JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

MD and NMR Analyses of Choline and TMA Binding to Duplex DNA: On the Origins of Aberrant Sequence-Dependent Stability by Alkyl Cations in Aqueous and Water-Free Solvents

Guillem Portella,[†,§] Markus W. Germann,[∥] Nicholas V. Hud,[*,‡] and Modesto Orozco[*,†,⊥]

[†]Joint IRB-BSC Program on Computational Biology, Institute for Research in Biomedicine, Josep Samitier 1-5 and Barcelona Supercomputing Center, Barcelona 08028, Spain

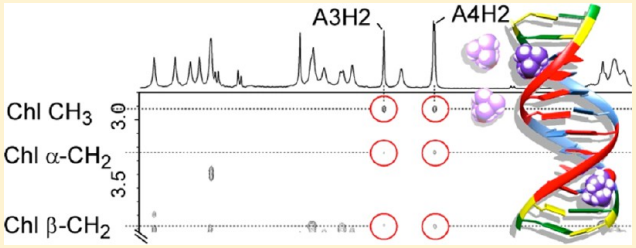
[‡]Department of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia, 30332, United States

[∥]Department of Chemistry, Georgia State University, Atlanta, Georgia, 30302, United States

[⊥]Departament de Bioquímica, Facultat de Biologia, Universitat de Barcelona, Avgda Diagonal 647, Barcelona 08028, Spain

Supporting Information

ABSTRACT: It has been known for decades that alkylammonium ions, such as tetramethyl ammonium (TMA), alter the usual correlation between DNA GC-content and duplex stability. In some cases it is even possible for an AT-rich duplex to be more stable than a GC-rich duplex of the same length. There has been much speculation regarding the origin of this aberration in sequence-dependent DNA duplex stability, but no clear resolution. Using a combination of molecular dynamics simulations and NMR spectroscopy we demonstrate that choline (2-hydroxy-*N*,*N*,*N*-trimethylethanaminium) and TMA are preferentially localized in the minor groove of DNA duplexes at A·T base pairs and these same ions show less pronounced localization in the major groove compared to what has been demonstrated for alkali and alkali earth metal ions. Furthermore, free energy calculations show that single-stranded GC-rich sequences exhibit more favorable solvation by choline than single-stranded AT-rich sequences. The sequence-specific nature of choline and TMA binding provides a rationale for the enhanced stability of AT-rich sequences when alkyl-ammonium ions are used as the counterions of DNA. Our combined theoretical and experimental study provides one of the most detailed pictures to date of cations localized along DNA in the solution state, and provides insights that go beyond understanding alkyl-ammonium ion binding to DNA. In particular, because choline and TMA bind to DNA in a manner that is found to be distinct from that previously reported for Na^+, K^+, Mg^{2+}, and Ca^{2+}, our results reveal the important but underappreciated role that most other cations play in sequence-specific duplex stability.



1. INTRODUCTION

The positive correlation between DNA duplex melting temperature and GC content has been known since the late 1950s.[1] The three hydrogen bonds of a G·C base pair, compared to two for an A·T base pair, is often cited as the origin of this differential stability, but stacking interactions[2,3] and secondary hydrogen bonds[4] also contribute to the enhanced stability of GC-rich sequences. Given the sequence dependence of cation localization along the DNA helix,[5] it is likely that cations play an important role in determining sequence-dependent duplex stability. Indeed, ∼40 years ago von Hippel et al. reported that some small alkylammonium ions, such as tetramethyl ammonium (TMA), can greatly reduce and even reverse the correlation between duplex stability with GC content.[6,7] This discovery was consistent with earlier experiments by Felsenfeld et al., which demonstrated that TMA binds to AT-rich DNA more favorably than to GC-rich DNA.[8] On the basis of these studies it was proposed decades ago that alkyl ammonium cations might preferentially bind in one of the DNA grooves in a sequence-specific manner. However, computational and experimental methods were not available at the time to directly test this hypothesis.

Recent investigations of DNA dissolved in a water-free deep eutectic solvent (DES) composed of one part choline chloride (ChCl) and two parts urea,[9] and DNA in a so-called hydrated ionic liquid (also with choline as the only cation),[10] have likewise shown a reduction in the thermal stability of duplexes with increased GC content. The close structural similarity of choline (2-hydroxy-*N*,*N*,*N*-trimethylethanaminium) and TMA immediately suggests a similar mode of interaction between these cations with DNA,[9] which could be distinct from the mode of interaction with most other cations. These investigations were motivated, in part, by the desire to expand the applicability of DNA-based technologies to nonaqueous solvents, including solvents of low volatility. In this context, understanding the interaction of choline with DNA in detail

Received: October 19, 2013
Published: February 3, 2014

   dx.doi.org/10.1021/ja410698u | *J. Am. Chem. Soc.* 2014, 136, 3075−3086

could prove essential for the design and use of complex DNA-based nanostructures in choline-based solvents.

Previous circular dichroism (CD) studies of DNA in the ChCl-urea DES showed that DNA is able to maintain a duplex secondary structure when transferred into this water-free solvent.[5] However, the helical structure of DNA was apparently altered in a sequence-dependent manner when brought into the DES. It is well-known that water activity and cations influence DNA structure and stability and that the degree of such changes can depend on nucleotide sequence,[7] but the mechanism by which choline alters the structure and stability is still not clear. DNA duplexes in solution with alkylammonium ions have been investigated by UV-monitored thermal denaturation,[9,11] CD spectroscopy,[9] and calculations,[12] but there has not been sufficient data to establish the sequence-dependent nature of alkyl ammonium ion association with DNA.

Here we report a comprehensive study of choline localization and exchange dynamics with DNA duplexes of different sequences. A combination of results from extensive molecular dynamics (MD) simulations and NMR spectroscopy provides an in-depth view of the sequence-specific association of choline and related cations. We have also analyzed the subtle structure changes of the DNA helix when moved from Na^+ to alkyl ammonium counterions (choline and TMA) by both MD and NMR spectroscopy. Overall, our results provide a detailed picture of how the structure of choline and TMA cause these cations to preferentially bind in the minor groove of AT-rich DNA duplexes, which must be considered as an integral part of any explanation for the observed increase in stability of these DNA duplexes in water-free deep eutectic and aqueous solvents where choline or TMA are the primary counterions. Additionally, MD calculations suggest that the decreased stability of GC-rich DNA duplexes in solutions containing high concentrations of choline as the only cation is due to less favorable interactions of these counterions with GC-rich duplexes and more favorable solvation of GC-rich single-stranded oligonucleotides as compared to AT-rich duplexes and single-stranded oligonucleotides, respectively. Our results therefore have important implications regarding the fundamental contribution that base-specific cation interactions provide to the sequence-dependent DNA stability.

2. RESULTS AND DISCUSSION

MD Simulations Reveal Sequence-Dependent and Groove-Specific Localization of Choline. A series of duplex DNA dodecamers with various self-complementary sequences were subjected to extensive MD simulations with Na^+, choline, or TMA as the only cationic species. The nucleotide sequences of these dodecamers varied in GC content from 33% to 100% (Table 1). All of these duplexes have two G·C base pairs at their termini to limit end fraying; those with the lowest GC content contain a core sequence comprising only A·T base pairs. MD-averaged structures of these duplexes with choline as the DNA counterion have global helical parameters that are very similar to the structures obtained for the same sequences with Na^+ as the DNA counterion (Table 1). This result is consistent with previous experimental results, which indicated that DNA duplexes of various GC content exhibit very similar CD spectra for DNA in solutions containing 3.7 M NaCl or 3.7 M ChCl.[9] However, CD spectroscopy is not able to reveal minor changes in DNA structure, which are evident in MD samplings and in detailed analysis of NMR data (see below).

Table 1. Average Helical Parameters of Duplexes Subjected to 0.5/0.4 μs Molecular Dynamics Simulations in the Presence of Choline/Na^+ Ions (full details in Table S1 in SI)[a]

sequence	cation	twist (deg)	tilt (deg)	roll (deg)
D1	choline	33.7	1.2	2.9
GCAAAATTTTGC	Na^+	33.8	1.1	2.8
D2	choline	34.1	2.7	3.2
GCTTTTAAAAGC	Na^+	33.8	2.1	2.5
D3	choline	30.9	0.4	3.9
GCGGGGCCCCGC	Na^+	31.4	0.0	5.5
D4	choline	33.3	1.3	3.0
GCAGAATTCTGC	Na	32.7	0.8	3.1
D5	choline	29.0	0.9	3.9
GCAGAGCTCTGC	Na^+	32.7	1.1	4.3
D6	choline	31.7	0.8	4.9
GCGGAATTCCGC	Na^+	31.8	0.5	3.9
D7	choline	33.4	1.0	3.0
GCAAGGCCTTGC	Na^+	32.7	0.8	1.8
D8	choline	33.8	1.1	2.8
GCAAGATCTTGC	Na^+	33.6	1.1	2.2

[a]All values for rise are 0.33−0.34 nm, shift values range from 0.01 to 0.03 nm, slide values are around −0.03 nm; sugar pseudorotation angle varied between 125 and 136° (average for all nucleotides in a duplex), but only 3° or less for any given sequence; standard deviations for each parameter, as well as translational degrees of freedom, are given in the extended table provided in the SI (Table S2). The sequence-dependent variations of the helical parameters are reported in Figure S2a−c in SI.

Our MD simulations show that all DNA duplexes are stable in all the conditions simulated (Table S1, Figure S1, Supporting Information [SI]), that choline interactions with duplex DNA are distinct from those of Na^+, and that the degree of these differences depends strongly on DNA sequence.

