Register shifted structures in base flipped 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, FL 33620, United States

Keywords: DNA, base flipping, uracil, DNA damage, simulation.

ABSTRACT: We report the occurrence of register shifted structures in simulations of uracil-containing dsDNA. These occur when the 3' base vicinal to uracil is thymine in U:A base paired DNA. Upon base flipping of uracil, this 3' thymine hydrogen bonds with the adenine across the uracil instead of its complementary base. The register shifted structure is persistent, and sterically blocks re-entry of uracil into the helix stack. Register shifting might be important for DNA repair, since the longer exposure of the lesion in register shifted structures could facilitate enzymatic recognition and repair.

Cellular DNA is subjected to thousands of spontaneous lesions a day.[1] A common lesion is uracil, which can arise from misincorporation of dUMP, or from the spontaneous deamination of cytosine. Misincorporation, which is stimulated by folate deficiency,[2] can lead to chromosome breaks, while cytosine deamination introduces G:C to A:T transition mutations that are 100% mutagenic.[3] Uracil lesions are repaired by enzymes from the base excision repair pathway. Repair is initiated by uracil DNA glycosylase (UDG), which excises the lesion by employing a base flipping mechanism that extrudes the lesion extrahelically.[4]

UDG excision rates depend on sequence.[5] Using a combination of time-resolved fluorescence spectroscopy, NMR imino proton exchange measurements, and molecular dynamics (MD) simulations, we recently demonstrated that this dependence stems from differences in DNA deformability.[5f] k_{cat}/K_M values for U:A base paired strands with AUT, TUA, AUA and TUT motifs were strongly correlated to DNA rigidity, with higher excision efficiency for more flexible strands. This link is widely valid, as shown by a comparison

of relative excision efficiencies[5a] to simulated rigidities of a large library of uracil-containing strands.[6]

Here we report the occurrence of a rare event observed in one of these simulations, and verified in other, that may have important biological implications. 146 U:A and U:G base paired uracil-containing sequences were simulated, spanning over 100 μs in total (Table S1).[5f, 6] The MD simulations were performed at 1 bar and 300 K with OPENMM,[7] in rectangular boxes of either 100 mM NaCl or 150 mM KCl TIP3P water[8] with 15 Å solvation layers, using SHAKE for bonds involving hydrogen atoms,[9] a nonbonded cutoff of 12 Å and the particle mesh Ewald (PME) method with an error tolerance of 0.00005 for long range electrostatics,[10] Langevin dynamics with a friction coefficient of 5 ps^{-1} and a 2 fs time step, and the AMBER OL15 force field.[11] Heating took place in the presence of harmonic restraints, which were slowly released during equilibration and absent in production; details are given in Refs. [5f, 6]. DNA did not interact with its periodic images in any of the simulations; apart from fraying of the terminal base pairs and the transitory conformational changes described below, DNA was stable and maintained its structural integrity.

Toward the end of a simulation of U:A base paired 5'-GCGC$(UTGC)_3$G-3' dsDNA (base pairing and numbering are shown in Fig. 1a), U_9, the central lesion, spontaneously flipped toward the major groove (Fig. 1b). Base flipping has been observed in simulations before,[12] and was assessed by monitoring a pseudodihedral angle,[13] which was < -40° for flipping toward the major and > 40° for flipping toward the minor groove. Base flipping of uracil (but no other base) was observed in many sequences, and occurred exclusively toward the major groove. Base flipping makes DNA more bendable, and 31 ns after flipping the strand was bent 60° toward the major groove. Over the next 4 ns it bent back to 15°; during this back bending, a dramatic structural change started to occur. Upon reaching a flipping angle of -80°, the tilt angle of the T_{10}:A_{25} base pair decreased from 0° to -10°, while its twist angle increased from 35° to 45°. At the same time, the buckling angle of T_{10} increased by nearly 20° from -5° to 15°. These movements were enabled by the flipped U_9, which led to a loss of steric

interactions of T_{10} with its 5' vicinal base. The motions led to transient, weak hydrogen bonds of T_{10} with two bases: A_{25}, its complementary base, and A_{26}, the base complementary to U_9 (Fig. 1c). The hydrogen bond with A_{25} subsequently weakened, while the hydrogen bond with A_{26} gained in strength. The T_{10} − A_{26} hydrogen bond broke 41 ns after the flip, upon which T_{10} and A_{26} base paired with two Watson-Crick hydrogen bonds. The result was a highly unusual register shifted structure, where T_{10} base paired with A_{26} instead of its complementary base, A_{25} was unpaired, and U_9 flipped (Fig. 1d).

