barriers of 1.3-2.2 kcal/mol; these barriers were between 1.7-4.2 kcal/mol less than in U:A strands. For U:A strands, minor groove flipping barriers were largest for AUT and TUT sequences, followed by TUA and AUA sequences. Minor groove flipping barriers did not follow the same trends in U:G strands: TUT sequences had the largest barriers, but no other patterns were found. Some U:G strands had higher minor groove flipping barriers than U:A strands; moreover, in the U:G strands, the difference between minor and major groove flipping (7.5-12.9 kcal/mol) was larger than for U:A strands (1.7-8.5 kcal/mol). Overall, sequence trends for the preferred direction of flipping (i.e. flipping toward major groove) were similar to the those observed in unbiased simulations: a 5' T with 3' A facilitated base flipping, while a 5' A with a 3' T kept the uracil better stacked. The difference in flipping for AUT and TUA sequences is also consistent with findings from the unbiased simulations.

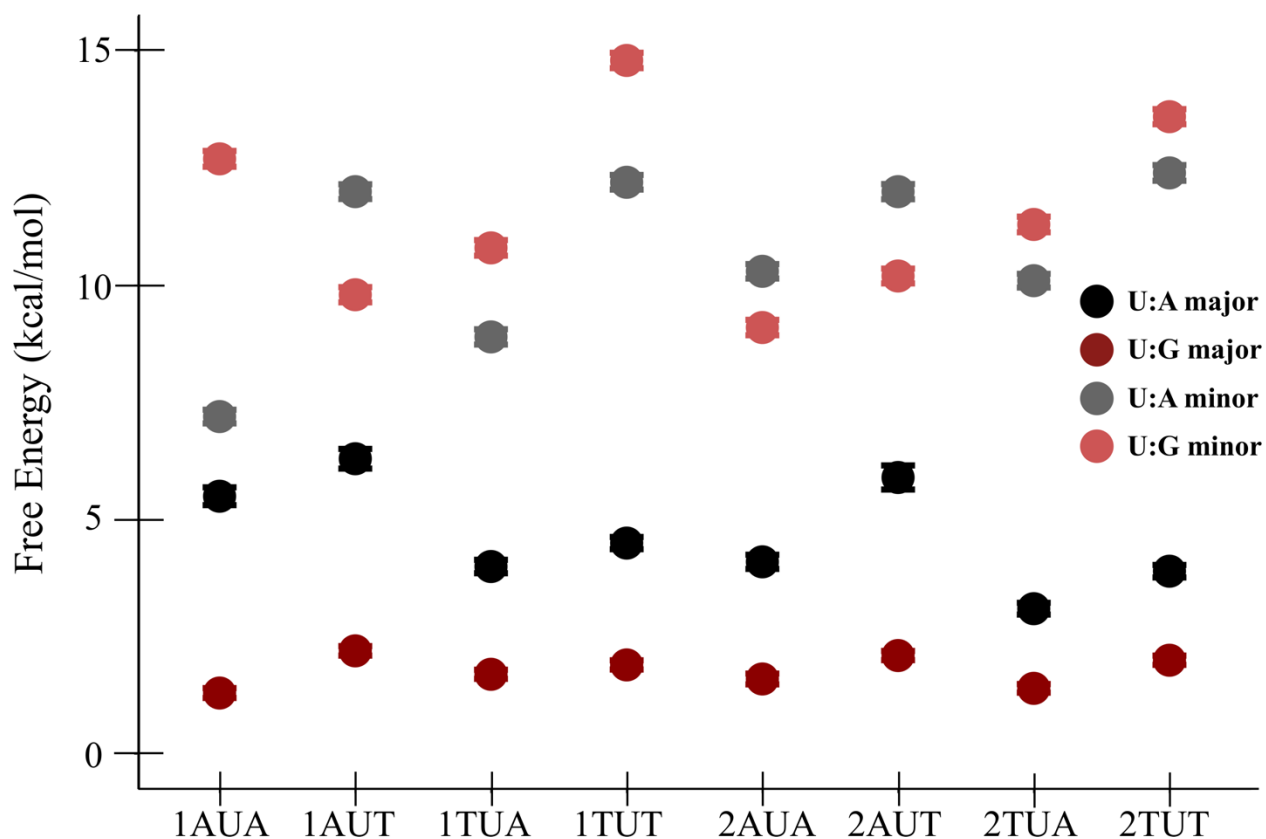


Figure 8: Flipping free energy barriers for each of the sequences in Table 1 where uracil is base paired to either A or G.