Plots of cation localization, as calculated by a radial distribution function (RDF, Methods) applied to MD trajectories, are shown in Figure 1 for the DNA sequences of Table 1. When duplexes D1−D8 are compared against the same RDF scale, it is clear that Na^+ localization is most favorable in the major groove at GpC steps (e.g., in D3, D5, D7; Figure 1), which is consistent with previous MD simulations.[13,14]

Reducing the RDF scale to less than one-half of that in Figure 1 reveals weaker Na^+ localization sites at the major groove side of GpG, GpA, and GpC steps, but not at steps comprising only A·T base pairs (e.g., ApA steps, SI, Figure S3). Na^+ is not localized in the minor groove as favorably as it is in the major groove.[13−16] A notable exception is the ApT step of sequence D8, a sequence element that has previously been identified as a favorable site for Na^+ localization in the minor groove.[15−17]

Choline localization by duplexes D1−D8 is clearly distinct from Na^+ localization. The most obvious difference is the more pronounced choline localization in the minor groove at sequences containing A·T base pairs. These minor groove sites are often the same as those observed for Na^+, but choline localization is typically stronger (Figure 1). Quantitatively, integration of ion density around DNA (Figure 1) shows that, on average, the binding of choline to DNA is ∼4 kJ/mol more stable than that of Na^+ (details on the binding calculations can be found in SI).

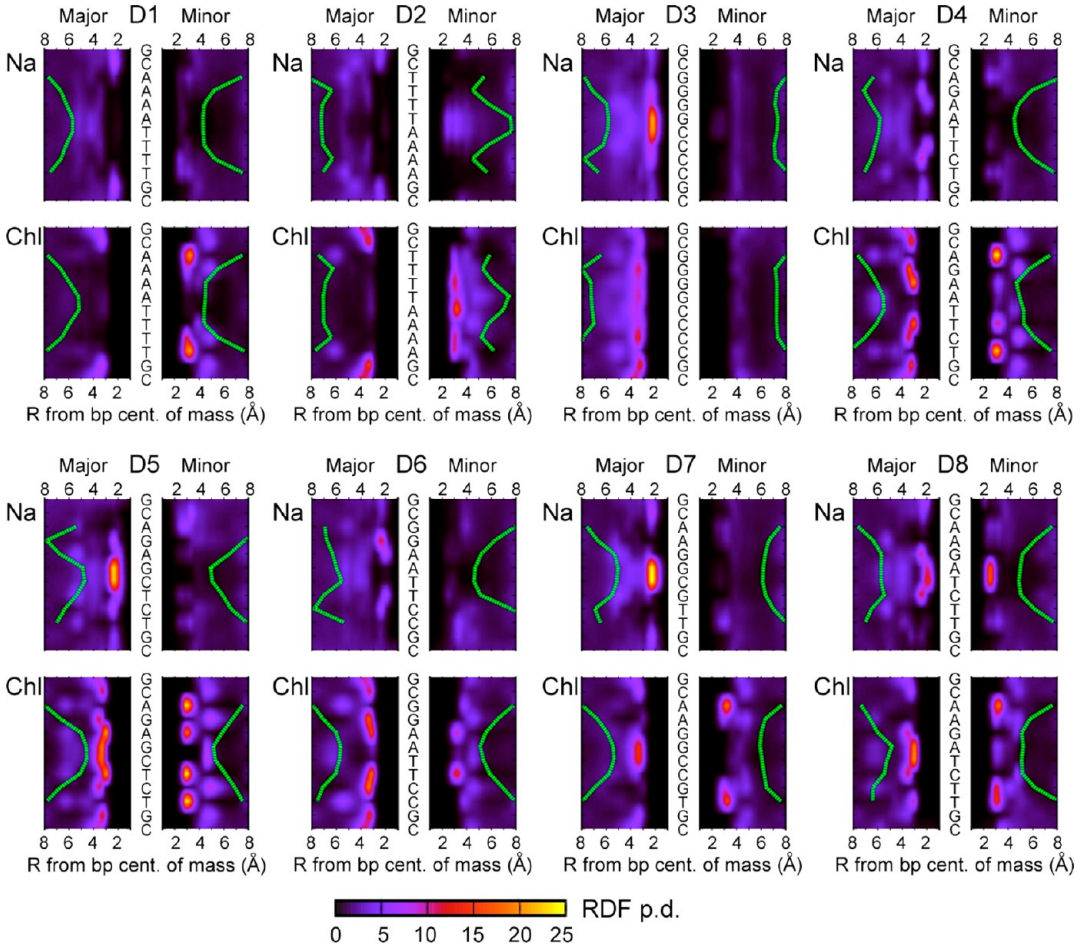


Figure 1. Choline and sodium ion binding positions along the major and minor grooves for eight different sequences studied by MD simulations. Sequences are provided along the vertical axis of each graph, in a top-to-bottom, 5′-to-3′, orientation. The two-dimensional heat maps display the radial distribution function (RDF) of particle (i.e., choline and sodium ions) density with respect to the radial position in the major and minor grooves (plotted against the horizontal axis) and the position along the duplex sequence (vertical axis). Note that the horizontal axis of the major groove is reversed from that of the minor groove. The reference of the RDF for both grooves is placed close to the center of the mass of the base pair. From these reference positions at each base pair, the density of solute was quantified as a function of the distance to the reference by searching a spherical section of 1 nm cutoff. To avoid capturing densities from both grooves, we restricted the search of solute using an angular cutoff. The green line overlaid on the maps displays the groove width (after subtracting 0.6 nm for the major groove). The profiles were not symmetrized. Further details, such as cutoff selection and grid size, can be found in the SI.

Choline localization in the major groove is most pronounced at base steps containing G·C base pairs, which is also observed for Na^+. However, relative propensity for Na^+ versus choline localization in the major groove, based on RDF signals, does not show as regular a trend as the minor groove. For example, choline and Na^+ RDF values are similar for some sequences (e.g., D5), but for other sequences Na^+ shows stronger localization (e.g., D7), while for others choline exhibits stronger major groove localization compared to Na^+ (e.g., D4, D8).

The major and minor groove widths of D1-D8 are overlaid on the RDF plots in Figure 1 (green lines). Inspection of these plots reveals that choline localization in the minor groove is correlated with nucleotide sequence composition and groove width. For example, sequences D1 and D2 have homo-AT stretches of 5′-A_4T_4-3′ and 5′-T_4A_4-3′ that localize choline in the minor groove. Both of these sequence elements are primarily composed of ApA/TpT base steps that create very electronegative minor grooves.[18] In particular, these bases lack the electropositive amine proton that is found on the minor groove edge of each G·C base pair. The electronegativity of the minor groove becomes even more pronounced as the groove becomes narrower, which can create the most favored sites for the binding of small cations (e.g., Na^+).[18] However, choline localization is seen to be strongly favored at the widest regions of these two sequence elements, near the outer regions of the 5′-A_4T_4-3′ element and near the center of the 5′-T_4A_4-3′ element. The major groove also reveals a correlation between choline localization and nucleotide sequence, but there is no clear correlation with major groove width (e.g., compare choline localization by D4 and D8).

The importance of minor groove width in localization of choline ions is further suggested by the excellent match between the size of a choline molecule and the width of the minor groove at its localization sites. The MD structure presented in Figure 2A illustrates this point, where two symmetry-related choline molecules are shown as space-filling models within the minor groove of duplex D1 (sequence GCAAAATTTTGC). These choline ions are located at positions within the minor groove with optimal width for van der Waals interactions and hydrogen bonding with the atoms of the A·T base pairs of nucleotides A3 and A4 (and symmetry-related T9 and T10). Closer to the center of the A_4T_4 sequence

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

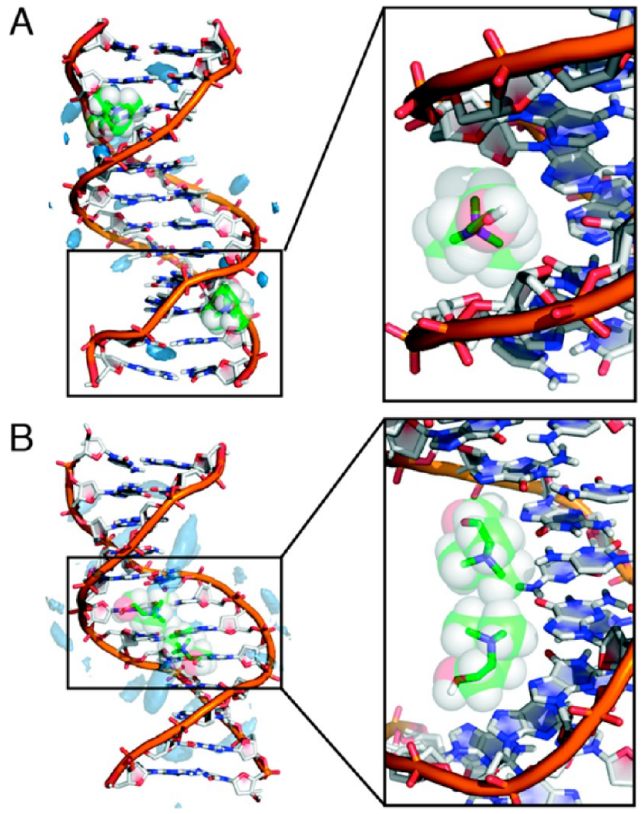


Figure 2. Examples of choline ions in their principal localization positions in (A) the minor groove of D1, sequence GCA_4T_4GC and (B) in the major groove of D3, sequence GCG_4C_4GC. Since the bulk solution is dense in choline, and the major groove is large enough, the presence of two choline ions is common, but they do exchange rapidly with bulk and they appear delocalized.

element of D1, the minor groove becomes too narrow to accommodate a choline ion without expansion of the groove.

MD simulations also reveal that the interaction of choline ions within the major groove is very different from that in the minor groove. As illustrated in Figure 2B by a structure taken from the MD trajectory of D3 (sequence GCGGGGCCCCGC), the major groove is wide enough for two choline ions to simultaneously occupy the same location along the helix, placed side-by-side along a vector that crosses the major groove perpendicular to the backbones.

Free Energy Calculations Confirm the Origins of Aberrant Stability of AT-Rich and GC-Rich Duplexes in Choline Solutions. We have computed differential duplex stability in ChCl and NaCl for two duplexes that contain different AT/GC content. For this calculation, the sequence GCAAAATTTTGC (i.e., D1) was compared with GCGAGATCTCGC, i.e. four A·T base pairs are changed to G·C base pairs, a sequence with increased GC content and the same pattern of purine and pyrimidine bases along each strand. Calculations were done in 4 M ChCl and 4 M NaCl using the molecular dynamics discrete thermodynamic integration method (MD/DTI) and thermodynamic cycles (see SI Methods and Figure S4). As expected from previous experiments in this concentration regime,[11] replacing four A·T base pairs with four G·C base pairs in 4 M NaCl increases duplex stability ($\Delta\Delta G = -9 \pm 4$ kJ/mol, SI Table S3). However, the same base pair changes are *destabilizing* in 4 M ChCl ($\Delta\Delta G = 15 \pm 3$ kJ/mol, SI Table S3).

The origins of these cation-dependent relative duplex stabilities is the result of differential cation interactions with the two sequences in both their duplex and single-stranded states. First, the differential free energy due to changes in solvent of the AT-rich sequence, with respect to the GC-rich sequence, is -17 ± 3 kJ/mol for solvation with choline ions compared to solvation with sodium ions. That is, A·T base pairs are more favorably stabilized by choline than the corresponding G·C base pairs. Second, the single-stranded state of the GC-rich sequence exhibits a differential energy of solvation by choline versus sodium ions that is -7 ± 3 kJ/mol with respect to the AT-rich sequence. That is, the single stranded GC-rich sequence exhibits more favorable solvation by choline than the single stranded AT-rich sequences. These results are fully consistent the experimental studies of sequence-dependent base pair stability in TMA solutions by Riccelli and Benight.[19]

Our MD/DTI calculations, combined with unbiased MD trajectories, further indicate that choline ions increase the stability of duplexes with A·T base pairs, relative to sodium ions, due to the tighter binding of choline to duplex DNA at A·T base pairs, mainly in the minor groove. The decrease in melting temperature of GC-rich sequences in choline solution, with respect to sodium solution, is most likely caused by a more favorable solvation of unpaired G and C bases compared to unpaired A and T bases. Specifically, MD/DTI calculations (see Table S4 in SI) show that unpaired guanosine is stabilized by 11.6 ± 1.5 kJ/mol with respect to adenosine when moving from 4 M NaCl to 4 M ChCl solution and that cytidine is slightly stabilized over thymidine by 2.2 ± 2.0 kJ/mol. Direct estimates of the solvation free energies of all four free bases in two solvents at 4 M concentration also support this conclusion (Table S4 in SI).

NMR Spectroscopy Reveals Choline Localization in the Minor Groove near A·T Base Pairs. Choline's 13 nonexchangable protons (nine methyl and four ethylene protons) provide an outstanding opportunity to study the interaction of this cation with DNA by ^{1}H NMR. In addition to being able to monitor choline proton resonances, there is also the potential to observe NOE transfers between choline ions and DNA protons, which would provide direct evidence of cations localized within the DNA grooves.

Two DNA duplexes were selected for NMR investigations from the set of sequences explored by MD simulations: D1 (sequence GCAAAATTTTGC) and D5 (sequence GCAGAGCTCTGC). These two duplexes were chosen because their MD trajectories indicated some of the better-defined sites for choline localization, in both the major and minor grooves (Figure 1). The contrast in choline localization by sequences D1 and D5 also looked potentially informative, as MD simulations of D1 revealed two choline localization sites in the minor groove, whereas the minor groove of D5 exhibited four such sites. Additionally, MD simulations of D1 did not show strong choline localization in the major groove, whereas choline localization in the major groove of D5 is among the most prominent of all sequences studied by MD.

In Figure 3 selected regions of nuclear Overhauser effect spectroscopy (NOESY) spectra are shown for D1 and D5 with choline as the only cation in the samples. These spectra contain choline−DNA cross peaks for a subset of DNA protons. For D1 and D5 the only aromatic protons showing NOE transfers from choline are AH2 protons of adenine bases. For the sugar protons, choline cross peaks are only observed to a subset of H1′ protons. The strongest NOE transfers to DNA are from

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

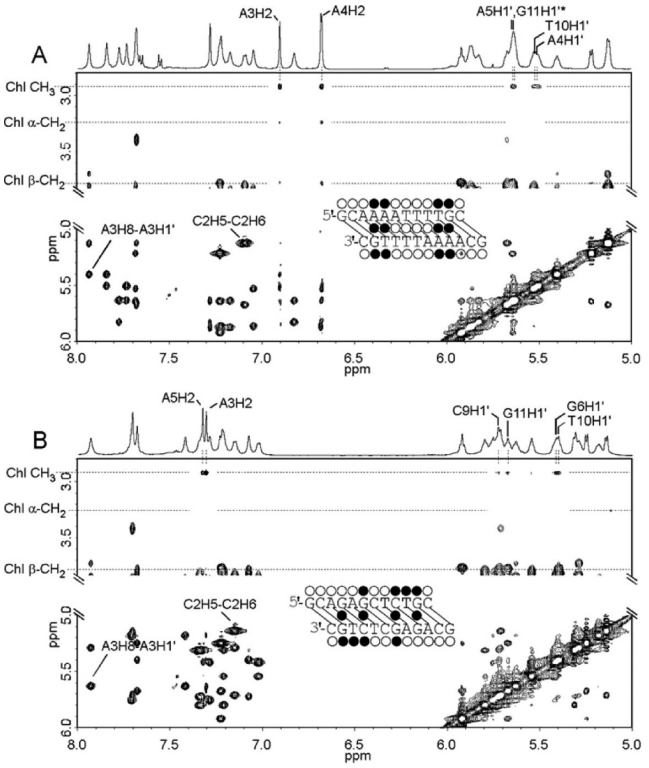


Figure 3. Selected regions of the NOESY spectra of DNA duplexes in the presence of choline. (A) Regions from the NOESY of duplex D1, GCAAAATTTTGC. (B) Regions from the NOESY of duplex D5, GCAGAGCTCTGC. Insets in A and B illustrate positions of H1′ and AH2 protons for which cross peaks to choline methyl protons are observed for each duplex (filled circles). NOESY spectra were collected at 280 K with 100 ms mixing time.

the nine equivalent methyl protons of choline (Figure 3). Smaller cross peaks are observed from the α-ethylene and the β-ethylene protons, but only to the AH2 protons of D1.

The AH2 and H1′ protons both are located in the minor groove. Schematic maps of NOE transfers to these protons on "unwound" D1 and D5 duplexes illustrate that choline has defined localization sites within the minor grooves of these two duplexes (Figure 3). In the case of D1, two well-defined sites are observed near the outermost ApA/TpT steps of the 5′-AAAATTTT-3′ sequence element. In the case of D5, two broader and less-well-defined minor groove sites are observed. A comparison of these localization sites of D1 with the RDF plots for D1 in Figure 1 shows excellent agreement between simulations and experiment. The NMR results for D5 are also in agreement with simulations, which predicted four closely spaced choline localization sites. That is, within the resolution limitations of mapping choline localization based on cross peaks to the H1′ protons, the NMR spectra reveal a wider choline localization site for D5 compared to that for D1.

It is interesting to note that NOE transfers are observed between the α and β ethanol CH_2 protons of choline to the A3H2 and A4H2 protons of D1, but NOE transfers are not observed to the AH2 protons of D5 from these same protons (Figure 3). The origin of this difference could be due to a different orientation of choline in the minor grooves of D1 and D5, or it could be due to the choline molecules in D5 being less ordered and more mobile.

In order to quantify the ability of MD simulations to reproduce NMR observables, and accordingly to define cation placement, we calculated distance-weighted probabilities for choline protons near DNA protons: $\sum p_r \cdot r^{-6}$. In these calculations the probability that a choline methyl proton is at a distance r from a DNA proton (p_r) is multiplied by radial distance between the two protons raised to the inverse sixth power, which captures the r^{-6} dependence of the NOE. As shown in Figure S4 and Table S5 in SI, simple distance-weighted probabilities predict NOE transfers from the choline methyl protons to be strongest for the A3H2 and A4H2 protons of D1. For duplex D5, A3H2 and A5H2 are predicted to have the strongest NOE transfers from choline. Reassuringly, these are the strongest DNA−choline cross peaks observed in our NOESY spectra (Figure 3). The relative magnitudes of the predicted choline methyl-AH2 NOE transfers are also in good quantitative agreement with experimental data. Specifically, for D5 the predicted intensity of the A5H2-choline NOE relative to the A3H2-choline NOE is 0.8, compared to an experimental ratio of 0.63. For D1, the predicted intensity of A4H2 relative to A3H2 is 0.44, compared to an experimental ratio of 1.0. Given the r^{-6} dependence of NOE transfers, and the highly dynamic nature of DNA−cation interactions, this agreement between experiment and simulation is surprisingly good.