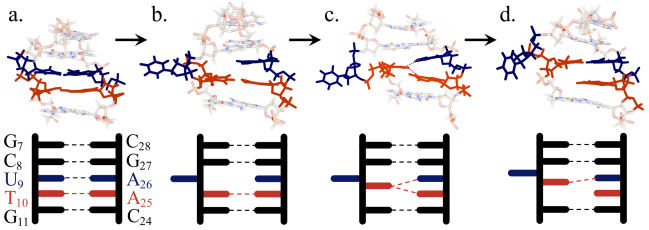

Figure 1. Pathway leading to the register shifted structure for the 5'-GCGC(UTGC)$_3$G-3' sequence. Base numbering and complementary base pairing are shown in panel a, and the register shifted structure in panel d; other panels are discussed in the text.

Further extension of the simulation showed that the register shifted structure was stable and persistent for 125 ns. Structural changes were limited to the $U_9T_{10}/A_{25}A_{26}$ step and did not affect the structure or dynamics of the remaining bases (Fig. 2). The register shifted structure spontaneously disappeared by base pairing of T_{10} to its complementary base A_{25}, while U_9 remained base flipped. During this process, the buckling angle for T_{10} decreased from 25° to -25° over a time period of 1 ns; during this timeframe the T_{10}:A_{26} hydrogen bonds broke, while the T_{10}:A_{25} base pair reformed.

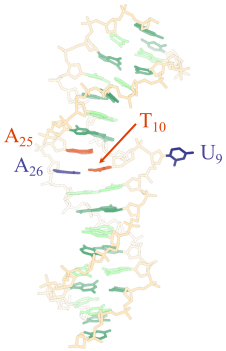

Figure 2. 5'-GCGC(UTGC)$_3$G-3' register shifted structure; hydrogen atoms not shown for clarity.

Canonical T_{10}:A_{25} pairing persisted for 50 ns, while U_9 remained extrahelical. Then, a second register shift occurred. This structure also involved T_{10}, formed in the same manner, but quicker (over 2 ns), and persisted for 230 ns. U_9 was the only base that was flipped; it remained flipped while the register shifted structure occurred and disappeared. 27 ns later, a third register shifted structure with T_{10} appeared. 231 ns after this third structure disappeared, U_9 remained flipped and another noncanonical structure emerged in which T_{10} stacked between A_{25} and A_{26} (Fig. 3a). This unusual stacking lasted for 200 ns; during this time, U_9 remained extrahelical. Similar stacking behavior was seen in long MD simulations of undamaged DNA.[14]
Since the register shift was observed in only one out of many simulations, several additional MD simulations of the 5'-GCGC(UTGC)$_3$G-3' sequence were performed. No register shift was observed in any of these. However, register shifts were observed in 2 out of 5 simulations that were started from a snapshot of the register-shifted trajectory. This snapshot was taken 21 ns after the U_9 flip, but well before the aforementioned changes in DNA bending, tilt, twist, and buckle of T_{10} (Fig. 3b). Reheating and production followed the same protocol as the original trajectory, but initial velocities were redrawn at random. For these two simulations, the register shifts occurred in a manner similar to that of the original trajectory, and persisted well over 100 ns. In one of these simulations, U_9 flipped back intrahelically and reformed hydrogen bonds with A_{26} 50 ns after the register shifted structure disappeared. In the other simulation, the register shifted structure appeared another 3 times, separated by ~20 ns; each lasted for ~155 ns, while U_9 remained flipped. U_9 flipped back and base paired with A_{26} in the simulations that did not display a register shift.