Discussion

A systematic simulation study of uracil-damaged DNA showed that the lesion significantly increases the flexibility of DNA in a sequence-dependent manner. Apart from the bending persistence length and minor groove flipping barriers of U:G base paired sequences, sequence trends in the flexibilities of U:A and U:G base paired DNA largely followed trends of undamaged DNA. For U:G base paired strands, sequence trends in local flexibilities, groove widths, and major groove flipping were similar to undamaged DNA, but motions were more amplified and also significantly larger than in U:A paired strands. Therefore, by and large, the lesion mostly enhances trends in flexibility that are already apparent in undamaged DNA. In dictating deformability of grooves and steps, DNA bending and base flipping, the 3' base immediately adjacent to the lesion is generally the most influential, and 3' purines induce more flexibility than 3' pyrimidines. Flexibility is strongly promoted by a 5' thymine immediately

adjacent to the lesion. Other 5' bases are less influential, although a 5' cytosine mildly promotes flexibility.

It was recently established that UDG excision efficiency is controlled by substrate flexibility for DNA with U:A base paired AUT, TUA, AUA and TUT motifs, with higher rates for more flexible substrates (Orndorff et al., 2023). The current simulations echo trends in flexibility of those sequences. More importantly, when combined with previously published UDG excision rates, they indicate that the link between substrate flexibility and repair efficiency can be generalized to larger sequence contexts. Relative excision efficiencies have been reported for uracils embedded within a long (> 6 kbp) viral dsDNA (Nilsen et al., 1995), but information on the flexibility of most constructs was lacking until now. Inspection of these efficiencies shows a clear correspondence to our established trends in flexibility: sequences with a vicinal 3' T are generally poorer substrates, while sequences with a 5' T or 3' A or 3' G are generally better substrates. These observations imply that the link between substrate flexibility and UDG efficiency is much more widely valid.

Sequences in which the uracil is base paired to guanine resemble a mismatch and exist in a naturally open state. This not only decreases base stacking and promotes base flipping, but also contributes to the flexibility of the entire sequence. Sequences in which uracil is base paired to adenine are better stacked; while much more flexible than undamaged DNA, these sequences bend and flip less than U:G base paired sequences. It is of interest that K_M but not k_{cat} was correlated to substrate flexibility in the aforementioned study (Orndorff et al., 2023), which indicates that more flexible substrates are better bound. Combined with our finding that U:G strands are much more flexible than U:A strands, this may explain why UDG binds U:G base paired strands better than U:A strands (Panayotou et al., 1998).

DNA bending and base flipping are strongly coupled, with more flipping for more bendable sequences. DNA bending primes uracil damaged strands toward major groove flipping, and DNA bending by 30-50° facilitates flipping. Flipping for unbent or highly bent DNA is unfavorable, because of steric interactions of the flipped base with the groove. In fact, flipping at low angles induces a larger DNA bending angle, and flipping at high angle a back bend. The calculations

reaffirm the notion that in the UDG complex DNA is bent to facilitate flipping (Franco et al., 2013; Orndorff et al., 2023), but also explain why bending is only moderate (Bianchet et al., 2003; Burmeister et al., 2015; Earl et al., 2018; H. Kosaka et al., 2007; H. Kosaka, Nakagawa, N., Masui, R., Kuramitsu, S., Hoseki, J., 2007; S. S. Parikh et al., 1998; Sudip S. Parikh et al., 2000; Parker et al., 2007; Pedersen et al., 2015; Geir Slupphaug et al., 1996). Moreover, in the free energy simulations, the bending angle at which the free energy for major groove flipping was minimized was $32 \pm 6^\circ$ for the simulated sequences (where the error is from the grid size of the PMF). While the PMF simulations were only done for a subset of sequences, it is intriguing that this value is the same for all. Moreover, this bending angle is the same as the average bending angle in the crystal structures ($33 \pm 8^\circ$). These observations suggest that UDG binds the substrate in a bend conformation that is optimal for flipping, and that this particularly bending angle is optimal for flipping of all tested substrates.

