JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Dynamical Signature of Abasic Damage in DNA

Kristina E. Furse and Steven A. Corcelli*

Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States

Supporting Information

ABSTRACT: Time-dependent Stokes shift (TDSS) responses in proteins and DNA exhibit a broad range of long time scales (>10 ps) that are not present in bulk aqueous solution. The physical interpretation of the long TDSS time scales in biomolecular systems is a matter of considerable debate because of the many different components present in the sample (water, biomolecule, counterions), which have highly correlated motions and intrinsically different abilities to adapt to local perturbations. Here we use molecular dynamics (MD) simulations to show that the surprisingly slow (∼10 ns) TDSS response of coumarin 102 (C102), a base pair replacement, reflects a distinct dynamical signature for DNA damage. When the C102 molecule is covalently incorporated into DNA, an abasic site is created on the strand opposite the C102 probe. The abasic sugar exhibits a reversible interchange between intra- and extrahelical conformations that are kinetically stable on a nanosecond time scale. This conformational change, only possible in damaged DNA, was found to be responsible for the long time scales in the measured TDSS response. For the first time, a TDSS measurement has been attributed to a specific biomolecular motion. This finding directly contradicts the prevailing notion that the TDSS response in biomolecular contexts is dominated by hydration dynamics. It also suggests that TDSS experiments can be used to study ultrafast biomolecular dynamics that are inaccessible to other techniques.

Time-dependent Stokes shift (TDSS) experiments interrogate the collective dynamics of a molecular environment responding to the electronic excitation of a fluorescent probe molecule. A tremendous amount of experimental, theoretical, and simulation effort to understand the physical origins of the TDSS response in single-component liquids has resulted in substantial insight into the dynamic behavior of bulk solvents.[1,2] When the same probes are incorporated into proteins and DNA, TDSS responses exhibit a broad range of long time scales not found in bulk aqueous solution, indicating dynamic processes occurring on much slower time scales (>10 ps).[3−7] Unlike neat liquids, biomolecular systems contain many different components (water, biomolecule, counterions) that have highly correlated motions and intrinsically different abilities to adapt to the change in charge distribution of the probe molecule induced by its electronic excitation. As a result, the physical meaning of the long TDSS time scales is a matter of considerable debate,[8] and dramatic differences in experimental results, such as the 3 orders of magnitude discrepancy in the TDSS response for DNA measured using base pair replacement versus groove-bound probes, are not well understood.[4−7]

TDSS experiments work by creating a perturbation in the charge distribution of a probe molecule via electronic excitation and then monitoring the maximum fluorescence emission frequency, $\nu(t)$, which shifts to the red over time as the environment equilibrates to the excited-state charge distribution. The time-dependent shift is a direct reflection of motions in the environment in the immediate vicinity of the probe. Generally, a response function is constructed,

$$S(t) = \frac{\nu(t) - \nu(\infty)}{\nu(0) - \nu(\infty)} \quad (1)$$

which exhibits several time scales as it decays to zero. For organic dyes in aqueous solution, the time scales are well understood and have been connected to specific microscopic motions: a ∼200 fs decay corresponds to librational motions as water molecules reorient their dipole moments relative to the excited-state charge distribution of the dye, and a 1−2 ps decay corresponds to a collective rearrangement of the water hydrogen-bonding network near the dye.