The register shift blocks the extrahelical uracil from flipping back into the helical stack. Using cut off values of ±40° of the pseudodihedral angle as the criterion for base flipping, the average time a uracil was flipped in structures without register shift was 49 ± 33 ns. The average lifetime of the register shifted structure was much longer, 166 ± 58 ns; during this time the uracil remained extrahelical. Thus, the register shift blocked uracil from flipping back in and extended its extrahelical exposure.

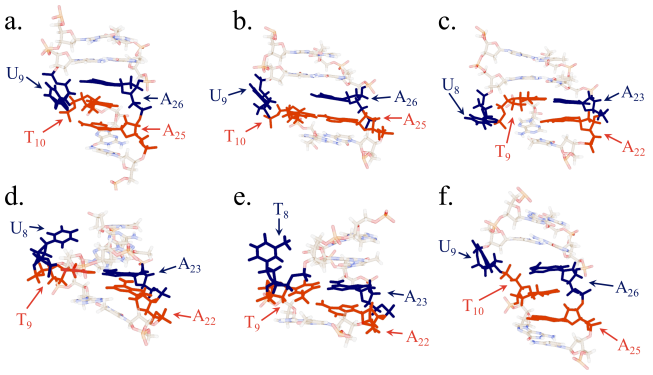

Figure 3. Simulation structures. Noncanonical base stacking (a) and the snapshot 21 ns after base flipping (b) for the 5'-GCGC(UTGC)$_3$G-3' sequence simulated with the AMBER force field. Register shifted structures for flipping toward the major (c) and minor (d) grooves of the 5′- CGTGCAAUTATGACG-3′ sequence in AMBER-based umbrella sampling simulations. (e) Register shifted structure in AMBER-based umbrella simulations of 5′- CGTGCAA**T**TATGACG-3′ dsDNA; base flipping of the bolded T_8 was enhanced. (f) Noncanonical base stacking in CHARMM-based simulations of 5'-GCGC(UTGC)$_3$G-3'.

We had previously observed register shifted structures in umbrella sampling[15] simulations of U:A containing dsDNA (unpublished). In these either sampling of the uracil flipping angle or, since bending aids flipping,[5f, 6, 16] sampling of both the flipping and DNA bending angle were enhanced by the use of harmonic restraining potentials. This was seen for 5′-CGTGCAAUTATGACG-3′ (1D and 2D) and 5′-CGTGAAAUTGTTACG-3′ (1D only). At the time, these simulations were discarded because of the occurrence of these noncanonical register shifted structures. Register shifting was not observed when the simulations were rerun using different random seeds for heating. Register shifting was also not observed in umbrella sampling simulations of other uracil-containing sequences, including other sequences with a vicinal 3' thymine. Sequences were simulated with OPENMM[7] at 300 K in rectangular boxes of 100 mM NaCl TIP3P[8] water with 15 Å solvent layers, using SHAKE,[9] PME,[10] Langevin dynamics, and the AMBER OL15 force field;[11] simulation settings were the same as described above. Flipping angles in the 1D umbrella simulations were biased

in windows of 5° with a force constant of 200 kcal/(mol·rad^2); in the 2D umbrella simulations bending and flipping were biased in 15° windows with a force constant of 65 kcal/(mol·rad^2). Simulations were run sequentially, starting from windows in which uracil was fully stacked and unflipped.
Identical to what was observed in the unbiased simulations, the 3' thymine vicinal to uracil would register shift by hydrogen bonding to the adenine that was complementary to the uracil. The register shift occurred when uracil was flipped to either the major (Fig. 3c) or minor groove (Fig. 3d), in a manner similar to the unbiased simulations. Once formed, the register shifted structure persisted. Even when uracil was biased to restack in subsequent simulation windows, register shifting prevailed: restacking of uracil was sterically blocked by the register shifted 3' vicinal thymine. This block happened irrespective whether uracil was flipped toward the major or minor groove.

We also observed a register shift in a previously discarded 2D umbrella simulation of 5′-CGTGCAA**T**TATGACG-3' dsDNA, where base flipping of the (bolded) central thymine and DNA bending were enhanced (Fig. 3e). Upon flipping of the central thymine, the 3' vicinal thymine formed a hydrogen bond with the adenine that was complementary to the flipped thymine. Formation of this register shifted structure proceeded in a similar manner as for the uracil-containing strands, and restacking of the flipped thymine was prevented by the register shifted 3' vicinal thymine. Register shifting was not observed in reruns of the same sequence or in umbrella or unbiased simulations of other undamaged dsDNA sequences, which suggests that register shifting is easier for uracil than thymine. This is likely due to the lower base flipping propensity of thymine; the increased polarity of uracil increases solvent accessibility, flexibility and base flipping.[6, 12a, c, 17]

The effect of the register shift on re-entry of uracil is further illustrated in Fig. 4, where the free energy as a function of the uracil flipping and DNA bending angles, calculated from 2D umbrella sampling, is shown for the 5′-CGTGCAAUTATGACG-3′ sequence. When the register shift occurred, a deep basin emerged for the flipped-out state (Fig.

4a). In contrast, uracil preferred to be stacked when the register shift did not occur (Fig. 4b).

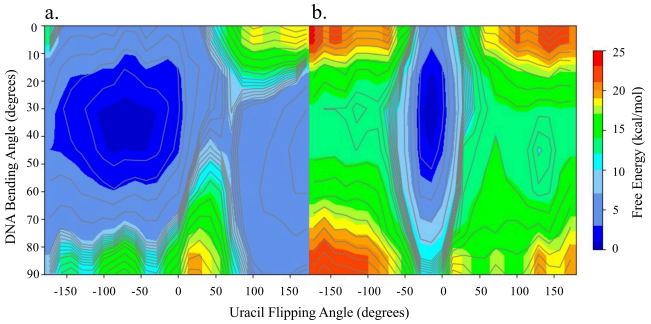

Figure 4. Free energy as a function of uracil flipping angle and DNA bending angle for 5′-CGTGCAAUTATGACG-3′ dsDNA with (a) and without (b) a register shift.

To investigate the effect of the force field, we ran five unbiased MD simulations of U:A base paired 5'-GCGC(UTGC)$_3$G-3' dsDNA with the CHARMM force field,[18] using the same setup as described above. In these simulations, no register shifts occurred. We then ran CHARMM simulations starting from the same snapshot that was used for the additional unbiased AMBER simulations (Fig. 3b). After heating and equilibration, the register shift occurred in three out of five simulations, in a manner similar to the AMBER simulations. In one of the CHARMM simulations without register shift, unusual base stacking was observed while U_9 was flipped (Fig. 3f); this unusual stacking was similar to that observed in the AMBER simulation (Fig. 3a), and lasted 61 ns. No register shift was observed in CHARMM-based umbrella sampling simulations.

In conclusion, we have observed a register shift in base pairing upon base flipping in simulations of dsDNA. The register shift occurred most frequently for base flipping of a uracil lesion in U:A base paired DNA, but was also observed for base flipping of a thymine in T:A base paired DNA. In all cases, the 3' base vicinal to the flipped base was a thymine; in the register shifted structure, this thymine base paired with the adenine across from the flipped base instead of its own complementary base. Register shifting is a rare event, since it was only observed in a few out of many sequences and trajectories. While the shift was first observed in umbrella sampling simulations, it is

unlikely a biasing artifact since it was also observed in unbiased simulations. Moreover, the noncanonical structure is likely not a force field artifact since it was observed in both AMBER and CHARMM simulations.

To the best of our knowledge, this is the first time a register shift concurrent with base flipping has been reported. While rare, register shifting might have significant biological implications for DNA repair, since the register shift blocks re-entry of the flipped base, and exposes it longer. Prolonged exposure of uracil might facilitate UDG recognition and increase the efficiency of repair. A dependence on local sequence is likely, and register shifting might also occur for U:G base paired DNA. Sequence effects in register shifting might be important for the genesis of mutation hotspots and molecular evolution, and will be studied in future work.

ASSOCIATED CONTENT

Supporting Information

A table of simulated sequences is provided in the Supporting Information. The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author

* Arjan van der Vaart, avandervaart@usf.edu.

Author Contributions

The manuscript was written through contributions of all authors.

Funding Sources

National Science Foundation MCB-1919096.

ACKNOWLEDGMENT

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

ABBREVIATIONS

MD, molecular dynamics; UDG, uracil DNA glycosylase.

REFERENCES

(1) Hoeijmakers, J. H. J. Molecular origins of cancer DNA. Damage, Aging, and Cancer. *N. Engl. J. Med.* **2009**, *361* (15), 1475-1485. DOI: 10.1056/NEJMra0804615.

(2) Blount, B. C.; Mack, M. M.; Wehr, C. M.; MacGregor, J. T.; Hiatt, R. A.; Wang, G.; Wickramasinghe, S. N.; Everson, R. B.; Ames, B. N. Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: Implications for cancer and neuronal damage. *Proc. Natl. Acad. Sci. USA* **1997**, *94* (7), 3290-3295. DOI: 10.1073/pnas.94.7.3290.

(3) Krokan, H. E.; Drablos, F.; Slupphaug, G. Uracil in DNA--occurrence, consequences and repair. *Oncogene* **2002**, *21* (58), 8935-8948. DOI: 10.1038/sj.onc.1205996.

(4) (a) Brooks, S. C.; Adhikary, S.; Rubinson, E. H.; Eichman, B. F. Recent advances in the structural mechanisms of DNA glycosylases. *Biochim. Biophys. Acta Proteins Proteom.* **2013**, *1834* (1), 247-271. DOI: 10.1016/j.bbapap.2012.10.005. (b) Parikh, S. S.; Walcher, G.; Jones, G. D.; Slupphaug, G.; Krokan, H. E.; Blackburn, G. M.; Tainer, J. A. Uracil-DNA glycosylase–DNA substrate and product structures: Conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl. Acad. Sci. USA* **2000**, *97* (10), 5083-5088. DOI: 10.1073/pnas.97.10.5083. (c) Stivers, J. T.; Jiang, Y. L. A Mechanistic Perspective on the Chemistry of DNA Repair Glycosylases. *Chem. Rev.* **2003**, *103* (7), 2729-2760. DOI: 10.1021/cr010219b.

(5) (a) Nilsen, H.; Yazdankhah, S. P.; Eftedal, I.; Krokan, H. E. 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 Lett.* **1995**, *362* (2), 205-209. DOI: Doi 10.1016/0014-5793(95)00244-4. (b) Eftedal, I.; Guddal, P. H.; Slupphaug, G.; Volden, G.; Krokan, H. E. Consensus Sequences for Good and Poor Removal of Uracil from Double-Stranded DNA by Uracil-DNA Glycosylase. *Nucleic Acids Res* **1993**, *21* (9), 2095-2101. DOI: DOI 10.1093/nar/21.9.2095. (c) Slupphaug, G.; Eftedal, I.; Kavli, B.; Bharati, S.; Helle, N. M.; Haug, T.; Levine, D. W.; Krokan, H. E. Properties of a recombinant human uracil-DNA

glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. *Biochemistry* **1995**, *34* (1), 128-138. DOI: 10.1021/bi00001a016. (d) Hölz, K.; Pavlic, A.; Lietard, J.; Somoza, M. M. Specificity and Efficiency of the Uracil DNA Glycosylase-Mediated Strand Cleavage Surveyed on Large Sequence Libraries. *Sci. Rep.* **2019**, *9* (1), 17822. DOI: 10.1038/s41598-019-54044-x. (e) Bellamy, S. R. W.; Baldwin, G. S. A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1. *Nucleic Acids Res* **2001**, *29* (18), 3857-3863. DOI: 10.1093/nar/29.18.3857. (f) Orndorff, P. B.; Poddar, S.; Owens, A. M.; Kumari, N.; Ugaz, B. T.; Amin, S.; Van Horn, W. D.; van der Vaart, A.; Levitus, M. Uracil-DNA glycosylase efficiency is modulated by substrate rigidity. *Sci. Rep.* **2023**, *13* (1), 3915. DOI: 10.1038/s41598-023-30620-0.

