Supplementary Material

Ribonucleotide-containing DNA – DeRose etal.

I: Details for Experimental Methods:

2D NMR Experiments

The proton 2D ^{1}H-^{1}H NOESY, ROESY and TOCSY experiments were acquired at 25 °C on a Varian (Santa Clara, CA) INOVA 600 MHz spectrometer using a Varian triple resonance Z-gradient room-temperature probe. 2D ^{1}H-^{1}H COSY experiments were also acquired on a second Varian INOVA 600 MHz spectrometer using a Varian triple resonance Z-gradient Cold Probe. The 2D ^{1}H/^{31}P COSY, H3'-selective ^{1}H/^{31}P HSQC and 2D ^{1}H-^{31}P constant-time NOESY (CT-NOESY) difference experiments were acquired at 25 °C on a Varian INOVA 500 MHz NMR spectrometer, using a Nalorac ^{1}H/X/^{31}P triple resonance actively shield Z-gradient probe. Proton and ^{31}P chemical shifts were referenced to external DSS and trimethyl phosphate, respectively.

Proton Homonuclear Experiments

2D NOESY spectra in D_2O were obtained with mixing times of 50, 75, 100, 150 and 200 ms. 2D T-ROESY experiments were acquired with mixing times of 50 and 100 ms. The T-ROESY experiment was used to reduce TOCSY contributions to the cross-peak intensity.[1,2] A z-filtered 2D TOCSY spectrum of the D_2O sample was also acquired with mixing time of 100 ms, using a DIPSI-2 spinlock sequence[3] at a field strength of 6000 Hz. All the above 2D spectra were acquired with acquisition times of 171 ms in the F2 dimension and 42.7 ms in the F1 dimension, with sweep widths of 10 ppm in both dimensions. In these experiments, the residual water peak was suppressed using presaturation during the 2 second recovery delay between scans. 2D NOESY spectra in H_2O were acquired with mixing times of 80, 120 and 200 ms, using WATERGATE water suppression.[4] 2D phase sensitive COSY experiments were measured in isotropic and liquid crystalline (20 mg/ml Pf1) media to measure the proton-proton scalar

couplings and RDCs. These COSY experiments were acquired with an acquisition time of 342 ms in the F2 dimension and 42.7 ms in the F1 dimension, with sweep widths of 10 ppm in both dimensions, using an 8 second delay between scans to allow the spin systems to fully relax for ACME simulation of the coupling constants.[5] RDCs involving the adenosine H2 protons were not used due to the long relaxation times of these protons.

^{1}H/^{31}P Experiments

All ^{1}H/^{31}P experiments were acquired on a Varian INOVA 500 MHz NMR spectrometer. A 2D ^{1}H/^{31}P COSY spectrum was acquired to assign the ^{31}P chemical shifts.[6] These assignments were confirmed by an H3'-selective ^{1}H/^{31}P HSQC experiment.[7] In the ^{1}H/^{31}P COSY experiment, the ^{1}H and ^{31}P acquisition times were 512 ms and 210 ms, respectively; in the ^{1}H/^{31}P HSQC experiments the acquisition times were 128 ms and 210 ms, respectively. In both experiments, the ^{1}H and ^{31}P sweep widths were 10 ppm and 3 ppm, respectively. H3'-^{31}P scalar couplings and RDCs were measured using 2D ^{1}H-^{31}P CT-NOESY difference experiments.[8] In these experiments, the F2 and F1 acquisition times were 204.6 ms and 69.5 ms, respectively; the sweep widths were 10 ppm and 2.5 ppm, respectively; WET water suppression was used during the 300 ms mixing time to suppress the residual water signal.[9]

Processing and Assignment of Spectra

All spectra were processed using NMRPipe[10] and assigned using NMRViewJ.[11] The spectra were processed using cosine-bell squared apodization in both dimensions and zero filling in the indirectly detected dimensions. In some cases, polynomial baseline correction was used to improve the appearance of the spectrum, but no post acquisition water suppression was used.

NMR Structure Calculations

Structure calculations were carried out using the simulated annealing protocol described by Kuszewski et al.[12] with the refine_full.inp script provided with XPLOR-NIH version 2.25.[13]

Ten structures were computed starting from classical A-form DNA structure and the five best structures, having the lowest energies and fewest number of experimental restraint violations, were deposited with the RCSB Protein Data Bank (PDB ID code 2L7D). 212 distance restraints were generated from the 100 ms NOESY cross-peaks using only the cross-peaks that appeared in the 50 ms T-ROESY spectrum to reduce the effect of spin diffusion. These peak volumes were converted to distance restraints using the NMRViewJ structure method with r^{-4} distance dependence to further account for spin diffusion. The integrals of cross-peaks involving the thymine methyl groups were reduced by a factor of 2.0. Integrals involving the rG4 H8 proton were increased by a factor of 2.0, due to deuterium exchange of this proton.[14] This factor was determined by integrating the 1D ^{1}H spectrum of the oligomer. The initial model used to calibrate these NOE-derived distances was obtained from a calculation using the NOE distance restraints provided with the sample XPLOR-NIH Dickerson dodecamer calculation.[15] The distance and dihedral restraints involving the rG4 nucleotide were not used, and no RDC or CSA restraints were used to generate the initial model. 158 ^{1}H-^{1}H RDCs, D_{HH}, were obtained from 2D COSY spectra acquired in isotropic and 20 mg/ml Pf1 media using the ACME fitting method.[16] The sign of the RDC values was determined from preliminary structure calculations and by comparison with RDC values obtained by Wu et al.[17] for the Dickerson dodecamer, since all the NMR data point to a structure that is very similar to the Dickerson dodecamer (see below). Twenty-two H3'-^{31}P scalar couplings, $^{3}J_{3'P}$, and RDCs, $D_{3'P}$, were measured using 2D ^{1}H-^{31}P CT-NOESY difference experiments obtained in isotropic and 20 mg/ml Pf1 media as described by Wu et al.[8] The $^{3}J_{3'P}$ scalar couplings were used to directly restrain the H3'-^{31}P dihedral angle, ϕ, via the standard Karplus relation.[17,18,19] The 22 deoxyribose sugar torsion angles, δ, were constrained to the S-type range ($140 \pm 35°$), based on the $^{3}J_{HH}$ couplings obtained from ACME simulations of the cross-peaks in the 2D COSY spectrum.[12,20,21] The two ribose sugar torsion angles, δ, were constrained to the N-type range ($80 \pm 35°$), based on the absence of the H1'-H2' cross-peaks in the 2D COSY and TOCSY spectra.[21,22] Since the imino proton region of the 200 ms NOESY spectrum in H_2O was virtually unchanged from that reported for the

Dickerson dodecamer,[23] the same Watson-Crick base pairing distance restraints used by Kuszewski et al.[12] in their calculation of the Dickerson dodecamer were used in the current refinement of the rG4-substituted dodecamer; six restraints per base pair were used. Initial calculations produced structures with close contacts (~1.6 Å) between the rG4 O2' and A5 H5' atoms in both strands of the symmetric duplex. The close contacts were removed by restraining the atom separation to be at least equal to the sum of their Van der Waals radii.

In order to compare the structure to the unsubstituted Dickerson dodecamer, a similar calculation was performed on the Dickerson dodecamer using only the NOE distance, $^{3}J_{3'P}$, D_{HH}, and $D_{3'P}$ restraints used by Schweiters and Clore,[19] provided with the XPLOR-NIH release, and originally obtained from Tjandra et al.[15] and Wu et al.[17] In addition, to make the two calculations as similar as possible, all the deoxyribose sugar torsion angles, δ, were constrained to the S-type range (140 ± 35º), as in the calculation of the rG4-substituted structure. In the remainder of this manuscript, the NMR calculated rG4-substituted structure will be referred to as rG4-DNA, and the newly calculated unsubstituted Drew-Dickerson dodecamer structure as dd-DNA.

Molecular Dynamics Simulations

Each dodecamer system was dissolved in a bath of water so that the closest nucleotide atom was at least 15Å away from the box boundary. Sodium counterions were introduced to neutralize the simulation system. The PMEMD module of the Amber.11 suite of programs was used in the trajectory calculations with the Amber99 force field.[28] The time step was set to 1 fs and the standard particle mesh Ewald (PME) procedure was used to accommodate the long range interactions.[29] Neighbor lists were updated at every MD step and after 2 ns of equilibration at constant temperature and volume, MD trajectories were calculated for at least 13 ns at constant temperature and constant pressure. Configurations extracted at 100 ps intervals of the last 10 ns were used for analysis.

II: Base Stacking in rG4-DNA

The six unique base-pair stacking diagrams for the ribose-substituted rG4-DNA structure and the dd-DNA structure are shown in Figure S1. These diagrams were generated using 3DNA v2.0.[24,25] The base stacking for both structures is similar, except for the base-pair steps containing the ribonucleotide substitution. In base-pair step 3, the rG4 base moves away from the C3 base and closer to the G10 base (labeled G22 in Figure S1) in the complementary strand, while the G10 base moves away from the C11 base (labeled C21 in Figure S1) in the complementary strand. The primary perturbation in stacking interaction occurs in base-pair step 4, where the rG4 and A5 bases show a greater degree of overlap in the rG4-DNA structure compared to the dd-DNA structure. The stacking interaction between C9 and T8 (labeled C21 and T20, respectively, in Figure S1) is reduced in the complementary strand. These changes may account for the slight increase in the chemical shifts of the imino proton of rG4 (Table 1b) and the H1' and H8 protons of A5 (Table 1a). All other base-pair step stacking interactions are virtually unperturbed by the ribose substitutions.

References

1. Hwang, T.-L. and Shaka, A. J. (1992) "Cross Relaxation without TOCSY: Transverse Rotating-Frame Overhauser Effect Spectroscopy. *J. Am. Chem. