Using NMR and Molecular Dynamics to Link Structure and Dynamics Effects of the Universal Base 8-aza, 7-deaza, N8 linked Adenosine Analog
Alexander M. Spring-Connell¹, Marina G. Evich¹, Frank Seela², and Markus W. Germann¹,*

Supplemental Information

Detailed materials and methods for NMR experiments and computational methods.

NMR Experimental:

NMR experiments were conducted on a Bruker Avance 600 MHz NMR equipped with a 5mm or 8mm QXI probe, $^1$H {$^{13}$C, $^{15}$N, $^{31}$P}; HSQC and $^{13}$C $T_1$ experiments were conducted on 500 MHz Avance systems equipped with a 5mm TXI cryoprobe $^1$H{$^{13}$C, $^{15}$N}. Standard protocols were utilized for the collection of $^1$D $^1$H and $^{31}$P and 2D COSY, TOCSY, CTNOESY, HPCOR, HSQC and NOESY spectra as previously described. (1-4) Exceptions and unique details are as follows: for D$_2$O NOESY spectra, mixing times of 75, 150, and 250 ms with an 8s interscan delay were used; supercooled aqueous imino proton spectra were completed using jump-and-return for water suppression and a NOESY mixing time of 150ms. HSQC and $^{13}$C $T_1$ spectra were collected for both the sugar and base regions and each utilized an interscan delay of 5s. For $^{13}$C $T_1$ relaxation measurement, the Bruker pulse program hsqct1etgpsi was used with delays of 0.01, 0.04, 0.08, 0.120, 0.160, 0.300, and 0.500 s for the base region and 0.01, 0.05, 0.100, 0.150, 0.200, 0.300, and 0.500 for the sugar region. (5,6) Experimental peak intensities from $T_1$ spectra were fit to a standard 1st order exponential decay model in Kaleidagraph and resulting $T_1$ values were calculated. (Equation 1)

$$M_z = M_0 \ e^{(-\tau/T_1)} \quad [\text{equation 1}]$$

where $M_z$ is the measured peak intensity, $\tau$ is the inversion recovery delay, and $M_0$ is the full intensity of the resonance.

Base-pair lifetime experiments were conducted and calculated as previously described. (7,8) Briefly, for imino proton exchange measurements, the buffer was 10 mM NaP, 100 mM NaCl, with a pH of 8.0 and 8.1 for the control and UB sequences, respectively. Ammonia was added from stock solutions ranging from 1 – 5 M. Imino proton recovery time was used to calculate exchange rates as a function of the NH$_3$ concentration; 1 ms I-burp and E-burp selective pulses centered on the imino proton region were used in semi selective $T_1$ inversion recovery experiments at 3 °C for the control G6 and UB sequence G4 imino proton resonances and 15 °C for the control T14 imino proton resonance. Plots of exchange time vs 1/catalyst concentration were used to calculate the base-pair lifetime. (8)
Assignments were completed using Sparky 3.33. (9) Referencing for $^1$H and $^{31}$P was completed using internal DSS and an external capillary containing 85% $\text{H}_3\text{PO}_4$ respectively. RDC coupling values were determined for the sugar and base regions using f2 coupled HSQC spectra collected in the presence and absence of pf1 phage (ASLA) ($^2$H splitting of 18.0 and 32.2 Hz giving concentrations of 20.3 and 36.3 mg/ml for the UB and control samples respectively). (10)

rMD Structure Calculations

Standard $^1$H, $^{13}$C, and $^{31}$P assignment protocols were utilized as described previously. (2) NOE volumes were calculated in Sparky 3.33 using Gaussian or Sum-over-box methodologies. (9) Structure calculation methods using an iterative RANDMARDI process with MARDIGRAS, CORMA, and AMBER 9.0 cycles were completed as described previously. (1,2,11-16) Total $R^x$ values were calculated in CORMA using $\tau_c$ times ranging from 2.5 - 4.0 ns; the $\tau_c$ with the lowest overall $R^x$ values was 3.6 ns. Sugar puckering pseudorotation and fraction south calculations using measured 3-bond scalar couplings $^3J$ from low-flip angle COSY and $^{31}$P COSY experiments was completed using a graphical method. (17) Epsilon torsion values were calculated from the ratio of peak heights collected from CTNOESY (constant time NOESY) spectra. (4) Standard backbone torsion angle ($\alpha$, $\beta$, $\gamma$, and $\zeta$) and Watson-Crick hydrogen bonding and angle restraints were included using standard B-type DNA values. (1,18) The B-type backbone torsion angles restraints (except for the experimentally determined $\epsilon$ torsion angles) were not included for the UB5 residue. RDC restraints were measured from the base, H1', H3', and H4' regions and implemented as described previously. (2)

The AMBER parm99 force-field was appended to include specific parameters for the UB. (19) Briefly, charge derivations were completed using the R.E.D. script, which uses an iterative process to calculate atomic point charges using Gaussian03 QM (HF 6-31G*) and AMBER RESP calculations. (11,20-22) The unique C7-N8-N9 bonds were parameterized using the GAFF (General Amber Force-Field) atom definitions. (23) AMBER 9.0’s ANTECHAMBER module was used to build the force-field modification files and AMBER 9.0’s NUCGEN module was used to build the sequences for both the control and UB oligonucleotides. (11,24)