The much broader range of TDSS time scales that have been measured in proteins and DNA are less well understood. In proteins, the long time TDSS decay can span from ∼10 ps to ∼10 ns, with fully buried probes exhibiting longer response times than probes that are fully or partially solvent-exposed.[3] Compared to proteins, DNA has less structural and dynamic diversity, especially the small 10−40 base pair oligonucleotides that are typically studied. Yet, Berg and co-workers, who were the first to perform TDSS measurements in DNA, reported an incredible 4 decade decay, 40 fs to 40 ns, for a synthetic base pair analogue, coumarin 102 (C102), incorporated within a DNA 17-mer.[4] This is dramatically broader and slower than TDSS responses reported for DNA using other types of probes, including minor groove binders and single base analogues. Pal, Zhao, and Zewail measured two much faster time scales, 1.4 and 19 ps, for the drug Hoechst 33258 (H33258) bound to the minor groove of a dodecamer, $d(CGCAAATTTGCG)_2$.[7] Similar time scales, 1.5 and 12 ps, were reported by Pal et al. for 2-aminopurine (2AP) covalently incorporated in place of an adenine base near the center of the same DNA sequence, d(CGCA(2AP)ATTTGCG).[6] Even when Pal, Verma, and Sen increased the time range of the measurement out to 10 ns, they still saw rapid convergence to an equilibrium value after ∼100 ps for minor groove binding probe DAPI bound to a related sequence, $d(CGCGCAATTGCGCG)_2$.[5] The 40 ns decay times measured with C102−DNA have not

Received: October 28, 2010
Published: December 22, 2010

been observed with any other probe−DNA system, groove-bound or base-stacked.

MD simulations are a powerful complement to experiment for unraveling the complex interactions in multicomponent systems. Although tremendous progress has been made in utilizing MD simulations to understand TDSS measurements of DNA, agreement has not been consistent.[9] We previously used simulations to connect directly to the H33258−DNA experiments of Pal, Zhao, and Zewail, reproducing the experimental time scales and attributing the slowest dynamics to DNA motion.[10] In direct contrast, using an MD simulation of native DNA (no explicit probe) and a new strategy to decompose the total response into component contributions, Berg and co-workers attributed the 40 ns response in the C102−DNA experiments to extraordinarily slow water motion.[11] A variety of different experimental techniques, including NMR, demonstrate constricted water at the DNA interface, especially the minor groove. A significant number of these water molecules are displaced when a small molecule binds to the minor groove, which could explain the lack of slow TDSS decay times for H33258−DNA and DAPI−DNA.[5,7] The slow water interpretation does not, however, explain the lack of slow TDSS decay times for 2AP−DNA.[6] The hydration of C102−DNA and 2AP−DNA should be qualitatively similar, since both probes are covalently incorporated within the DNA base stack. Moreover, when Pal et al. altered the hydration state of 2AP−DNA by binding the drug molecule pentamidine to the minor groove, the TDSS response was nearly unchanged.[6] Resolving these discrepancies is critical if we hope to gain useful information from these experiments about dynamics in multicomponent environments.

In order to systematically investigate these questions, we have collected and analyzed a self-consistent set of explicit biomolecular simulations and solution phase controls totaling 2 μs (15 TB) of data. All MD simulations were performed with AMBER 9.0,[12] using the parm99 force field[13] with parmbsc0 nucleotide modifications,[14] and the SPC/E water model.[15] To facilitate comparison, all of our DNA simulations feature the same dodecamer, $d(\mathrm{CGCAAATTTCGC})_2$, which is simulated with[16] and without[17] H33258 bound to the minor groove, and with C102 in place of five nonterminal base pairs (Figure 1). Results for H33258 in solution and bound to DNA have been described in detail and validated against experiment previously.[10,18−20] Here we will focus on broad conclusions using the full set of simulations.

The most practical and widely used approach for computing the TDSS response from MD simulations involves the application of linear response theory to calculate the equilibrium solvation time correlation function,[2,21,22]

$$C(t) = \frac{\langle \delta\Delta E(0)\Delta E(t)\rangle_g}{C(0)} \qquad (2)$$

where $\Delta E = E_e - E_g$ is the difference in the probe−environment interaction energy with the probe electronic charge distribution in its excited and ground electronic states, and $\delta\Delta E = \Delta E(t) - \langle \Delta E\rangle_g$ is the fluctuation in ΔE over time. For the purposes of computing ΔE and performing the MD simulations, the ground and excited states are modeled classically as two different collections of atomic-centered partial charges that are precomputed from quantum chemistry calculations. Figure 2 shows the four probe models designed for this work: two based on real,

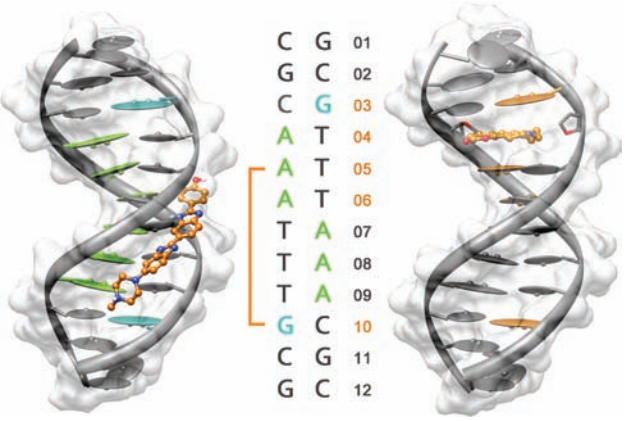

Figure 1. Schematic of probe locations in simulations. The unmodified, native DNA sequence was simulated with and without H33258 bound to the minor groove, where it contacts base pairs 05−10 (left). For the 150 ns unbound native control simulation, the eight highlighted nonterminal purine bases were used as probes. For the five 150 ns C102−DNA simulations (right), C102 was inserted in place of a purine base, and the partner pyrimidine was removed, creating an abasic site analogue. The C102 replacement sites, highlighted in orange, include unique base pairs, 03−06, plus position 10, which is identical to 03 in this palindromic sequence. The structure with C102 in position 04 is shown.

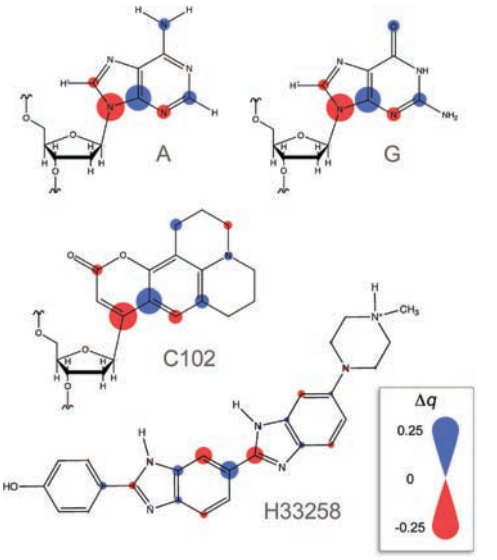

Figure 2. The four molecular probes used in this work. Probe excitation is modeled as a change in atomic charge distribution, Δq. The magnitude of Δq for individual atoms is indicated here by the size of the colored circle, red for atoms that become more negative in the excited state, blue for atoms that become more positive. H33258 and C102 are experimentally viable fluorescent probes. G and A are theoretical probes designed for this work by applying major features of the coumarin change in charge distribution to atoms of native DNA bases A and G.

experimentally viable probes, H33258 and C102, and two novel, theoretical probes created by imposing the C102 change in charge distribution, Δq, on the native purine bases, A and G. The advantage of this approach compared to other theoretical probes modeled as a single point on a native base is that spatially distributed changes in charge distribution tend to be more robust and less subject to linear response failures than single-point probes.[2,23] Since fluorescence relaxation in bulk solution is

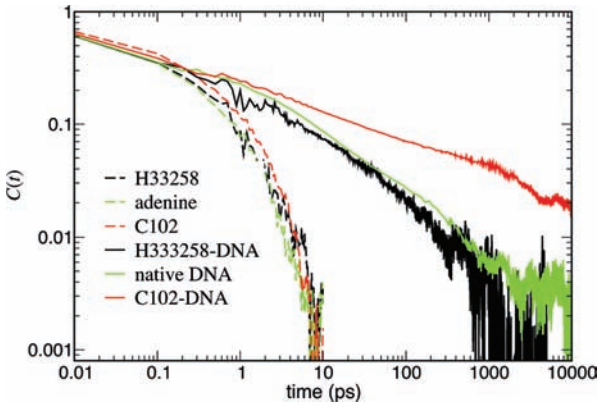

Figure 3. Collective TDSS responses modeled from MD simulations of the probes, H33258 (black) and C102 (red), free in solution and bound to DNA dodecamer, d(CGCAAATTTGCG)$_2$. The DNA bases were also used as probes in a control simulation of native DNA (green). Simulations with C102 in place of a base pair (red) are the only ones to exhibit significant long time decays beyond 1 ns.