Supplemental Online Material

Figures on the rmsds, distribution of Pearson correlation coefficients for uracil motion, differences in groove widths between damaged and undamaged strands, differences in step parameters between damaged and undamaged strands, standard deviation of the roll angle over all simulation snapshots, and DNA base flipping and bending of the 2AUT A:U sequence, as well as a table with groove widths for palindromic steps are provided as Supplemental Online Material.

Disclosure Statement

The authors report there are no competing interests to declare.

Acknowledgments

This work was supported by the NSF under grant MCB-1919096. Computer time was provided by USF Research Computing.

References

- Banavali, N. K., & MacKerell, A. D. (2002). Free Energy and Structural Pathways of Base Flipping in a DNA GCGC Containing Sequence. *Journal of Molecular Biology*, 319(1), 141-160. doi: [https://doi.org/10.1016/S0022-2836\(02\)00194-8](https://doi.org/10.1016/S0022-2836(02)00194-8)
- Baumann, C. G., Smith, S. B., Bloomfield, V. A., & Bustamante, C. (1997). Ionic effects on the elasticity of single DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 94(12), 6185-6190. doi: 10.1073/pnas.94.12.6185
- Bellamy, S. R. W., & Baldwin, G. S. (2001). A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1. *Nucleic Acids Research*, 29(18), 3857-3863. doi: 10.1093/nar/29.18.3857
- Beveridge, D. L., Barreiro, G., Byun, K. S., Case, D. A., Cheatham, T. E., Dixit, S. B., . . . Young, M. A. (2004). Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(C(p)G) steps. *Biophysical Journal*, 87(6), 3799-3813. doi: 10.1529/biophysj.104.045252
- Beveridge, D. L., Cheatham, T. E., III, & Mezei, M. (2012). The ABCs of molecular dynamics simulations on B-DNA, circa 2012. *Journal of Biosciences*, 37(3), 379-397. doi: 10.1007/s12038-012-9222-6
- Bianchet, M. A., Seiple, L. A., Jiang, Y. L., Ichikawa, Y., Amzel, L. M., & Stivers, J. T. (2003). Electrostatic Guidance of Glycosyl Cation Migration along the Reaction Coordinate of Uracil DNA Glycosylase. *Biochemistry*, 42(43), 12455-12460. doi: 10.1021/bi035372+
- Brooks, S. C., Adhikary, S., Robinson, E. H., & Eichman, B. F. (2013). Recent advances in the structural mechanisms of DNA glycosylases. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1834(1), 247-271. doi: 10.1016/j.bbapap.2012.10.005
- Brunet, A., Tardin, C., Salomé, L., Rousseau, P., Destainville, N., & Manghi, M. (2015). Dependence of DNA Persistence Length on Ionic Strength of Solutions with Monovalent and Divalent Salts: A Joint Theory–Experiment Study. *Macromolecules*, 48(11), 3641-3652. doi: 10.1021/acs.macromol.5b00735
- Burmeister, W. P., Tarbouriech, N., Fender, P., Contesto-Richefeu, C., Peyrefitte, C. N., & Iseni, F. (2015). Crystal Structure of the Vaccinia Virus Uracil-DNA Glycosylase in Complex with DNA*. *Journal of Biological Chemistry*, 290(29), 17923-17934. doi: <https://doi.org/10.1074/jbc.M115.648352>
- Bustamante, C., Bryant, Z., & Smith, S. B. (2003). Ten years of tension: single-molecule DNA mechanics. *Nature*, 421(6921), 423-427. doi: 10.1038/nature01405
- Carbonnaux, C., Fazakerley, G. V., & Sowers, L. C. (1990). An NMR structural study of deaminated base-pairs in DNA. *Nucleic Acids Research*, 18(14), 4075-4081. doi: 10.1093/nar/18.14.4075
- Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E., III, Cruzeiro, V. D. W., . . . Kollman, P. A. (2018). AMBER 2018. University of California, San Francisco.
- Chhetri, K. B., Sharma, A., Naskar, S., & Maiti, P. K. (2022). Nanoscale structures and mechanics of peptide nucleic acids. *Nanoscale*, 14(17), 6620-6635. doi: 10.1039/D1NR04239D

- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089-10092. doi: 10.1063/1.464397
- Dixit, S. B., Beveridge, D. L., Case, D. A., Cheatham, T. E., Giudice, E., Lankas, F., . . . Varnai, P. (2005). Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II: Sequence context effects on the dynamical structures of the 10 unique dinucleotide steps. *Biophysical Journal*, 89(6), 3721-3740. doi: 10.1529/biophysj.105.067397
- Earl, C., Bagn  ris, C., Zeman, K., Cole, A., Barrett, T., & Savva, R. (2018). A structurally conserved motif in γ -herpesvirus uracil-DNA glycosylases elicits duplex nucleotide-flipping. *Nucleic Acids Research*, 46(8), 4286-4300. doi: 10.1093/nar/gky217
- Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., . . . Pande, V. S. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Computational Biology*, 13(7), e1005659. doi: 10.1371/journal.pcbi.1005659
- Eftedal, I., Guddal, P. H., Slupphaug, G., Volden, G., & Krokan, H. E. (1993). Consensus Sequences for Good and Poor Removal of Uracil from Double-Stranded DNA by Uracil-DNA Glycosylase. *Nucleic Acids Research*, 21(9), 2095-2101. doi: DOI 10.1093/nar/21.9.2095
- Fadda, E., & Pom  s, R. (2011). On the molecular basis of uracil recognition in DNA: comparative study of T-A versus U-A structure, dynamics and open base pair kinetics. *Nucleic acids research*, 39(2), 767-780. doi: 10.1093/nar/gkq812
- Franco, D., Sgrignani, J., Bussi, G., & Magistrato, A. (2013). Structural Role of Uracil DNA Glycosylase for the Recognition of Uracil in DNA Duplexes. Clues from Atomistic Simulations. *Journal of Chemical Information and Modeling*, 53(6), 1371-1387. doi: 10.1021/ci4001647
- Fuxreiter, M., Luo, N., Jedlovsky, P., Simon, I., & Osman, R. (2002). Role of Base Flipping in Specific Recognition of Damaged DNA by Repair Enzymes. *Journal of Molecular Biology*, 323(5), 823-834. doi: [https://doi.org/10.1016/S0022-2836\(02\)00999-3](https://doi.org/10.1016/S0022-2836(02)00999-3)
- Galindo-Murillo, R., Robertson, J. C., Zgarbov  , M., Šponer, J., Otyepka, M., Jure  ka, P., & Cheatham, T. E. (2016). Assessing the Current State of Amber Force Field Modifications for DNA. *Journal of Chemical Theory and Computation*, 12(8), 4114-4127. doi: 10.1021/acs.jctc.6b00186
- Hagerman, P. J. (1988). Flexibility of DNA. *Annual Review of Biophysics and Biophysical Chemistry*, 17, 265-286
- Herrero-Gal  n, E., Fuentes-Perez, M. E., Carrasco, C., Valpuesta, J. M., Carrascosa, J. L., Moreno-Herrero, F., & Arias-Gonzalez, J. R. (2013). Mechanical Identities of RNA and DNA Double Helices Unveiled at the Single-Molecule Level. *Journal of the American Chemical Society*, 135(1), 122-131. doi: 10.1021/ja3054755
- Hoeijmakers, J. H. J. (2009). MOLECULAR ORIGINS OF CANCER DNA Damage, Aging, and Cancer. *New England Journal of Medicine*, 361(15), 1475-1485. doi: 10.1056/NEJMr0804615
- H  lz, K., Pavlic, A., Lietard, J., & Somoza, M. M. (2019). Specificity and Efficiency of the Uracil DNA Glycosylase-Mediated Strand Cleavage Surveyed on Large Sequence Libraries. *Scientific Reports*, 9(1), 17822. doi: 10.1038/s41598-019-54044-x

- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2), 926-935. doi: 10.1063/1.445869
- Kosaka, H., Hoseki, J., Nakagawa, N., Kuramitsu, S., & Masui, R. (2007). Crystal Structure of Family 5 Uracil-DNA Glycosylase Bound to DNA. *Journal of Molecular Biology*, 373(4), 839-850. doi: <https://doi.org/10.1016/j.jmb.2007.08.022>
- Kosaka, H., Nakagawa, N., Masui, R., Kuramitsu, S., Hoseki, J. (2007). Crystal structure of uracil-DNA glycosylase in complex with AP:C containing DNA. *RIKEN Structural Genomics/Proteomics Initiative (RSGI)*. doi: 10.2210/pdb2dp6/pdb
- Krokan, H. E., Drablos, F., & Slupphaug, G. (2002). Uracil in DNA--occurrence, consequences and repair. *Oncogene*, 21(58), 8935-8948. doi: 10.1038/sj.onc.1205996
- Lavery, R., Zakrzewska, K., Beveridge, D., Bishop, T. C., Case, D. A., Cheatham, T., III, . . . Spomer, J. (2010). A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Research*, 38(1), 299-313. doi: 10.1093/nar/gkp834
- Lipfert, J., Kerssemakers, J. W. J., Jager, T., & Dekker, N. H. (2010). Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments. *Nature Methods*, 7(12), 977-980. doi: 10.1038/nmeth.1520
- Lu, X.-J., & Olson, W. K. (2008). 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nature Protocols*, 3(7), 1213-1227. doi: 10.1038/nprot.2008.104
- Lu, Y., Weers, B., & Stellwagen, N. C. (2002). DNA persistence length revisited. *Biopolymers*, 61(4), 261-275. doi: <https://doi.org/10.1002/bip.10151>
- Ma, N., & van der Vaart, A. (2016). Anisotropy of B-DNA Groove Bending. *Journal of the American Chemical Society*, 138(31), 9951-9958. doi: 10.1021/jacs.6b05136
- Ma, N., & van der Vaart, A. (2017). Free Energy Coupling between DNA Bending and Base Flipping. *Journal of Chemical Information and Modeling*, 57(8), 2020-2026. doi: 10.1021/acs.jcim.7b00215
- Ma, N., & van der Vaart, A. (2017). KCI Dependence of B-DNA Groove Bending Anisotropy. *Journal of Physical Chemistry B*, 121(21), 5322-5330. doi: 10.1021/acs.jpcc.7b01957
- Mardt, A., Gorris, R. F., Ferraro, F., Ulrich, P., Zahran, M., & Imhof, P. (2022). Effect of a U:G mispair on the water around DNA. *Biophysical Chemistry*, 283, 106779. doi: <https://doi.org/10.1016/j.bpc.2022.106779>
- Marin-Gonzalez, A., Vilhena, J. G., Perez, R., & Moreno-Herrero, F. (2021). A molecular view of DNA flexibility. *Quarterly Reviews of Biophysics*, 54. doi: 10.1017/s0033583521000068
- Mazur, A. K. (2007). Wormlike Chain Theory and Bending of Short DNA. *Physical Review Letters*, 98(21), 218102. doi: 10.1103/PhysRevLett.98.218102
- Nilsen, H., Yazdankhah, S. P., Eftedal, I., & Krokan, H. E. (1995). Sequence Specificity for Removal of Uracil from U-Center-Dot-A Pairs and U-Center-Dot-G Mismatches by Uracil-DNA Glycosylase from Escherichia-Coli, and Correlation with Mutational Hotspots. *Febs Letters*, 362(2), 205-209. doi: Doi 10.1016/0014-5793(95)00244-4
- Orndorff, P. B., Poddar, S., Owens, A. M., Kumari, N., Ugaz, B. T., Amin, S., . . . Levitus, M. (2023). Uracil-DNA glycosylase efficiency is modulated by substrate rigidity. *Scientific Reports*, 13, 3915. doi: 10.1038/s41598-023-30620-0