Our MD simulations indicate that choline ions localized in the major groove of D5, yet no NOE cross peaks are observed between choline protons and major groove protons of this duplex. In particular, the central AGCT sequence element of D5 is predicted to localize choline, based on our RDF metric, with an occupancy that is comparable to the minor groove choline localization sites of the same duplex (Figure 1). However, no NOEs are observed in the spectra, when the noise level of the NOESY spectra presented in Figure 3 should *a priori* allow detection of NOE cross peaks that are as low as 0.2× the intensity of the choline methyl-A3H2 cross peaks of both duplexes D1 and D5. Thus, if only distance-weighted probability calculations are used as predictors, then cross peaks should be observed to at least the methyl protons of residues T8 and T10 of duplex D5 (Figure S5 in SI). If these NOE transfers do take place, their associated cross peaks are below the level of detection in our NOESY spectrum (Figure S6 in SI). Likewise, cross peaks to the other base protons, which are predicted to be of even lower intensity, are not observed.

To further enrich the synergy between the NMR and MD analysis we refined the prediction of the NOE signal for DNA protons due to choline protons by computing the rates of magnetization transfer from the spectral densities of the time-dependent dipole−dipole relaxation.[20,21] The spectral density can be obtained from the Fourier transform of the sum of autocorrelations functions for the dipole−dipole interactions between equivalent resonant spins i and j,[20] including both a distance and a rotational contribution (further details are in the SI). An example of the decay of the autocorrelation function of the dipolar interaction is shown in Figure 4A. The NOE rates computed for D1 and D5 from a 1 μs MD simulation agree well with the NMR cross-correlation signal intensities for the minor groove AH2 protons, as can be seen in Figure 4B, Table S4 in SI, validating calculations. Furthermore, the fastest dynamics of major groove choline molecules is captured by the quick decay of the autocorrelation function $C_{ij}(t)$ associated with the methyl group protons of T8 and T10, which translate into a smaller buildup rate (Figure 4), and might explain the absence of experimental detection. Similarly, the α and β protons of choline did not produce significant density on the D5 protons to clearly quantify an associated relaxation time.

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

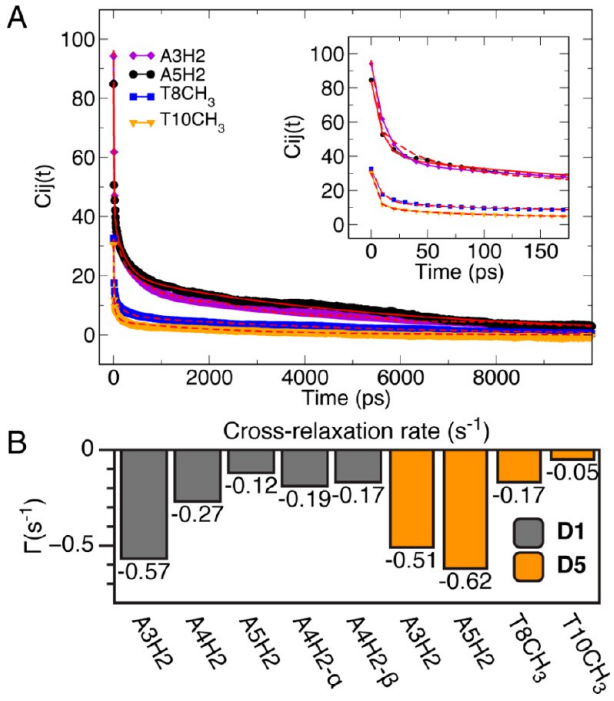


Figure 4. (A) Autocorrelation of the magnetic dipole−dipole interaction between minor groove (AH2) and major groove (TCH_3) protons with the choline methyl protons for D5. The inset shows the same data at a shorter time scale. Each curve is the result of averaging over all equivalent protons. The dashed red lines correspond to the best fit of the data to a sum of three exponentials, detailed in Table S6B in SI). (B) Bar heights represent the rate of NOE cross-correlation for minor groove (AH2) and major groove (TCH_3) protons for the sequences D1 (gray bars) and D5 (orange bars).

NOESY and Rotating Frame Nuclear Overhause Effect Spectroscopy (ROESY) Spectra Provide Insights on Choline Dynamics in the Major and Minor Grooves. For two protons sufficiently close to transfer magnetization (within 5 Å or less) the sign and maximum NOESY cross peak intensity will depend on the correlation time τ_c (or tumbling rate) of the molecule (or molecular complex) and the ^1H precession frequency, ω. If $\omega^*\tau_c \approx 1.12$ the NOE transfer will be null, regardless of inter-proton distance. At a field strength of 600 MHz (i.e., B = 14.1 T) the null NOE condition is satisfied for molecules with $\tau_c \approx 0.3$ ns; molecules with $\tau_c < 0.3$ ns will produce negative NOESY cross peaks, while molecules with $\tau_c > 0.3$ ns will produce positive cross peaks.[22]

The effect of rotational τ_c on the NOESY cross peak sign is observed in our NOESY spectra for samples containing choline; the intramolecular NOESY cross peaks of bulk choline molecules are opposite in sign to the diagonal peaks (Figure S7 in SI), which is consistent with free choline having a τ_c that is less than 0.3 ns. In contrast, the intermolecular cross peaks of DNA are positive, which is also consistent with the τ_c for a DNA dodecamer duplex being around 5 ns.[23]

The methyl protons of choline exhibit only one chemical shift even though NOE cross peaks confirm that some choline molecules are localized in the DNA minor groove. This observation indicates that choline ions localized in the DNA grooves exchange with bulk choline ions at a rate that is fast compared to the chemical shift time scale (i.e., bound lifetimes less than ∼1 ms). However, the fact that the choline−DNA NOESY cross peaks are of the same sign as the intramolecular DNA cross peaks (and opposite in sign to the intramolecular choline cross peaks) indicates that the effective τ_c of choline ions bound in the minor groove is >0.3 ns. Thus, choline ions in the minor groove must be restricted in their rotation and are temporarily assuming the tumbling rate of the duplex—also consistent with our MD results.

To obtain additional information on the dynamics of cholines bound to DNA, and to potentially observe magnetization transfer from choline protons to major groove protons, we also acquired ROESY spectra of D1 and D5 with choline. The ROESY experiment is similar to a NOESY experiment in that magnetization transfer also occurs through space by cross-relaxation between protons with a r^{-6} distance dependence.[24] However, ROESY cross peaks appear negative with respect to the resonances along the diagonal regardless of the τ_c of a molecule, and there is no null condition.

The regions of ROESY spectra of D1 and D2 that contain the choline to aromatic and H1′ cross peaks are shown in Figure 5. Similar to the NOESY spectra shown in Figure 3, the

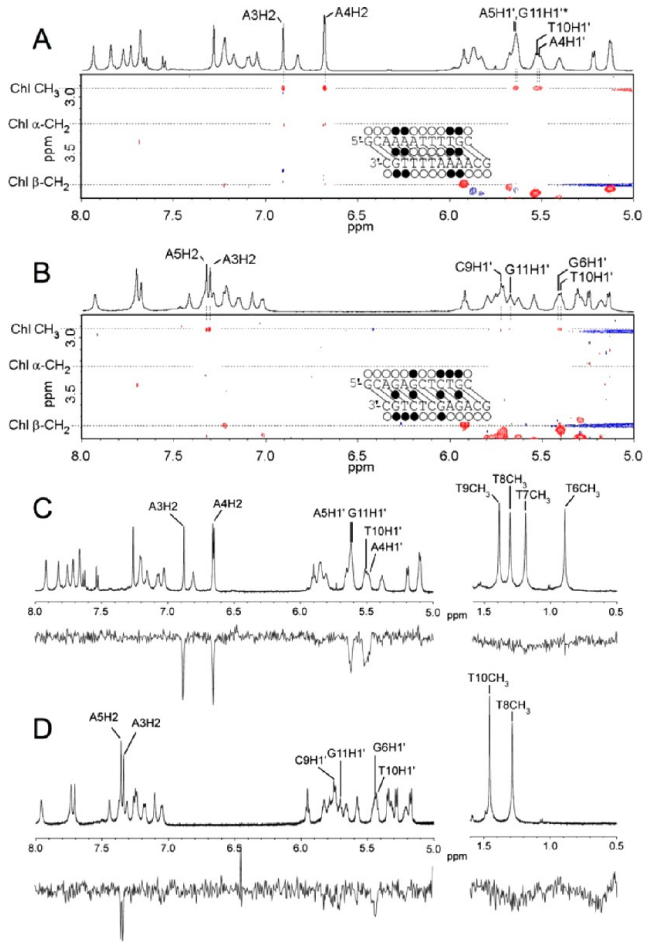


Figure 5. Selected regions of ROESY spectra for duplex D1 and D5 in the presence of choline. (A) ROESY of D1. (B) ROESY of D5. Red contours are negative with respect to diagonal, indicating direct (not relayed) magnetization transfers. ROESY mixing time 50 ms. (C, D) Rows extracted from ROESYs of duplex D1 (C) and D5 (D) at the chemical shift of the choline methyl protons. ROE transfers from choline methyl protons to AH2 and H1′ protons are similar to NOE transfers shown in Figure 3. T methyl regions are shown to illustrate that ROEs are not observed between choline methyl protons and these major groove protons.

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

ROESY spectra contain cross peaks between choline methyl protons and the AH2 and H1′ protons of D1 and D5. All of these cross peaks in the ROESY spectrum are negative, which provides additional support for the assignment of the smaller choline α- and β-CH_2-AH2 cross peaks in the NOESY spectrum as resulting from direct magnetization transfer between choline and DNA, i.e. not from chemical exchange or spin diffusion.[22] It is not possible to distinguish such cross peaks from direct magnetization transfers in a NOESY spectrum.

A row from the ROESY spectra corresponding to the chemical shift of the choline methyl protons is shown in Figure 5 for the chemical shift regions containing the aromatic/H1′ proton resonances and the methyl proton resonances. The absence of ROE transfers from choline methyl protons to T methyl protons further illustrates a significant difference between choline molecules localized in the minor groove versus any localized in the major groove; the nature of which becomes clearer with additional analysis by NMR and MD.

Because cross peak sign and magnitude are affected differently in NOESY and ROESY experiments by τ_c, it is sometimes possible to obtain information about the motion of a molecule or a complex by measuring the ratio of NOE versus ROE cross peaks for the same proton pairs. In the case of bulk choline, the ratio of intramolecular choline NOESY and ROESY cross peaks reveals that free choline molecules in our samples have a τ_c of 0.1 ns (Figure S7 in SI). Previously, this approach was used by Wüthrich and co-workers to establish limits on the bound lifetimes of water molecules associated with the surface and interior of a protein.[25] In an analogous manner, we used NOESY and ROESY data to further investigate the dynamic nature of choline ions associated with DNA. Specifically, cross peaks measured for intermolecular choline−DNA $^1H-^1H$ cross relaxation indicate that choline ions localized in the minor groove of D1 and D5 exhibit NOE:ROE peak ratios that are all consistent with a τ_c that is ≥1 ns (Figure S7 in SI).

Our data clearly indicates that the choline molecules localized in the minor groove have τ_c values that are substantially longer than the τ_c of bulk choline, which illustrates how the rotation of choline ions in the minor groove at A·T base pairs is highly restricted compared to the rotation of those in solution. As noted above, the τ_c of a dodecamer DNA duplex is expected to be on the order of 5 ns (between 3 to 7 ns according to our MD simulations, Table S7 in SI).

The dipole−dipole autocorrelation function $C_{ij}(t)$ obtained from our MD simulations can be approximated by a sum of exponential decays, with correlation times that can be attributed to different rotational and diffusive motions. We observe a fast decay, in the regime of a few tenths of picoseconds, followed by slower decay modes of hundreds of picoseconds and around 5 ns. The qualitative match between the typical rotational autocorrelation times computed for choline molecules in the grooves and free in solution indicates that the fast decay is most likely due to the rotational motion of bound choline methyl groups (Tables S6 and S7 in SI). The next decay time appears to be connected to the rotation of choline itself within the DNA grooves, and to the half-life of short-range contacts (see Figure 6 and the following section). Reassuringly, the longest decay observed is very likely due to the overall motion of DNA, as the time scales of around 5 ns coincides to the proton−proton vectors and DNA rotation. Thus, our NMR and MD measurements for the correlation times of choline molecules

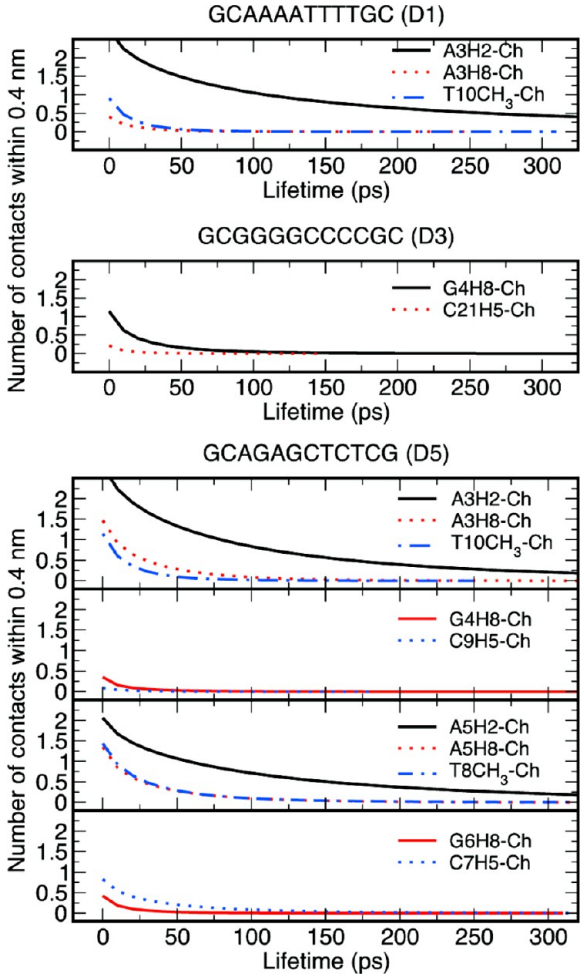


Figure 6. Contact lifetimes for choline ions within 0.4 nm of a nonexchangeable DNA proton in the major (near TCH_3 and AH8 protons) and minor grooves (near AH2 protons). The overall residence time of the choline contacts in the grooves is longer than the half-life for a short-ranged proton−proton contact, specially for the minor groove cholines; thus, the analysis reflects short-lived cations but mostly captures the choline rotating or diffusing away within the groove.

bound in the minor groove would be consistent with these molecules having their tumbling rate correlated with the tumbling rate of the DNA.

MD Reveals Sequence-Dependent Dynamics for Choline Ions Localized in the Major and Minor Grooves. We analyzed the mobility of choline ions in MD trajectories in the major and minor grooves of duplexes D1 and D5. In Figure 6 plots are presented for the lifetimes of close contacts (within 0.4 nm) of choline methyl protons and nonexchangeable base protons. As illustrated by these plots, choline ions in the major groove rarely exhibit close contacts with DNA protons that persist for more than 50 ps, whereas choline ions in the minor groove can remain within 0.4 nm of a DNA proton for more than 300 ps. Qualitative agreement between these MD-derived residence times and NMR data above clearly supports major groove choline molecules being less tightly associated with duplexes D1 and D5 compared to choline molecules bound in the minor groove.

In order to compare the dynamics of choline ions in the grooves of D1 and D5 with a duplex containing a homo-GC sequence, which are known to bind cations in the major

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

groove,[5] we present in Figure 6 a plot of residence lifetimes for choline ions in the major groove of D3 (sequence GCGGGGCCCCGC) near the GH8 and CH5 protons of a G·C base pair. While the residence lifetimes for these choline ions appear to be slightly greater compared to those of choline ions localized in the major groove of D5, they are still much less than those observed for choline ions in the minor groove. These results are consistent with the RDF plots shown in Figure 1, which are calculated on the basis of the probability of choline ions being localized near the center of mass of the Watson−Crick base pair, thereby not being biased by the different accessibility of AH2 and GH8 protons to choline ions localized in the minor and major grooves, respectively. Thus, taken together, both the dynamics and probability distribution of choline ions as analyzed by MD trajectories support greater localization (i.e., tighter binding) of choline ions in the minor groove of AT-rich DNA compared to the major groove of GC-rich DNA.

MD Simulations Reveal Subtle Changes in Minor Groove with Choline Binding. There has been much discussion about whether or not cation localization in the minor groove of A-tract DNA drives closure of the minor groove, or if the minor groove of A-tract DNA is intrinsically narrow and thereby a favorable site for cation binding.[5,15,16,18,26] We note that A-tract DNA sequences, based upon structural properties, have been defined as being homo-AT sequences greater than four base pairs in length that are composed of ApA or ApT base steps, but not TpA steps.[18,27] D1 contains the A-tract sequence A_4T_4.

To investigate whether or not the minor groove width of this sequence element is affected by choline binding, we analyzed the minimal distances across this groove at the A4·T21 base pair in simulations with choline or Na^+ as the only counterion (Figure 7). An intriguing feature is evident for the two different cations. A well-defined subpopulation with a 7 Å groove width is observed when choline is the counterion, but not when Na^+ is the counterion (Figure 7). Sorting structures from the MD trajectory of D1 according to minor groove width around the A4·T21 base pair reveals that larger groove size correlates with greater choline density (Figure S9 in SI). Similarly, the minor groove width has a probability distribution shifted toward a wider groove for those structures showing high choline occupancy at this base pair position (Figure S9 in SI). Although less dramatic, the shorter A-tract of D6 (sequence GCGGAATTCCGC) also shows a widening of the minor groove when associated with choline (Figure 7), which also appears to be the result of a choline-induced conformational selection.

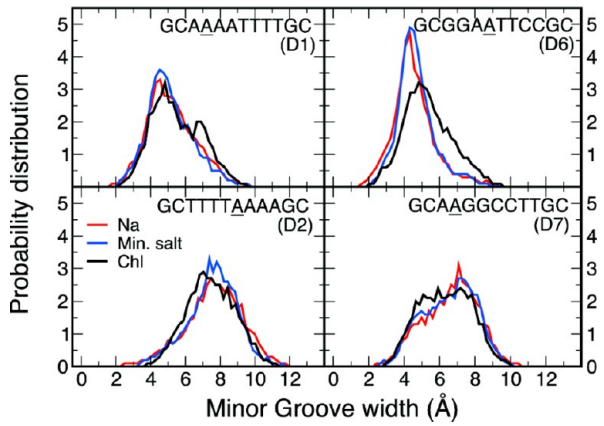


Figure 7. Changes in minor groove width probability distribution for selected choline binding positions (A·T base pairs underlined). The minimum salt (22 neutralizing Na^+) profiles show that widening of the minor groove is not directly related to variations in ionic strength. Note sequences displayed in the upper row D1 and D6 present a rather narrow minor groove in sodium solution and that it appears to widen upon choline binding (e.g., bump at 0.7 nm in D1, shifted maximum position for D6). Further correlation analyses between groove width and cation binding can be found in Figures S9 and S10 in SI.