(6) Orndorff, P. B.; van der Vaart, A. Systematic assessment of the flexibility of uracil damaged DNA. *J. Biomol. Struct. Dyn.* **2023**, *in press*. DOI: 10.1080/07391102.2023.2217683.

(7) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Comput. Biol.* **2017**, *13* (7), e1005659. DOI: 10.1371/journal.pcbi.1005659.

(8) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926-935. DOI: 10.1063/1.445869.

(9) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints – Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327-341. DOI: 10.1016/0021-9991(77)90098-5.

(10) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089-10092. DOI: 10.1063/1.464397.

(11) Galindo-Murillo, R.; Robertson, J. C.; Zgarbová, M.; Šponer, J.; Otyepka, M.; Jurečka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12* (8), 4114-4127. DOI: 10.1021/acs.jctc.6b00186.

(12) (a) Fadda, E.; Pomes, R. 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 Res* **2011**, *39* (2), 767-780. DOI: 10.1093/nar/gkq812. (b) Yin, Y. D.; Yang, L. J.; Zheng, G. Q.; Gu, C.; Yi, C. Q.; He, C.; Gao, Y. Q.; Zhao, X. S. Dynamics of spontaneous flipping of a mismatched base in DNA duplex. *Proc. Natl. Acad. Sci. USA* **2014**, *111* (22), 8043-8048. DOI: 10.1073/pnas.1400667111. (c) Fuxreiter, M.; Luo, N.; Jedlovszky, P.; Simon, I.; Osman, R. Role of Base Flipping in Specific Recognition of Damaged DNA by Repair Enzymes. *J. Mol. Biol.* **2002**, *323* (5), 823-834. DOI: 10.1016/S0022-2836(02)00999-3. (d) Priyakumar, U. D.; MacKerell, A. D. Computational Approaches for Investigating Base Flipping in Oligonucleotides. *Chem. Rev.* **2006**, *106* (2), 489-505. DOI: 10.1021/cr040475z.
(13) Banavali, N. K.; MacKerell, A. D. Free Energy and Structural Pathways of Base Flipping in a DNA GCGC Containing Sequence. *J. Mol. Biol.* **2002**, *319* (1), 141-160. DOI: 10.1016/S0022-2836(02)00194-8.
(14) Galindo-Murillo, R.; Roe, D. R.; Cheatham, T. E. Convergence and reproducibility in molecular dynamics simulations of the DNA duplex d(GCACGAACGAACGAACGC). *Biochim. Biophys. Acta Gen. Subj.* **2015**, *1850* (5), 1041-1058. DOI: 10.1016/j.bbagen.2014.09.007.
(15) Torrie, G. M.; Valleau, J. P. Non-physical sampling distributions in Monte-Carlo free-energy estimation - Umbrella sampling. *J. Comput. Phys.* **1977**, *23* (2), 187-199. DOI: 10.1016/0021-9991(77)90121-8.
(16) (a) Ramstein, J.; Lavery, R. Energetic coupling between DNA bending and base pair opening. *Proc. Natl. Acad. Sci. USA* **1988**, *85* (19), 7231-7235. DOI: 10.1073/pnas.85.19.7231. (b) Ma, N.; van der Vaart, A. Free Energy Coupling between DNA Bending and Base Flipping. *J. Chem. Inf. Model.* **2017**, *57* (8), 2020-2026. DOI: 10.1021/acs.jcim.7b00215.
(17) (a) Parker, J. B.; Stivers, J. T. Dynamics of Uracil and 5-Fluorouracil in DNA. *Biochemistry* **2011**, *50* (5), 612-617. DOI: 10.1021/bi101536k. (b) Carr, C. E.; Khutsishvili, I.; Gold, B.; Marky, L. A. Thermodynamic Stability of DNA Duplexes Comprising the Simplest T -> dU Substitutions. *Biochemistry* **2018**, *57* (39), 5666-5671. DOI: 10.1021/acs.biochem.8b00676.

(18) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **2012**, *8* (1), 348-362. DOI: 10.1021/ct200723y.