primarily determined by solvent attributes,[2] we would expect all of the probes to give similar results when simulated free in aqueous solution. Indeed, these control simulations reveal only minor differences in the modeled TDSS responses, and all decay to zero within 5 ps (Figure 3). When the probes are incorporated into DNA, differences emerge, and we reproduce the discrepancy seen in experiment: the C102−DNA response is significantly slower than that of H33258−DNA (Figure 3), which indicates that our simulations are capturing experimentally relevant differences between the two systems. In particular, we observe an approximately power law decay in agreement with experiment.[4]

PolyAT DNA, like the sequence studied here, is characterized by an unusually narrow minor groove, which is known to contain highly organized and constrained water molecules.[17,24−28] Detailed studies of water motion in our simulations confirm the presence of highly confined minor groove water in the native DNA simulation.[19] When we bind H33258 to the minor groove, a significant fraction of this water is displaced, and when we replace a base pair with C102, the minor groove widens, accompanied by a substantial increase in water mobility. Therefore, if constrained groove-bound water were the major source of slow dynamics in DNA, we would expect the native DNA simulation to exhibit the slowest response. In fact, we see the opposite. The TDSS response for native DNA is one of the fastest—similar to the response for H33258−DNA (Figure 3). These results are consistent with the experiments of Pal et al., who reported similar TDSS responses for 2AP−DNA with and without pentamidine in the minor groove.[6] Taken together, our simulations demonstrate the presence of constricted water in specific regions of the DNA interface but do not support the idea that this water is responsible for the slow TDSS response. The slowest response observed is for the C102−DNA simulations, in which we know the water molecules are relatively more mobile.[19] It is also important to note that the C102−DNA response is not only slower than that of H33258−DNA, as predicted by experiment, but also significantly slower than the TDSS response calculated for native DNA using the purine bases, G and A, as probes. This means that the anomalous C102−DNA behavior cannot be attributed to some fundamental difference between base-stacked and groove-bound probes, as previously thought,[5,11] but rather that there is something special about C102−DNA that

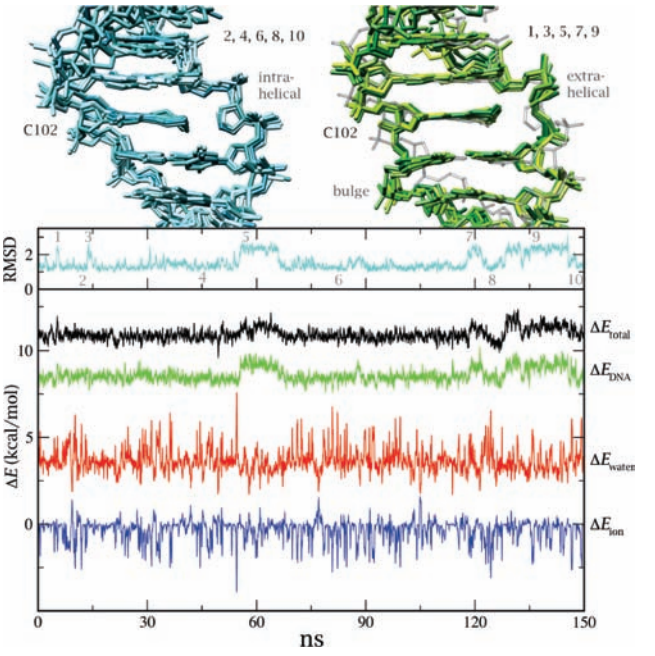

Figure 4. Physical origin of slow dynamics in C102−DNA simulations. Average structures sampled from regions of high and low rmsd comprise two conformational clusters, one in which the abasic site analogue opposite C102 is intrahelical (blue, low rmsd), and a second in which it is extra-helical (green, high rmsd), which interconvert reversibly on a tens of nanoseconds time scale. Our TDSS calculation is sensitive to this interconversion, since the same pattern appears in the time series of the DNA contribution to the instantaneous solvation energy, $\Delta E_{DNA}(t)$. By contrast, the water and ion contributions exhibit faster dynamics that are strongly anti-correlated. Decomposition of the total response into component contributions supports this interpretation; $C_{water}(t)$ is prominent at early times, but after ~100 ps, $C_{DNA}(t)$ is the primary contributor to the total response (Supporting Information). Time series data are shown at 0.1 ps resolution (1.5 million data points), smoothed with a 100 ps running average to highlight slow trends. ΔE data have been shifted vertically to facilitate comparison.

is not representative of native DNA on time scales beyond about 100 ps (Supporting Information). Our results also suggest that the impact of H33258 on the DNA dynamics reported by the TDSS measurements is modest relative to the effects caused by introducing an abasic site via C102.

The central quantity in computational TDSS studies is the instantaneous solvation energy, ΔE. The TDSS response is a collective property of the local environment as a whole, but computationally, we can investigate the individual contributions to ΔE from the components. Within the limit of pairwise additive interactions typically used in these calculations, the probe−environment interaction energy decomposes into a sum, $\Delta E = \Sigma_\alpha \Delta E_\alpha(t)$, where α represents each component present in the system (DNA, water, and counterions). Comparing the time series of the different ΔE components for the simulation in which C102 replaces the fourth base pair, d(CGC(C102)AATTTGCG), reveals a pattern of slow, tens of nanoseconds, changes in $\Delta E_{total}(t)$ that clearly originate from $\Delta E_{DNA}(t)$. This pattern also appears in a time series plot of the root-mean-square deviation (rmsd) of the instantaneous DNA conformation near the probe (Figure 4). Sampling structures from regions of high and low rmsd reveals reversible flipping of the abasic site analogue created in the partner strand across from C102 to

accommodate the probe, accompanied by lateral shifting of C102 relative to its nearest neighbor base pairs. Although we have focused our attention on C102 in the 04 position within the strand, similar reversible conformational changes were reported in the other C102−DNA simulations, but not in native DNA where the sugar is covalently tethered to a base.[19] It is important to note that the number of transitions observed in Figure 4 is insufficient to quantify the free energy difference or the kinetic rate constant for the flipping motion. Nevertheless, the correlation between the rmsd(t) and $\Delta E(t)$ is consistent with qualitatively connecting the conformational change to the long time scales in the TDSS measurements.

According to this analysis, the long time decay in our C102−DNA simulations is reporting on anomalous DNA dynamics, not slow water. Although Berg and co-workers used a different DNA sequence in their C102−DNA experiments,[4] our results suggest that the extremely long decays they found are not reporting on water behavior or even normal DNA dynamics but rather reflect a distinct dynamic signature for DNA damage in the form of an abasic site analogue. More generally, this work suggests that TDSS experiments are sensitive to biomolecular dynamics and can be used to provide valuable information on how DNA flexibility and dynamics change when subjected to damage, or to characterize other unique structures in DNA, RNA, or proteins. This sensitivity also means extreme caution must be used in designing and interpreting both experiments and calculations. Just as with bulk solvents, relating TDSS responses in these complex multicomponent systems to microscopic motions that have actual biological significance is tremendously challenging and will require a strong interaction between experiment, theory, and simulation. Continued methods development and refinement to facilitate the connection of theory to experiment are vital.