The T_4A_4 sequence element of D2, which is not a continuous A-tract due to the central TpA step, has a considerably wider minor groove compared to the A_4T_4 sequence element of duplex D1 (Figure 1 and Figure 7), a relationship between these two sequence elements that is well-known.[18] Here we observe that the center of the T_4A_4 sequence element D2, where the minor groove is widest, narrows slightly in the presence of choline compared to the same sequence in the presence of Na^+. The same correlation is also observed for D7 (sequence GCAAGCCCTTGC) at the ApA step of this duplex (Figure 7).

The ensemble average minor and major groove widths along all eight sequences studied by MD, in choline, TMA, and Na^+ solutions are provided in Figure S11 in SI. Overall, changes in the minor groove widths are relatively small, which support the idea that the minor groove is already of near optimal size for choline binding and that when the groove is an AT-rich sequence choline binds tightly. The much closer match of choline size with minor groove width immediately suggests that choline ions in the minor groove have the potential for more favorable interactions compared to choline ions in the major groove. This supposition is given support by our NMR and MD results that indicate much greater residence times for choline ions bound in the minor groove of certain sequences. These residence times are in contrast with the more labile association of sodium and potassium ions in the minor groove.[28,29]

^{1}H and ^{31}P Chemical Shifts of GCA_4T_4GC Duplex Also Confirm Sequence and Groove-Specific Binding of Choline/TMA. Changes in ^{1}H chemical shifts can reveal very small changes in DNA helical structure at single nucleotide resolution. For example, Braunlin and co-workers showed that the partial structural conversion of the DNA duplex $(CCCCGGGG)_2$ from a pure B-form helix to an estimated 30% A-form helix was associated with changes in the cytosine H6 chemical shifts ($\Delta\delta$) between 0.06 and 0.1 ppm, and $\Delta\delta$ values of 0.06 to 0.14 ppm for the guanine H8 protons.[30]

In Figure 8 the aromatic, H1′ and methyl regions from 1D ^{1}H spectra of duplex D1 (GCAAAATTTTGC) are shown for samples containing either Na^+, choline, or TMA as the counterion of DNA. The H1′ region shows the most significant changes in chemical shifts. In particular, the H1′ resonances of residues C2, A3, and T10 exhibit the greatest change between the Na^+ and choline samples, 0.05 ppm for C2H1′, 0.07 ppm for A3H1′, and 0.1 ppm for T10H1′. For the six pyrimidine H6 protons, the two with the largest change in chemical shift are C2H6 (0.02 ppm) and T10H6 (0.02 ppm); for the six purine H8 protons the two with the most change are A3H8 (0.01 ppm) and A4H8 (0.01 ppm), and of the four H2 protons the two with the most change in chemical shift are A3H2 (0.01 ppm) and A4H2 (0.02 ppm). The similar aromatic chemical shifts and the virtually identical T methyl chemical shifts observed in the three spectra are consistent with D1

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

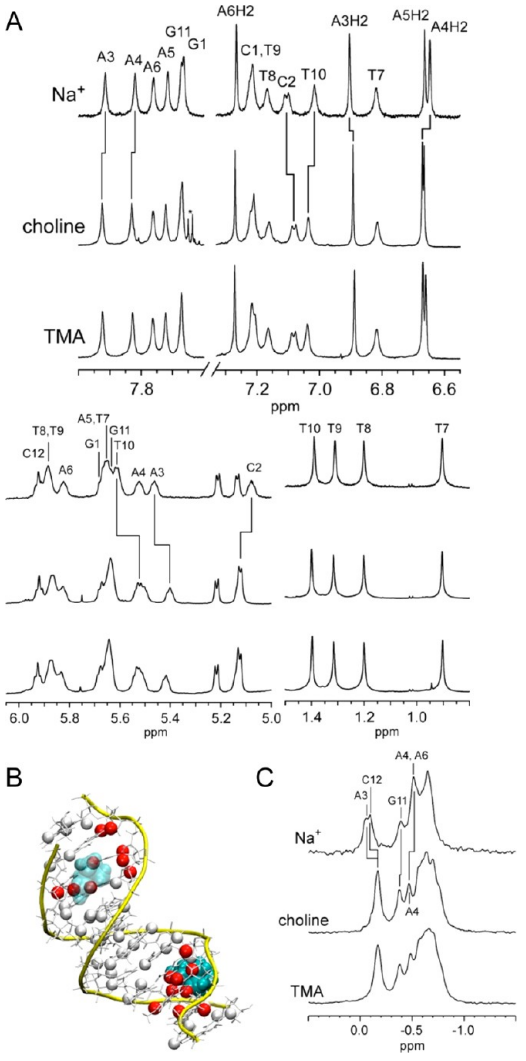


Figure 8. (A) Aromatic, H1′, and methyl region of ^{1}H spectra at 280 K for duplex D1 in the presence of Na^+, choline, or TMA. (B) MD structure of duplex D1 with bound choline molecules. Spheres indicate H6, H8, H2, and H1′ protons, with red spheres designating a subset of protons for each type that exhibit the largest changes in chemical shift (indicated in 1D spectra, chemical shift values given in text). (C) ^{31}P spectra of D1 with three different counterions.

maintaining the same overall helical structure for all three cations.

In Figure 8 we show a structure of D1 with choline, extracted from the MD trajectory with choline molecules localized at their locations of highest occupancy. On this structure all H6, H8, H2, and H1′ protons are represented as spheres, with the red spheres corresponding to the protons listed above that exhibit the greatest chemical shift differences between the Na^+ and choline sample of D1 (for each particular proton type). As can be readily appreciated, there is an excellent match between the symmetry-related locations observed by MD for choline binding to D1 and the regions of greatest changes in ^{1}H chemical shift.

In Figure 8 we also provide the aromatic, H1′, and methyl regions of duplex D1 in the presence of TMA. The chemical shifts of all protons in these regions are extremely close to those of choline, and the same cross peaks are observed in NOESY spectra between the methyl groups of TMA and the D1 protons as those identified above for the choline sample.[31] The near perfect match between the 1D and 2D ^{1}H spectra for the choline and TMA samples of D1 confirms that the interactions of these two cations with D1 are virtually identical. MD simulations also provide this observation for TMA with D1 (Figure S12 in SI). Simulations for other duplexes likewise show the same helical structures and the same positions of TMA and choline localization (Figure S12 in SI), which is consistent with the similar, yet aberrant, nature of DNA duplex stabilization by these two alkyl cations (discussed further below). The averaged MD structures for DNA in choline and TMA are very similar compared with the structures found in Na^+ solution (RMSD values between averaged structures in Table S6 in SI).

Although the TMA and choline ^{1}H spectra are essentially identical, close inspection reveals that there is a small difference in the chemical shifts of protons A3H2 and A4H2. This observation is intriguing, as the minor groove of A-tract DNA (such as AAAATTTT) is known to have several highly organized water molecules that have been named the spine of hydration.[15,32,33] Given that the MD simulations show the OH group of choline interacting with the minor groove in a similar fashion as waters in the spine of hydration (see Figure 2), it is possible that the small differences in AH2 chemical shifts between the choline and TMA samples of D1 could be due to the replacement of a water molecule in the minor groove by the OH group of choline.

While the chemical shifts of protons are very sensitive to both helical structure and solvent environment, the ^{31}P chemical shifts of DNA phosphate groups are determined more by backbone geometry (i.e., dihedral bond angles). Accordingly, ^{31}P NMR has been used in previous studies to monitor changes in DNA helical structure, including subtle changes associated with replacing Na^+ by K^+.[34] In Figure 8 we show the ^{31}P spectra of duplex D1 in the presence of Na^+, choline, and TMA. As with the ^{1}H spectra, minor differences are seen between the Na^+ and choline spectra, whereas the choline and TMA spectra are nearly identical. Of the resonances that can be assigned in both the Na^+ and the choline spectra the one that changes the most is associated with A3 (i.e., the 5′ phosphate of residue A3). Although several residues could not be assigned due to spectral overlap, their clustering within a small chemical shift range in both the Na^+ and choline/TMA samples indicates that the backbone structure around these residues changes less than the backbone around residue A3, as well as A4 and C12, which also show detectable changes in ^{31}P chemical shift.

While we can use chemical shift changes to map sites of cation association, the amount of change detected is very small and even localized to a subset of protons on a nucleotide. Thus, consistent again with MD simulations, our NMR results indicate that only very small changes in DNA structure occur upon choline binding in the minor groove. These results are also consistent with previous CD studies which have shown that the duplex formed by two strands of $(AT)_{16}$ show almost no change in helical structure when this DNA is brought from an aqueous buffer with 3.7 M NaCl into the deep eutectic solvent ChCl:urea that has choline as the only counterion at 3.7 M.[9] On the basis of these observations, and the results presented here, the duplex of $(AT)_{16}$ is expected to have choline localization sites along the minor groove that do not change appreciably upon choline binding. Moreover, the observation of choline localization sites in the minor groove of various sequences and only small changes in minor groove

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

width upon choline binding indicate that these very common minor groove binding sites for choline (and TMA) are preorganized at a width that is optimal for these cations.

Choline and TMA Interaction with Duplex DNA Is Unique among Monovalent Cations Studied Thus Far. Various experimental and theoretical investigations have demonstrated the localization of monovalent cations in the minor groove of AT-rich sequences.[5] In particular, the most favored site for Na^+, K^+, Rb^+, Tl^+, and NH_4^+ localization in a sequence element of the form A_nT_n (where $n \geq 2$) is at the central ApT step, which is also the point where the minor groove is narrowest for this A-tract sequence.[32,35−38] In contrast, here we have shown that choline and TMA are preferentially localized in the minor groove of AT-rich sequences near the widest points of the minor groove. Specifically, at the 5′-ApA and the 3′-TpT ends of A_4T_4 in D1, and at the central TpA step of T_4A_4 in D2. These favored localization sites for choline and TMA localization are the same as those identified for *divalent* cations (e.g., Mn^{2+} and Mg^{2+}) binding deep in the minor groove in the same or closely related sequence elements.[39] Since Mg^{2+} typically binds to DNA as a hexaaqua−metal complex, and choline/TMA are much larger than alkali and NH_4^+ monovalent cations, choline may be subjected to the same size-selective localization in the minor groove as hydrated divalent cations.[39,40]

To better appreciate the apparent mimicry of choline/TMA with hexaaqua-divalent cation binding the minor groove, the D1 conformation with choline bound at an ApA step was compared with the same sequence element with a bound hexaaqua−Mg^{2+} (PDB entry 1D49). These two DNA−cation complexes exhibit remarkably similar binding volumes and positioning of cations in the minor groove (Figure S13 in SI). We see from these structures that the methyl groups of choline effectively replace the waters of solvation in the minor groove that are shared by divalent ions as inner shell ligands. Thus, it appears that the same principle found for divalent cations applies to choline and TMA localization, *except that these are monovalent cations*.

X-ray crystallography and NMR solution-state studies of DNA have identified the most favored major groove sites for cation localization at GpN steps. These sites result from the electronegative O6 and N7 atoms of each guanine base which create a local electronegative patch for cation binding.[41,42] An analysis of B- and A-form duplex DNA crystal structures found that Mg^{2+} cations bound at these base steps always make water-mediated contacts with the major groove edge of G. That is, an inner-shell coordination water of Mg^{2+} acts as an anchor between the cation and the O6/N7 atoms of G. At the same location Ca^{2+} is also known to make direct, inner-shell contacts with G by partially dehydrating and utilizing O6 and N7 as inner shell ligands.[43] Likewise, monovalent cations can also partially dehydrate to accept O6 and N7 of G as inner-shell ligands. Our MD simulations do indicate that choline and TMA enter the major groove of GC-rich sequences, but these alkyl ammonium ions remain highly mobile, and apparently, unlike the alkali earth cations, the N^+ of choline/TMA and TMA cannot accept the O6 and N7 of G as stable inner shell ligands, or even share a water of hydration with G, due to the nonexchangeable methyl and ethanol groups. Consequently, choline exhibits an almost homogeneous ion density in the major groove of D3 (sequence GCGGGGCCCCGC), whereas Na^+ has a very defined site at the central GpC step (Figure 1). Thus, choline binding in the major goove of GC-rich sequences is not as well-defined as even Na^+ binding (which is less defined in its binding to DNA than divalent cations).[31] Overall, our MD and NMR studies very clearly indicate that the strength of the interaction of choline with the minor groove of A·T base pairs is generally stronger (Figures 4 and S3 in SI, free energy estimates in Table S3 in SI) than for the same ion with the major groove at G·C steps.

Finally, recent attempts to understand the nature of akyl−ammonium ion interactions with DNA in so-called hydrated ionic liquids have also been carried out using theoretical methods, but without the aid of a high-resolution experimental technique. In the study by Tateishi-Karimata and Sugimoto it was concluded, on the basis of molecular docking studies, that choline binds in the major groove at A·T base pairs and interacts with the T methyl group.[11] This proposal is not consistent with either our MD simulations or our NMR data, which is in better agreement with the binding sites derived by Senapati and co-workers from MD simulations.[44]

3. CONCLUSIONS

It has been known since the late 1960s that TMA binds more tightly to AT-rich DNA than to GC-rich DNA. In the 1970s von Hippel reported the aberrant sequence-dependent stability of DNA duplexes in solutions containing high concentrations of alkyl-ammonium ion. However, it has taken several decades for our understanding of cation−DNA interactions, the development of MD simulations, and advances in NMR instrumentation, to reach the point where we could finally provide a detailed explanation for the origin of the aberrant thermal stability of DNA duplexes in the presence of alkyl-ammonium ions.

We have shown that the akyl-ammonium cations choline and TMA act differently in their association with DNA compared to the previously reported interaction of monovalent alkali metals, NH_4^+, and the alkali earth divalent cations. These differences in groove and sequence preference provide a reasonable explanation for the origin of increased thermal stability of AT-rich duplex and reduced thermal stability of GC-rich sequences. For AT-rich sequences, the minor groove is the favored site for cation localization, and the fact that choline/TMA prefer the wider regions of an AT-rich minor groove greatly reduces the sequence restrictions for minor groove binding sites for these cations (compared to minor groove sites for alkali ions). Even a single A·T base pair is able to create a site for choline binding in the minor groove. In contrast, choline and TMA, like other cations, do not bind favorably in GC-rich minor grooves due to the presence of the electropositive hydrogen of the NH_2 group of G in the center of the minor groove. Furthermore, a peculiar property revealed for choline and TMA is their relatively weak localization at either AT-rich *or* GC-rich sequences in the major groove as compared to both alkali and alkali earth ions. Taken together, it is understandable that the substitution of akyl-ammonium cations for alkali or alkali earth cations would be expected to reduce the degree of cation association with GC-rich DNA but also with increased association with AT-rich DNA. Given that cation screening of phosphate charges is a necessary component of duplex stability, it stands to reason that GC-rich DNA duplexes would become less stable and AT-rich duplexes would become more stable with alkyl-ammonium ions as the only counterions. Indeed, our MD/DTI calculations support that the observed reversal of canonical stabilities of G·C and A·T base pairs in choline solution is explained by the aforementioned favorable

dx.doi.org/10.1021/ja410698u | J. Am. Chem. Soc. 2014, 136, 3075−3086

binding of choline in the minor grooves of A·T base pairs, together with more favorable choline solvation of unpaired guanine bases and, to some extend, unpaired cytosine bases.

In a broader context, the results presented here illustrate that the association of alkali and alkali earth cations with G·C base pairs contributes to the greater stability of G·C base pairs compared to A·T base pairs, along with the additional hydrogen bonds and greater base stacking energy.

4. METHODS

Molecular Dynamics Simulations. The DNA duplex structures and oligonucleotides were built using canonical parameters and the Nucleic Acid Builder language.[45] They were then immersed in a previously equilibrated mixture of water and solutes (choline chloride, sodium chloride, or TMA chloride) to define a 1.2 M cation concentration. Production trajectories were extended from 0.1 to 0.5 μs as indicated in Table S1 in SI. All simulations were carried out using the Gromacs-4.5 software,[46] with periodic boundary conditions and the particle mesh Ewald method for the long-range electrostatics,[47] together with a cutoff of 1.0 nm for the short-range repulsive and attractive dispersion interactions, which were modeled via a Lennard-Jones potential. The Settle algorithm was used to constrain bond lengths and angles of water molecules,[48] and P-Lincs was used for all other bond lengths, allowing a time step of 2 fs. The temperature was kept constant at 300 K by using the thermostat method of Bussi et al.[49] The pressure was controlled by coupling the simulation box to a pressure bath of 1 atm.[50] The force fields describing the interactions for DNA were generated based on the amber99SB+parmBSC0 parameters,[51,52] and we used the SPC/E model to describe the water molecules.[53] Choline ion parameters were extracted from Vrbka et al.[54] TMA ions were described following the parametrization of Heyda et al.,[55] sodium and chlorine ions, using Dang's parameters.[56]

Structural Analysis. Standard geometric analysis of the trajectories was carried out using Gromacs-4.5 program suite.[46] Curves+ was used for the helical analysis of the trajectories and for the determination of the groove widths.[57] Curves+ computes the minimum distance between phosphate groups of opposite strands, taking into account the excluded van der Waals volume of the phosphate groups.

Solvent Characterization. Ion location in the grooves was carried out by positioning the reference of a radial distribution function close to the center of mass of the base pair. To avoid capturing densities from both grooves, we restricted the search of solute using an angular cutoff (full details can be found in the SI). Solvent and cosolvent densities were obtained in a cubic grid, 0.05 nm^3 in side length, used to establish a reference density in the bulk both for the duplex DNA systems and the cosolvent/water mixtures. We characterized binding sites by detecting clusters of high-density regions. We have individually clustered the position of the cations on the binding sites using a single linkage algorithm with cutoff of 0.1 nm.

Free Energy Calculations. Choline and sodium binding free energies were computed from the densities of ions in the binding sites with respect to their densities in the bulk, corrected by the free energy required for transferring a solute molecule from a given standard volume V_0 to the binding site (full details in SI).

The relative free energy of binding of two sequences containing a different number of A·T and G·C base pairs ($d(GCAAAATTTTGC)_2$ and $d(GCGAGATCTCGC)_2$), both in sodium and choline chloride at 4 M concentrations, as well as the solvation preferences of unpaired nucleosides, were computed using thermodynamic cycles (sketch in Figure S4 in SI) and thermodynamic integration in the discrete formalism (DTI).[58] We described the interactions that are involved in the alchemical transformation between the bases/base pairs using the mixed-topology libraries of Seeliger et al.[59] adapted to the amber99SB +parmBSC0 force-field. Full details on these free energy calculations are given in the SI.

NMR Spectroscopy. The NMR data were collected at 280 K on a Bruker Avance 600 MHz NMR spectrometer equipped with a QXI probe. Acquisition and processing parameters are similar to those described in earlier studies.[60] NOESY experiments (10 × 10 ppm) were collected with 2048 × 512 data points in the two dimensions and 16 scans per t_1 increment for mixing times of 100 and 250 ms. Presaturation of the residual water peak was applied during the mixing and 3s relaxation delays. ROESY experiments were acquired with analogous parameters for 50, 250, and 500 ms spin locks. ^{31}P 1D NMR spectra (2048 scans) were recorded with ^{1}H decoupling and a capillary containing 85% H_3PO_4 was used for referencing. A sweep width of 12 ppm and a relaxation delay of 1.5 s was used together with a 45° read pulse). The ^{31}P resonances were assigned from $^{1}H-^{31}P$ correlation experiments as described previously.[61]

Sample Preparation for NMR Spectroscopy. DNA oligonucleotides were purchased from Integrated DNA Technologies, resuspended in distilled water, and passed over a column of Spehadex G-10 to remove residual protecting groups and cations. The desalted DNA was then concentrated by lyophilization and then passed over a column of Biorad AG50W cation-exchange resin that had been equilibriated with either sodium, TMA, or choline ions. Samples were then lyophilized and resuspended in D_2O three times to exchange protons for deuterons. At the last resuspension, samples were taken up in volumes that, based on UV absorbance, resulted in samples with concentrations of 2 mM in DNA strands. A comparison of integrated intensities of alkyl ammonium proton resonances with DNA resonances revealed that this sample preparation protocol resulted in a total cation concentration that was 2× the DNA nucleotide concentration.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Full details on the computational methods, six tables and thirteen figures with MD and NMR observables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

modesto.orozco@irbbarcelona.org (M.O.)
hud@gatech.edu (N.V.H.)

Present Address

[§]Department of Chemistry, Lensfield Road, Cambridge, CB2 1EW, UK.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Loren D. Williams for helpful discussions and Dr. Irena Mamajanov for preparation of DNA samples. This work was supported by the Spanish Ministry of Science and Innovation ([BIO2012-32868] and Consolider E-Science), Instituto de Salud Carlos III (INB-Genoma España and COMBIOMED RETICS), and the European Research Council (ERC). M.O is an ICREA Academia Fellow and G.P is a Sara Borrell Fellow. This research was jointly supported by the NSF and the NASA Astrobiology Program, under the NSF Center for Chemical Evolution, [CHE-1004570] (N.V.H.). The Consejo Superior de Investigaciones Científicas (CSIC) provided support for N.V.H. as a visiting scholar.

■ REFERENCES

(1) Marmur, J.; Doty, P. *Nature* **1959**, *183*, 1427−1429.
(2) Pérez, A.; Noy, A.; Lankas, F.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **2004**, *32*, 6144−6151.
(3) Parker, T. M.; Hohenstein, E. G.; Parrish, R. M.; Hud, N. V.; Sherrill, C. D. *J. Am. Chem. Soc.* **2012**, *135*, 1306−1316.
(4) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008−2010.

(5) Hud, N. V.; Engelhart, A. E. In *Nucleic Acid-Metal Ion Interactions*; Neidle, S., Hud, N. V., Eds.; The Royal Society of Chemistry: London, 2009; pp 71−113.
(6) Rees, W. A.; Yager, T. D.; Korte, J.; von Hippel, P. H. *Biochemistry* **1993**, *32*, 137−144.
(7) Melchior, W. B.; von Hippel, P. H. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 298−302.
(8) Shapiro, J. T.; Stannard, B. S.; Felsenfeld, G. *Biochemistry* **1969**, *8*, 3233−3241.
(9) Mamajanov, I.; Engelhart, A. E.; Bean, H. D.; Hud, N. V. *Angew. Chem.* **2010**, *49*, 6454−6458.
(10) Kamashev, D. E.; Mazur, A. K. *Biochemistry* **2004**, *43*, 8160−8168.
(11) Tateishi-Karimata, H.; Sugimoto, N. *Angew. Chem., Int. Ed.* **2012**, *51*, 1416−1419.
(12) Chandran, A.; Ghoshdastidar, D.; Senapati, S. *J. Am. Chem. Soc.* **2012**, *134*, 20330−20339.
(13) Auffinger, P.; Westhof, E. *J. Mol. Biol.* **2001**, *305*, 1057−1072.
(14) Auffinger, P.; Westhof, E. *J. Mol. Biol.* **2000**, *300*, 1113−1131.
(15) Pérez, A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2007**, *129*, 14739−14745.
(16) Rueda, M.; Cubero, E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2004**, *87*, 800−811.
(17) Marincola, F. C.; Denisov, V. P.; Halle, B. *J. Am. Chem. Soc.* **2004**, *126*, 6739−6750.
(18) Hud, N. V.; Plavec, J. *Biopolymers* **2003**, *69*, 144−159.
(19) Riccelli, P. V.; Benight, A. S. *Nucleic Acids Res.* **1993**, *21*, 3785−3788.
(20) Brüschweiler, R.; Wright, P. E. *Chem. Phys. Lett.* **1994**, 2614.
(21) Solomon, I. *Phys. Rev.* **1955**, *99*, 559−559.
(22) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; 2nd ed.; Wiley-VCH: New York, 2000.
(23) Aramini, J. M.; Mujeeb, A.; Germann, M. W. *Nucleic Acids Res.* **1998**, *26*, 5644−5654.
(24) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.
(25) Otting, G.; Liepinsh, E.; Wuthrich, K. *Science* **1991**, *254*, 974−980.
(26) Mazur, A. K. *J. Am. Chem. Soc.* **2000**, *122*, 12778−12785.
(27) Koo, H. S.; Wu, H. M.; Crothers, D. M. *Nature* **1986**, *320*, 501−506.
(28) Ponomarev, S. Y.; Thayer, K. M.; Beveridge, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14771−14775.
(29) Várnai, P.; Zakrzewska, K. *Nucleic Acids Res.* **2004**, *32*, 4269−4280.
(30) Xu, Q. W.; Shoemaker, R. K.; Braunlin, W. H. *Biophys. J.* **1993**, *65*, 1039−1049.
(31) *Nucleic Acid-Metal Ion Interactions*; Hud, N. V., Ed.; Royal Society of Chemistry Publishers: London, 2009.
(32) Shui, X. Q.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355.
(33) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535−556.
(34) Heddi, B.; Foloppe, N.; Hantz, E.; Hartmann, B. *J. Mol. Biol.* **2007**, *368*, 1403−1411.
(35) Hud, N. V.; Polak, M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 293−301.
(36) Woods, K. K.; McFail-Isom, L.; Sines, C. C.; Howerton, S. B.; Stephens, R. K.; Williams, L. D. *J. Am. Chem. Soc.* **2000**, *122*, 1546−1547.
(37) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 3590−3595.
(38) Hud, N. V.; Sklenar, V.; Feigon, J. *J. Mol. Biol.* **1999**, *286*, 651−660.
(39) Hud, N. V.; Feigon, F. *Biochemistry* **2002**, *41*, 9900−9910.
(40) Hud, N. V.; Schultze, P.; Feigon, J. *J. Am. Chem. Soc.* **1998**, *120*, 6403−6404.
(41) Montrel, M.; Chuprina, V. P.; Poltev, V. I.; Nerdal, W.; Sletten, E. *J. Biomol. Struct. Dyn.* **1998**, *16*, 631−637.
(42) Subirana, J. A.; Soler-Lopez, M. *Ann. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 27−45.
(43) Chiu, T. K.; Dickerson, R. E. *J. Mol. Biol.* **2000**, *301*, 915−945.
(44) Chandran, A.; Ghoshdastidar, D.; Senapati, S. *J. Am. Chem. Soc.* **2012**, *134*, 20330−20339.
(45) Macke, T., Case, D. A., Leontes, N. B., SantaLucia, J., Jr., Eds.; American Chemical Society: Washington, DC, 1998, pp 379−393.
(46) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845−854.
(47) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(48) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952−962.
(49) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101−014101.
(50) Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(51) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712−725.
(52) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., 3rd; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(53) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(54) Vrbka, L.; Jungwirth, P.; Bauduin, P.; Touraud, D.; Kunz, W. *J. Phys. Chem. B* **2006**, *110*, 7036−7043.
(55) Heyda, J.; Lund, M.; Oncak, M.; Slavicek, P.; Jungwirth, P. *J. Phys. Chem. B* **2010**, *114*, 10843−10852.
(56) Smith, D. E.; Dang, L. X. *J. Chem. Phys.* **1994**, *100*, 3757−3766.
(57) Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 5917−5929.
(58) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300−313.
(59) Seeliger, D.; Buelens, F. P.; Goette, M.; de Groot, B. L.; Grubmüller, H. *Nucleic Acids Res.* **2011**, 1−10.
(60) Aramini, J. M.; Cleaver, S. H.; Pon, R. T.; Cunningham, R. P.; Germann, M. W. *J. Mol. Biol.* **2004**, *338*, 77−91.
(61) Sklenar, V.; Miyashiro, H.; Zon, G.; Todd Miles, H.; Bax, A. *FEBS Lett.* **1986**, *208*, 94−98.