- Panayotou, G., Brown, T., Barlow, T., Pearl, L. H., & Savva, R. (1998). Direct measurement of the substrate preference of uracil-DNA glycosylase. *Journal of Biological Chemistry*, 273(1), 45-50. doi: 10.1074/jbc.273.1.45
- Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E., & Tainer, J. A. (1998). Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO Journal*, 17(17), 5214-5226. doi: DOI 10.1093/emboj/17.17.5214
- Parikh, S. S., Walcher, G., Jones, G. D., Slupphaug, G., Krokan, H. E., Blackburn, G. M., & Tainer, J. A. (2000). Uracil-DNA glycosylase-DNA substrate and product structures: Conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proceedings of the National Academy of Sciences*, 97(10), 5083-5088. doi: 10.1073/pnas.97.10.5083
- Parker, J. B., Bianchet, M. A., Krosky, D. J., Friedman, J. I., Amzel, L. M., & Stivers, J. T. (2007). Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature*, 449(7161), 433-437. doi: 10.1038/nature06131
- Parker, J. B., & Stivers, J. T. (2011). Dynamics of Uracil and 5-Fluorouracil in DNA. *Biochemistry*, 50(5), 612-617. doi: 10.1021/bi101536k
- Pedersen, H. L., Johnson, K. A., McVey, C. E., Leiros, I., & Moe, E. (2015). Structure determination of uracil-DNA N-glycosylase from *Deinococcus radiodurans* in complex with DNA. *Acta Crystallographica Section D*, 71(10), 2137-2149. doi: doi:10.1107/S1399004715014157
- Peguro-Tejada, A., & van der Vaart, A. (2017). Biasing Simulations of DNA Base Pair Parameters with Application to Propellor Twisting in AT/AT, AA/TT, and AC/GT Steps and Their Uracil Analogs. *Journal of Chemical Information and Modeling*, 57(1), 85-92. doi: 10.1021/acs.jcim.6b00660
- Peters, J. P., & Maher, L. J. (2010). DNA curvature and flexibility in vitro and in vivo. *Quarterly Reviews of Biophysics*, 43(1), 23-63. doi: 10.1017/s0033583510000077
- Priyakumar, U. D., & MacKerell, A. D. (2006). Computational Approaches for Investigating Base Flipping in Oligonucleotides. *Chemical Reviews*, 106(2), 489-505. doi: 10.1021/cr040475z
- Ramstein, J., & Lavery, R. (1988). Energetic coupling between DNA bending and base pair opening. *Proceedings of the National Academy of Sciences*, 85(19), 7231-7235. doi: 10.1073/pnas.85.19.7231
- Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. (1977). Numerical-Integration of Cartesian Equations of Motion of a System with Constraints – Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics*, 23, 327-341. doi: 10.1016/0021-9991(77)90098-5
- Seibert, E., Ross, J. B. A., & Osman, R. (2003). Contribution of Opening and Bending Dynamics to Specific Recognition of DNA Damage. *Journal of Molecular Biology*, 330(4), 687-703. doi: [https://doi.org/10.1016/S0022-2836\(03\)00598-9](https://doi.org/10.1016/S0022-2836(03)00598-9)
- Shirts, M. R., & Chodera, J. D. (2008). Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of chemical physics*, 129(12), 124105-124105. doi: 10.1063/1.2978177
- Slupphaug, G., Eftedal, I., Kavli, B., Bharati, S., Helle, N. M., Haug, T., . . . Krokan, H. E. (1995). Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. *Biochemistry*, 34(1), 128-138. doi: 10.1021/bi00001a016

- Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., & Tainer, J. A. (1996). A nucleotide-flipping mechanism from the structure of human uracil–DNA glycosylase bound to DNA. *Nature*, 384(6604), 87-92. doi: 10.1038/384087a0
- Stivers, J. T., & Jiang, Y. L. (2003). A Mechanistic Perspective on the Chemistry of DNA Repair Glycosylases. *Chemical Reviews*, 103(7), 2729-2760. doi: 10.1021/cr010219b
- Stivers, J. T., Pankiewicz, K. W., & Watanabe, K. A. (1999). Kinetic mechanism of damage site recognition and uracil flipping by Escherichia coli uracil DNA glycosylase. *Biochemistry*, 38(3), 952-963. doi: 10.1021/bi9818669
- Torrie, G. M., & Valleau, J. P. (1977). Non-physical sampling distributions in Monte-Carlo free-energy estimation - Umbrella sampling. *Journal of Computational Physics*, 23(2), 187-199. doi: 10.1016/0021-9991(77)90121-8
- Velasco-Berrelleza, V., Burman, M., Shepherd, J., Leake, M., Golestanian, R., & Noy, A. (2020). SerraNA: A program to determine nucleic acids elasticity from simulation data. *Physical Chemistry Chemical Physics*, 22. doi: 10.1039/D0CP02713H
- Visnes, T., Doseth, B., Pettersen, H. S., Hagen, L., Sousa, M. M. L., Akbari, M., . . . Krokan, H. E. (2009). Uracil in DNA and its processing by different DNA glycosylases. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364(1517), 563-568. doi: 10.1098/rstb.2008.0186
- Williams, J. S., Lujan, S. A., & Kunkel, T. A. (2016). Processing ribonucleotides incorporated during eukaryotic DNA replication. *Nature Reviews Molecular Cell Biology*, 17(6), 350-363. doi: 10.1038/nrm.2016.37