■ ASSOCIATED CONTENT

Supporting Information. Decomposition of the total solvation response, $C(t)$, into components for water, ions, DNA, and conformation of the C102 probe; absolute Stokes shifts derived from $C(t)$. Complete ref 12. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
scorcell@nd.edu

■ ACKNOWLEDGMENT

This work was supported by the National Science Foundation (CHE-0845736). The authors are also thankful for high performance computing support from the Center for Research Computing at the University of Notre Dame.

■ REFERENCES

(1) Fleming, G. R.; Cho, M. H. *Annu. Rev. Phys. Chem.* **1996,** *47*, 109.
(2) Stratt, R. M.; Maroncelli, M. *J. Phys. Chem.* **1996,** *100*, 12981.
(3) Abbyad, P.; Shi, X. H.; Childs, W.; McAnaney, T. B.; Cohen, B. E.; Boxer, S. G. *J. Phys. Chem. B* **2007,** *111*, 8269.
(4) Andreatta, D.; Lustres, J. L. P.; Kovalenko, S. A.; Ernsting, N. P.; Murphy, C. J.; Coleman, R. S.; Berg, M. A. *J. Am. Chem. Soc.* **2005,** *127*, 7270.
(5) Pal, N.; Verma, S. D.; Sen, S. *J. Am. Chem. Soc.* **2010,** *132*, 9277.
(6) Pal, S. K.; Zhao, L.; Xia, T. B.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100*, 13746.
(7) Pal, S. K.; Zhao, L. A.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100*, 8113.
(8) Halle, B.; Nilsson, L. *J. Phys. Chem. B* **2009,** *113*, 8210.
(9) Furse, K. E.; Corcelli, S. A. *J. Phys. Chem. Lett.* **2010,** *1*, 1813.
(10) Furse, K. E.; Corcelli, S. A. *J. Am. Chem. Soc.* **2008,** *130*, 13103.
(11) Sen, S.; Andreatta, D.; Ponomarev, S. Y.; Beveridge, D. L.; Berg, M. A. *J. Am. Chem. Soc.* **2009,** *131*, 1724.
(12) Case, D. A.; et al. *AMBER 9*; University of California: San Francisco, 2006.
(13) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000,** *21*, 1049.
(14) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007,** *92*, 3817.
(15) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987,** *91*, 6269.
(16) Vega, M. C.; Saez, I. G.; Aymami, J.; Eritja, R.; Vandermarel, G. A.; Vanboom, J. H.; Rich, A.; Coll, M. *Eur. J. Biochem.* **1994,** *222*, 721.
(17) Woods, K. K.; Maehigashi, T.; Howerton, S. B.; Sines, C. C.; Tannenbaum, S.; Williams, L. D. *J. Am. Chem. Soc.* **2004,** *126*, 15330.
(18) Furse, K. E.; Corcelli, S. A. *J. Chem. Theory Comput.* **2009,** *5*, 1959.
(19) Furse, K. E.; Corcelli, S. A. *J. Phys. Chem. B* **2010,** *114*, 9934.
(20) Furse, K. E.; Lindquist, B. A.; Corcelli, S. A. *J. Phys. Chem. B* **2008,** *112*, 3231.
(21) Carter, E. A.; Hynes, J. T. *J. Chem. Phys.* **1991,** *94*, 5961.
(22) Laird, B. B.; Thompson, W. H. *J. Chem. Phys.* **2007,** 126.
(23) Kumar, P. V.; Maroncelli, M. *J. Chem. Phys.* **1995,** *103*, 3038.
(24) Denisov, V. P.; Carlstrom, G.; Venu, K.; Halle, B. *J. Mol. Biol.* **1997,** *268*, 118.
(25) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981,** *151*, 535.
(26) Halle, B.; Denisov, V. P. *Biopolymers* **1998,** *48*, 210.
(27) Jana, B.; Pal, S.; Maiti, P. K.; Lin, S. T.; Hynes, J. T.; Bagchi, B. *J. Phys. Chem. B* **2006,** *110*, 19611.
(28) Liepinsh, E.; Otting, G.; Wuthrich, K. *Nucleic Acids Res.* **1992,** *20*, 6549.