Starting structures were canonical B-type DNA with the incorporation of UB at position 5. Fully restrained molecular dynamics (rMD) were conducted in explicit water (TIP3P) (truncated octahedral box measuring 144,000 Å³) and charges were neutralized with the addition of 16 Na+ counterions. Simulation times were 10 ns using periodic boundaries and a 1 fs step size; simulations were carried out under NPT conditions (isothermal, isobaric ensemble) at 300 K and 1 atm using the SHAKE (33) algorithm for TIP3P water. The final bundle of structures was selected from the last 100 ps at a rate of 1 structure per 10 ps. Each structure was minimized using all restraints; the structure with the lowest overall AMBER penalties and
CORMA $R_\text{x}$ value was selected as the final representative structure. Structural and helical parameters were determined using the Curves+ algorithm. (25)

**Other Computational Methods**

Molecular dynamics simulations using time-averaged restraints (MDtar) and AMBER 9.0 with a $\tau = 20$ ps were completed as described previously. (26) Structures were neutralized by the addition of 16 Na$^+$ ions and solvated using a truncated octahedral box of TIP3P water that extended 10 Å from the DNA helix in all directions (144,000 Å$^3$); MDtar simulations were run for 1 ns with a 1 fs step size and periodic boundaries. Only RANDMARDI vetted distance restraints were treated with the time-averaged protocols. Broad restraints involving exchangeable protons distance restraints, 3 atom angle restraints, and 4 atom dihedral restraints were imposed under constant enforcement. RDC restraints were not implemented. Simulations were carried out under NPT conditions (isothermal, isobaric ensemble) at 300 K and 1 atm using the SHAKE (33) algorithm for TIP3P water. Analysis was conducted using Curves+. (25) The PDQPRO algorithm was used in conjunction with CORMA to select high probability structures from MDtar trajectory of structures. (27)

Unconstrained extended molecular dynamic simulations were completed as follows. DNA duplexes were initially constructed using AMBER 12 Nucleic Acid Builder (NAB) and LEap, implementing AMBER’s parmbsc0 force field, an extension of AMBER parm99, with an additional reparameterization of $\alpha/\gamma$ torsional terms (28,29). The UB parameterization was included in the parmbsc0 force field. To prevent duplexes from unraveling at the ends, an observation made by our lab in previous unrestrained MD simulations $>100$ ns and in a recent study on fraying of terminal base pairs in extended MD simulations (30), we added additional “protecting” groups consisting of G-C base pairs on each of the duplexes. These additional base pairs can reduce destabilization from better base stacking, as well as elongating the sequence. The duplexes were neutralized with Na$^+$ ions and solvated using an octahedral box of TIP3P water molecules (186,000 Å$^3$). Additional Na$^+$ and Cl$^-$ ions were added to mimic a biological environment of 100 mM NaCl. A gentle equilibration protocol, modified from a ten step process by the Orozco lab (31) was implemented to yield initial structures for the 400 ns production run. The 400 ns unrestrained molecular dynamics simulations were run using AMBER 12 NVIDIA GPU PMEMD implementation (28) using periodic boundaries and the Particle Mesh Ewald method (32) for calculating electrostatic effects (PMEMD). All production simulations were carried out NPT (isothermal, isobaric ensemble) at 300 K and 1 atm, using SHAKE (33) algorithm for TIP3P water and 2 fs time steps.

NMR chemical shift calculations were completed with Gaussian 03 using a DFT B3LYP/6-31g(d,p) basis set and the GIAO method. Calculations using the polarizable continuum model (CPCM) resulted in large deviations in calculated chemical shifts for the UB5 H7 proton whereas calculations in vacuum resulted in small changes (experimentally determined UB H7 resonances in duplex and single stranded
conditions exhibit relatively small differences in chemical shifts). Hence, all chemical shift calculations for UB5 H2 were completed in vacuum. (22)

Supplemental Table and Figure Legends:

SI Table 1. Torsion angles observed for for populations A, B, and C from the 400 ns unconstrained MD trajectory for the UB sequence. Populations percentages were determined from UB α vs. UB γ and UB α vs UB β plots where the populations were readily determined. Population percentage results between the two plots were in agreement with one another.

SI Table 2. ¹H, ¹³C, and ³¹P chemical shift resonances (in ppm) for the UB sequence.

SI Table 3. ¹H, ¹³C, and ³¹P chemical shift resonances (in ppm) for the control sequence.

SI Table 4. Intrasugar ³J_H-H coupling values, Σ1’, Σ2’, and Σ2” couplings, fraction south (fₛ), and pseudorotation angles (P) for the UB sequence.

SI Table 5. ³J_H₃'-P coupling values and epsilon torsion angles for the UB and control sequences.

SI Figure 1. NOESY spectrum (150 ms mixing time, 21 °C) of the base to H1’ region for the UB structure. Lines represent connectivity for the different strands.

SI Figure 2: Additional data for structure determination. 2A) Differences in sugar proton chemical shifts. |Δ ppm| = |control δ ppm – UB δ ppm|. Color coding is indicated on the figure. 2B) NOESY distance restraints are detailed with green lines for UB5 H7 (left) and A5 H8 (right) establishing that each is well restrained. Atom colors are as follows: blue = nitrogen, grey = carbon, red = oxygen, white = hydrogen, and gold = phosphorous.

SI Figure 3. From top to bottom: graphs for alpha, gamma, epsilon-zeta (B₁ and B₁₀), and base opening helical parameter for all residues of the solved UB structure. Right, an overlay of the final bundle of structures for the UB sequence.

SI Figure 4. NOESY integration values of the base to H1’ crosspeaks. A) UB sequence, B) control sequence. NOESY spectra were collected at 21 °C with a 75 ms mixing time. Noted in red are the A1 H8-H1’ and A1 H1’-H8 crosspeak integrations.

SI Figure 5. Low-flip angle COSY spectrum for H1’ to H2’/H2” region of the UB sequence (21 °C). The crosspeaks associated with UB5 and T14 are noted. Colors represent the positive and negative phasing resulting from the crosspeak patterns.
(Note, for clarity, crosspeaks for other resonances have not been labeled. See SI Table 2 for assignments).

SI Figure 6. HSQC spectrum for the UB sequence base region at 25 °C. The inset represents 1D slices for the adenosine H8/C8 crosspeaks as well as the UB H7/C7 crosspeak extracted from the 2D HSQC spectrum.

SI Figure 7. HSQC spectrum for the UB sequence H1'/C1' region at 25 °C. Circled in red is the T14 H1'/C1' crosspeak. Also note that the G6 H1'/C1' crosspeak exhibits a slightly reduced intensity compared to the other Gs.

SI Figure 8. NMR imino proton spectrum at -8 °C for the UB sequence (supercooled aqueous conditions). The red arrow highlights the additional resonance appearing for the T14 imino proton.

SI Figure 9. Stacked imino proton spectra at various temperatures for the UB sequence. Temperatures ranging from 1 °C to -13 °C (all supercooled aqueous conditions) are stacked showing that peaks shift in an anticipated manner indicating that there is no global change in duplex conformation.

SI Figure 10: PDQPRO selected structures from the MDtar trajectory of the UB sequence. 5A) Overlay of the PDQPRO UB5 – T14 base pair and flanking residues. Views into major groove (left) and minor groove (right). 5B) Overlay of the PDQPRO UB5 – T14 base pair conformations. Sugars and backbone are in grey. 5C) Overlay of the backbone of G4, UB5, and G6 (top to bottom) from the PDQPRO structures. Torsion angles are labeled; sugars and bases are in grey. Atom colors for all panels (aside from greyed atoms): cyan = carbon, red = oxygen, gold = phosphorus, white = hydrogen, and blue = nitrogen.

SI Figure 11. A & B) Heavy atom RMSD vs trajectory time for the 400 ns unrestrained MD calculations. Heavy atom RMSD values for the base pairs were calculated in the ptraj routine of AMBER 12.0 based on a randomly selected snapshot from the MD trajectory. A) Control sequence; G4-C15 (top), A5-T14 (middle), and G6-C13 (bottom) B) UB sequence; G4-C15 (top), UB5-T14 (middle), and G6-C13 (bottom). For both A and B, left graphs are heavy atom RMSD vs trajectory time (X axis is trajectory time in ns and Y axis is heavy atom RMSD value in Å) and right graphs are histograms for RMSD occurrence (X axis is RMSD value in Å and Y axis is occurrence from 0 to 500). C) Scatter plot of the UB5 beta torsion angle vs. UB5:T14 heavy atom RMSD (X axis is RMSD in Å, Y axis is UB5 beta torsion angle). Populations A, B, and C have been noted on the graph.

SI Figure 12: Histograms for the backbone torsion angles from the 400 ns unconstrained MD trajectories. From top to bottom: G4, UB5/A5, G6, C13, T14, and C15 residues. From left to right: alpha, beta, gamma, epsilon, and zeta torsion angles. Within each set of torsion angles, the UB sequence is on the left and the
control is on the right. X and Y axes for each histogram are: Y axis is occurrence from 0 to 1200 (bottom to top); X axis is torsion angle from 0 to 360° (left to right).

SI Figure 13. B1 / BII sampling from 400 ns unrestrained MD. A) epsilon – zeta histograms for UB sequence. X axis for each are epsilon – zeta angle from -200 to +200° (left to right); Y axis is occurrence from 0 to 1000 (bottom to top) Inset graphs are epsilon – zeta vs. trajectory time (Y axis epsilon – zeta from -200 to +200° bottom to top and X axis is trajectory time). Percent BII are noted and defined as ε - ζ > +50. B) epsilon – zeta for the control sequence. X axis for each are epsilon – zeta angle from -200 to +200° (left to right); Y axis is occurrence from 0 to 1000 (bottom to top) Percent BII are noted and defined as ε - ζ > +50. C) UB5 alpha vs. T14 epsilon – zeta scatter plot. Note, for T14 there is no apparent correlation between UB alpha and T14’s B1 / BII transitions.

SI Figure 14. A) UB5 H2 proton and the G4 and G6 carbonyls. A view of the G4, UB5, and G6 bases from the major groove. The larger black spheres represent the O6 carbonyl oxygen of G4 and G6 and the large white sphere represents the UB H2 proton. The cones are qualitative representations of the deshielding area of the carbonyl. On the left is the structure with from the solved structure; on the right is a structure from the unconstrained MD simulations. Note the different orientations of the UB5 H2 proton (white sphere) with respects to the carbonyls. B) The UB5:T14 base pair conformations from the rMD solved structure and the unconstrained MD simulation (left and right, respectively). Atom colors are as follows: cyan = carbon, blue = nitrogen, red = oxygen, white = hydrogen, and gold = phosphorous.

SI Figure 15. Simulation of two site chemical exchange at 600 MHz for a system with 1 ppm (600 Hz) separation. Both symmetric (1:1) and 8:2 population were computed using the DNMR subprogram in Topspin 3.5.

SI Figure 16. Simulation of two site chemical exchange at 400 MHz for a system with 1 ppm (400 Hz) separation. Both symmetric (1:1) and 8:2 population were computed using the DNMR subprogram in Topspin 3.5. In addition, calculated coalescence points (1:1) exchange at different field strengths with a 1 ppm separation of exchanging species are:

| Field (MHz) | k_c (s^-1) |
|-------------|------------|
| 200 MHz     | 444 s^-1   |
| 400 MHz     | 488 s^-1   |
| 600 MHz     | 1333 s^-1  |
| 800 MHz     | 1777 s^-1  |

SI Figure 17. UV absorbance traces at 260 nm for the melting studies on the UB:dT and control samples in NMR buffer conditions. See main text, table 1 for sequences. For all traces, the Y axis is absorbance at 260 nm and the X axis is the temperature in K. The top 3 traces are for the control sequence; the bottom 3 traces are for the UB sequence. Conditions are 100 mM NaCl, 10 mM NaP, and a pH of 6.8. Total strand concentrations are noted above each trace.
SI Figure 18. Representative imino proton exchange and base opening calculation graphs. A) A representative graph of the imino proton T\_1 inversion recovery data (graph is of the UB sequence G4 imino proton recovery with 6.5 mM NH\_3 at 3 °C). X axis is recovery time in seconds; Y axis is peak intensity (arbitrary units). B) A representative graph of \(\tau_{ex}\) vs. 1/catalyst. (graph is of the UB sequence G4 imino proton).

SI Figure 19. Stacked imino proton recovery spectra for the UB sequence guanosine residues. 1D NMR spectra of a subset of the imino proton region highlighting the guanosine residues where each spectrum represents a different recovery delay. The G4 and G6/G12 resonances are noted with red lines through the spectra. Note that G6/G12 exhibits a slower recovery as compared to G4.

SI Figure 20. Stacked imino proton recovery spectra for the control sequence guanosine residues. 1D NMR spectra of a subset of the imino proton region highlighting the guanosine residues where each spectrum represents a different recovery delay. The G4 and G6 resonances are noted with red lines through the spectra. Note that they have similar recovery profiles.

SI Figure 21. Representative \(^{13}\)C T\_1 data set for UB5 C1'. A) An HSQC spectrum of the UB sequence H1'/C1' region. Circled in red is the UB5 H1'/C1' crosspeak. (Note, for clarity, the other resonances have not been labeled. See SI Figure 7 and SI table 2 for resonance assignment) B) A 1D \(^{1}\)H slice of the UB5 H1'/C1' HSQC crosspeak. The intensity was used for T\_1 analysis. The same procedure was used to extract the intensities of the crosspeak for all other relaxation delays. C) T\_1 recovery data graph showing the intensities for all delays for the UB5 H1'/C1' crosspeak. Y axis is normalized peak intensity; X axis is recovery time in seconds.

SI Figure 22. Backbone structures of G4, UB5, and G6 (top to bottom) for populations A, B, and C from the long-term unconstrained MD. Representative structures were selected for populations A, B, and C. Torsion angles and populations are labeled; sugars and bases are in grey. Atom colors for the backbone are: cyan = carbon, red = oxygen, gold = phosphorus, white = hydrogen, and blue = nitrogen.

REFERENCES:

1. Aramini, J.M., Cleaver, S., Pon, R., Cunningham, R. and Germann, M. (2004) Solution Structure of a DNA Duplex Containing an [alpha]-Anomeric Adenosine: Insights into Substrate Recognition by Endonuclease IV. Journal of molecular biology, 338, 77-91.

2. Johnson, C.N., Spring, A.M., Sergueev, D., Shaw, B.R. and Germann, M.W. (2011) Structural basis of the RNase H1 activity on stereo regular borano phosphonate DNA/RNA hybrids. Biochemistry, 50, 3903-3912.
3. Johnson, C.N., Spring, A.M., Desai, S., Cunningham, R.P. and Germann, M.W. (2012) DNA sequence context conceals alpha-anomeric lesions. *J Mol Biol*, **416**, 425-437.

4. Wu, Z., Tjandra, N. and Bax, A. (2001) Measurement of 1H3’-31P dipolar couplings in a DNA oligonucleotide by constant-time NOESY difference spectroscopy. *Journal of Biomolecular NMR*, **19**, 367-370.

5. Xu, Q. and Bush, C.A. (1996) Dynamics of uniformly 13C-enriched cell wall polysaccharide of Streptococcus mitis J22 studied by 13C relaxation rates. *Biochemistry*, **35**, 14512-14520.

6. Sklenar, V., Torchia, D. and Bax, A. (1987) Measurement of Carbon-13 Longitudinal Relaxation Using 1H Detection. *Journal of Magnetic Resonance*, **73**, 375-379.

7. Mazurek, A., Johnson, C., Germann, M. and Fishel, R. (2009) Sequence context effect for hMSH2-hMSH6 mismatch-dependent activation. *Proceedings of the National Academy of Sciences*, **106**, 4177.

8. Gueron, M. and Leroy, J.L. (1995) Studies of base pair kinetics by NMR measurement of proton exchange. *Methods in enzymology*, **261**, 383-413.

9. Goddard, T. and Kneller, D. (2008) SPARKY 3

10. Zweckstetter, M. and Bax, A. (2001) Characterization of molecular alignment in aqueous suspensions of P1 bacteriophage. *J Biomol NMR*, **20**, 365-377.

11. Case, D., Darden, T., Cheatham, T., III, Simmerling, C., Wang, J., Duke, R., Luo, R., Merz, K., Pearlman, D. et al. (2006) AMBER 9.0.

12. Borgias, B.A. and James, T.L. (1988) COMATOSE, a method for constrained refinement of macromolecular structure based on two-dimensional nuclear Overhauser effect spectra. *Journal of Magnetic Resonance (1969)*, **79**, 493-512.

13. Borgias, B.A. and James, T.L. (1990) MARDIGRAS-A procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. *Journal of Magnetic Resonance (1969)*, **87**, 475-487.

14. Keepers, J.W. and James, T.L. (1984) A theoretical study of distance determinations from NMR. Two-dimensional nuclear overhauser effect spectra. *Journal of Magnetic Resonance (1969)*, **57**, 404-426.

15. Thomas, P.D., Basus, V.J. and James, T.L. (1991) Protein solution structure determination using distances from two-dimensional nuclear Overhauser effect experiments: effect of approximations on the accuracy of derived structures. *Proc Natl Acad Sci USA*, **88**, 1237-1241.

16. James, T.L. (1991) Relaxation matrix analysis of two-dimensional nuclear Overhauser effect spectra. *Current Opinion in Structural Biology*, **1**, 1042-1053.

17. Rinkel, L.J. and Altona, C. (1987) Conformational analysis of the deoxyribose furanose ring in DNA by means of sums of proton-proton coupling constants: A graphical method. *J. Biomol. Struct. Dyn.*, **4**, 621-649.

18. Mujeeb, A., Kerwin, S.M., Kenyon, G.L. and James, T.L. (1993) Solution structure of a conserved DNA sequence from the HIV-1 genome: restrained
molecular dynamics simulation with distance and torsion angle restraints derived from two-dimensional NMR spectra. *Biochemistry, 32*, 13419-13431.

19. Cheatham, T.E., 3rd, Cieplak, P. and Kollman, P.A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn, 16*, 845-862.

20. Dupradeau, F.Y., Pigache, A., Zaffran, T., Savineau, C., Lelong, R., Grivel, N., Lelong, D., Rosanski, W. and Cieplak, P. (2010) The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys Chem Chem Phys, 12*, 7821-7839.

21. Pigache, A., Cieplak, P. and Dupradeau, F.Y. (2004), 227th ACS National Meeting, Anaheim, Ca, USA.

22. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery, J.A., Vreven, T., Kudin, K.N., Burant, J.C. et al. (2003) Gaussian 03, Revision C.02.

23. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A. (2004) Development and testing of a general amber force field. *J Comput Chem, 25*, 1157-1174.

24. Wang, J., Wang, W., Kollman, P.A. and Case, D.A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model, 25*, 247-260.

25. Lavery, R., Moakher, M., Maddocks, J.H., Petkeviciute, D. and Zakrzewska, K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res, 37*, 5917-5929.

26. Aramini, J.M., Mujeeb, A., Ulyanov, N.B. and Germann, M.W. (2000) Conformational dynamics in mixed alpha/beta-oligonucleotides containing polarity reversals: a molecular dynamics study using time-averaged restraints. *J Biomol NMR, 18*, 287-302.

27. Ulyanov, N.B., Schmitz, U., Kumar, A. and James, T.L. (1995) Probability assessment of conformational ensembles: sugar repuckering in a DNA duplex in solution. *Biophys J, 68*, 13-24.

28. Case, D.A., Darden, T.A., Cheatham, T.E., Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M. et al. (2012). University of California, San Francisco.

29. Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham Iii, T.E., Laughton, C.A. and Orozco, M. (2007) Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophysical Journal, 92*, 3817-3829.

30. Zgarbová, M., Otyepka, M., Šponer, J., Lankaš, F. and Jurečka, P. (2014) Base Pair Fraying in Molecular Dynamics Simulations of DNA and RNA. *Journal of Chemical Theory and Computation, 10*, 3177-3189.

31. Shields, G.C., Laughton, C.A. and Orozco, M. (1997) Molecular Dynamics Simulations of the d(T-A·T) Triple Helix. *Journal of the American Chemical Society, 119*, 7463-7469.

32. Cheatham, T.E., III, Miller, J.L., Fox, T., Darden, T.A. and Kollman, P.A. (1995) Molecular Dynamics Simulations on Solvated Biomolecular Systems: The
Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *Journal of the American Chemical Society*, **117**, 4193-4194.

33. Ryckaert, J.-P., Ciccotti, G. and Berendsen, H.J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, **23**, 327-341.
SI Figure 3
SI Figure 4
SI Figure 5
SI Figure 6
SI Figure 7
SI Figure 8
SI Figure 9
SI Figure 11 A and B
SI Figure 11C
For all histograms: X axis is torsion angle (0 to 360 deg.); Y axis is occurrence (0 to 1200)
SI Figure 13
SI Figure 14
600MHz symmetric

600MHz 80:20
**SI Figure 17**

Absorbance at 260 nm vs. Temperature (K) for different conditions:

- **Control BorAB Duplex, Ct 20 uM**
- **Control BorAB Duplex, Ct 10 uM**
- **Control BorAB Duplex, Ct 5 uM**
- **Universal Base Duplex, Ct 20 uM**
- **Universal Base Duplex, Ct 10 uM**
- **Control BorAB Duplex, Ct 5 uM**

The graphs show the absorbance values at 260 nm plotted against temperature in Kelvin (K). Each graph includes a line for the control A at different concentrations (20, 10, and 5 uM), and a line for the universal base duplex at 20 uM. The absorbance values range from 1.3 to 1.8, and the temperature ranges from 265 to 300 K.
SI Figure 18
SI Figure 19

The diagram shows a spectrum with peaks labeled G4, G3, and G6/G12. The x-axis represents ppm values ranging from 13.4 to 12.4.
SI Figure 20
SI Figure 21
| Populations | UB5 $\alpha$ | UB5 $\beta$ | UB5 $\gamma$ | G4 $\varepsilon$ | G4 $\zeta$ | Occurrence |
|-------------|--------------|-------------|--------------|-----------------|-----------|------------|
| B           | 240-270      | 40-80       | 160-200      | 180-240         | 260-300   | 50.5%      |
| A           | 290-300      | 160-190     | 40-80        | 160-210         | 260-300   | 13.6%      |
| C           | 50-90        | 175-250     | 165-220      | 240-300         | 60-230*   | 22.4%      |
| Trans.      | 100-180      | 160-220     | 165-220      | 180-240         | 240-290   | 13.5%      |
## Universal Base Sequence

| Residue | H1' | H2' | H2'' | H3' | H4' | H2 | H5 | H6 | H7 | H8 | Methyl | P  |
|---------|-----|-----|------|-----|-----|-----|----|----|----|----|--------|----|
| A1      | 6.215 | 2.637 | 2.790 | 4.872 | 4.269 | 7.995 |   |    |    |    |        |    |
| T2      | 5.536 | 2.068 | 2.301 | 4.832 | 4.142 | 7.318 |   |    |    |    | 1.446  | -0.567|
| G3      | 5.491 | 2.653 | 2.682 | 4.960 | 4.318 | 7.841 |   |    |    |    | -0.028 |    |
| G4      | 5.714 | 2.528 | 2.861 | 4.987 | 4.336 | 7.685 |   |    |    |    | 0.020  |    |
| UBS     | 6.035 | 2.552 | 2.758 | 4.961 | 4.415 | 8.108 |   |    |    |    | -0.198 |    |
| G6      | 5.768 | 2.543 | 2.627 | 4.935 | 4.394 | 7.660 |   |    |    |    | -1.037 |    |
| C7      | 5.903 | 2.098 | 2.507 | 4.698 | 4.190 | 5.279 | 7.386 |    |    |    |        | -0.296|
| T8      | 6.141 | 2.198 | 2.540 | 4.882 | 4.180 | 7.515 | 1.671 |    |    |    |        | -0.686|
| C9      | 6.297 | 2.288 | 2.288 | 4.590 | 4.034 | 5.829 | 7.677 |    |    |    |        | -0.309|
| G10     | 5.572 | 2.512 | 2.708 | 4.843 | 4.172 | 7.888 |   |    |    |    |        |    |
| A11     | 6.072 | 2.788 | 2.927 | 5.071 | 4.441 | 7.838 | 8.208 |    |    |    |        | -0.348|
| G12     | 5.831 | 2.547 | 2.655 | 4.981 | 4.423 | 7.704 |   |    |    |    | -0.461 |    |
| C13     | 6.000 | 1.936 | 2.510 | 4.804 | 4.220 | 5.222 | 7.326 |    |    |    |        | -0.400|
| T14     | 5.771 | 2.266 | 2.327 | 4.855 | 4.195 | 7.432 | 1.594 |    |    |    |        | -0.672|
| C15     | 6.012 | 2.223 | 2.462 | 4.881 | 4.218 | 5.764 | 7.622 |    |    |    |        | 0.001 |
| C16     | 5.450 | 1.994 | 2.323 | 4.837 | 4.097 | 5.701 | 7.469 |    |    |    |        | -0.447|
| A17     | 6.328 | 2.811 | 2.886 | 5.034 | 4.422 | 7.924 | 8.362 |    |    |    |        | -0.109|
| T18     | 6.126 | 2.176 | 2.176 | 4.526 | 4.037 | 7.301 | 1.596 |    |    |    |        | -0.474|

## 13C

| Residue | C1' | C2 | C5 | C6 | C7 | C8 |
|---------|-----|----|----|----|----|----|
| A1      | 87.96 | 154.54 |   |    |    | 142.74 |
| T2      | 85.93 |       | 139.14 |   |    |    |
| G3      | 84.40 |       |     | 138.64 |   |    |
| G4      | 84.82 |       |     | 137.94 |   |    |
| UBS     | 92.92 | *   |    | 132.04 |   |    |
| G6      | 84.40 |       |     | 137.84 |   |    |
| C7      | 87.01 | 98.09 | 142.24 |   |    |    |
| T8      | 86.07 |       | 139.94 |   |    |    |
| C9      | 86.87 | 98.96 | 144.64 |   |    |    |
| G10     | 85.30 |       |     | 138.84 |   |    |
| A11     | 84.91 | 154.34 |   |    |    | 141.24 |
| G12     | 84.15 |       |     | 137.34 |   |    |
| C13     | 86.83 | 98.49 | 142.24 |   |    |    |
| T14     | 85.95 |       | 138.94 |   |    |    |
| C15     | 86.84 | 99.41 | 143.84 |   |    |    |
| C16     | 86.41 | 99.02 | 143.24 |   |    |    |
| A17     | 85.43 | 155.14 |   |    |    | 142.04 |
| T18     | 85.99 |       | 139.24 |   |    |    |

* UBS H2/C2 resonance not observed
### SI Table 3

#### Control Sequence

| Residue | H1' | H2' | H2'' | H3' | H4' | H2 | H5 | H6 | H8 | Methyl | 31P |
|---------|-----|-----|------|-----|-----|-----|-----|-----|-----|--------|-----|
| A1      | 6.218 | 2.655 | 2.809 | 4.874 | 4.270 | 7.992 | 8.170 |
| T2      | 5.611 | 2.098 | 2.366 | 4.853 | 4.163 | 7.321 | 1.429 | -0.678 |
| G3      | 5.583 | 2.647 | 2.728 | 4.984 | 4.337 | 7.834 | 8.053 | -0.236 |
| G4      | 5.555 | 2.608 | 2.745 | 5.014 | 4.375 | 7.762 | 8.053 | -0.319 |
| A5      | 6.079 | 2.647 | 2.904 | 5.051 | 4.434 | 7.699 | 8.053 | -0.484 |
| G6      | 5.708 | 2.514 | 2.592 | 4.944 | 4.375 | 7.607 | 8.053 | -0.638 |
| C7      | 5.880 | 2.094 | 2.493 | 4.661 | 4.197 | 5.205 | 7.369 | -0.298 |
| T8      | 6.154 | 2.210 | 2.544 | 4.884 | 4.193 | 7.533 | 1.646 | -0.831 |
| C9      | 6.292 | 2.285 | 2.285 | 4.587 | 4.030 | 5.811 | 7.670 | -0.425 |
| G10     | 5.570 | 2.509 | 2.704 | 4.841 | 4.171 | 7.885 | 8.190 | -0.482 |
| A11     | 6.088 | 2.772 | 2.921 | 5.067 | 4.438 | 7.854 | 8.190 | -0.607 |
| G12     | 5.794 | 2.563 | 2.634 | 4.979 | 4.420 | 7.683 | 7.377 | -0.341 |
| C13     | 5.895 | 2.083 | 2.511 | 4.673 | 4.236 | 5.214 | 7.377 | -0.863 |
| T14     | 6.091 | 2.225 | 2.570 | 4.882 | 4.235 | 7.480 | 1.570 | -0.684 |
| C15     | 5.922 | 2.146 | 2.429 | 4.835 | 4.166 | 5.654 | 7.552 | -0.383 |
| C16     | 5.502 | 1.984 | 2.338 | 4.817 | 4.104 | 5.662 | 7.470 | -0.280 |
| A17     | 6.308 | 2.775 | 2.872 | 5.013 | 4.406 | 7.884 | 8.328 | -0.607 |
| T18     | 6.116 | 2.161 | 2.161 | 4.519 | 4.024 | 7.285 | 1.561 | -0.607 |

#### Sugar $^{13}$C

| Residue | C1' | C2 | C5 | C6 | C8 |
|---------|-----|----|----|----|----|
| A1      | 87.82 | 154.60 |    |    | 142.60 |
| T2      | 85.76 | 139.10 |    |    |       |
| G3      | 84.25 | 138.50 |    |    |       |
| G4      | 84.04 | 137.70 |    |    |       |
| A5      | 84.71 | 154.40 |    |    | 140.90 |
| G6      | 83.73 | 137.10 |    |    |       |
| C7      | 87.13 | 97.83 | 142.10 |    |    |
| T8      | 86.06 | 139.90 |    |    |       |
| C9      | 86.55 | 98.77 | 144.50 |    |    |
| G10     | 84.98 | 138.80 |    |    |       |
| A11     | 84.75 | 154.50 |    |    | 141.20 |
| G12     | 83.87 | 137.20 |    |    |       |
| C13     | 87.10 | 97.78 | 142.10 |    |    |
| T14     | 85.94 | 139.80 |    |    |       |
| C15     | 86.31 | 98.82 | 143.60 |    |    |
| C16     | 86.43 | 98.83 | 143.20 |    |    |
| A17     | 85.45 | 155.00 |    |    | 142.00 |
| T18     | 85.57 | 139.30 |    |    |       |
| Residue | J1'-2' (Hz) | J1'-2'' (Hz) | Σ1' (Hz) | Σ2' (Hz) | Σ2'' (Hz) | f_s (perc.) | P (deg.) |
|---------|-------------|--------------|----------|----------|-----------|-------------|----------|
| A1      | 8.3         | 6.0          | 14.1     | 24.8     | 21.6      | 65 - 90%    | 115 - 200°|
| T2      | 9.5         | 5.6          | 15.2     | 25.9     | 20.6      | 85 - 100%   | 120 - 180°|
| G3      | Overlapped  | Overlapped   | Overlapped| Overlapped| Overlapped| Overlapped  | 150 - 175°|
| G4      | 8.6         | 6.1          | 14.6     | 27.7     | 21.9      | 75 - 90%    | 150 - 175°|
| UB5     | 8.6         | 5.9          | 14.6     | 27.4     | 22.5      | 75 - 85%    | 125 - 170°|
| G6      | 9.4         | 4.7          | 14.4     | 23.8     | 20.4      | 75 - 95%    | 145 - 170°|
| C7      | 6.1         | 6.9          | 13.0     | 27.8     | 25.8      | 40 - 60%    | 120 - 165°|
| T8      | 7.3         | 6.6          | 13.8     | 28.6     | 22.6      | 55 - 75%    | 115 - 170°|
| C9      | Overlapped  | Overlapped   | Overlapped| Overlapped| Overlapped| Overlapped  | 150 - 175°|
| G10     | 9.5         | 5.6          | 15.0     | 24.7     | 21.0      | 85 - 100%   | 100 - 205°|
| A11     | 9.7         | 5.5          | 15.2     | 24.3     | 21.3      | 85 - 100%   | 125 - 210°|
| G12     | 10.0        | 5.3          | 15.3     | 23.4     | 21.1      | 90 - 100%   | 115 - 190°|
| C13     | 8.4         | 6.2          | 14.4     | 28.8     | 21.7      | 65 - 85%    | 100 - 175°|
| T14     | 9.8         | 4.2          | 15.1     | 23.5     | 23.0      | 75 - 100%   | 100 - 215°|
| C15     | 9.7         | 5.4          | 15.0     | 26.7     | 20.6      | 80 - 100%   | 100 - 200°|
| C16     | 8.9         | 5.7          | 14.7     | 24.8     | 21.7      | 75 - 95%    | 100 - 205°|
| A17     | 8.6         | 5.9          | 14.5     | 25.7     | 21.8      | 75 - 95%    | 105 - 215°|
| T18     | Overlapped  | Overlapped   | Overlapped| Overlapped| Overlapped| Overlapped  | 150 - 175°|
| Residue | $^{3}J_{\text{H3}^\prime-p}$ (Hz) | Epsilon (deg.) |
|---------|--------------------------------|----------------|
|         | Universal | Control | Universal | Control |
| A1      | 5.1       | 4.6     | -164 ± 5° | -167 ± 10° |
| T2      | 4.7       | 4.6     | -166 ±10° | -167 ± 10° |
| G3      | 5.3       | 4.0     | -163 ± 10° | -170 ± 10° |
| G4      | 4.0       | 3.8     | -170 ± 5° | -171 ± 10° |
| UB5/A5  | 3.8       | 3.3     | -171 ±11° | -174 ± 10° |
| G6      | 4.4       | 3.8     | -168 ± 5° | -171 ± 10° |
| C7      | 4.9       | 4.5     | -165 ± 5° | -167 ± 10° |
| T8      | 5.1       | 4.2     | -164 ± 8° | -169 ± 10° |
| G10     | 5.0       | 4.3     | -165 ± 7° | -168 ± 10° |
| A11     | 3.9       | 3.3     | -170 ± 5° | -174 ± 10° |
| G12     | 3.9       | 3.7     | -171 ± 7° | -172 ± 10° |
| C13     | 4.7       | 4.2     | -166 ± 9° | -169 ± 10° |
| T14     | 5.2       | 3.1     | -164 ± 7° | -175 ± 10° |
| C15     | 4.6       | 4.3     | -167 ± 7° | -169 ± 10° |
| C16     | 5.0       | 4.4     | -165 ± 3° | -168 ± 10° |
| A17     | 4.3       | 3.7     | -169 ± 5° | -172 ± 10° |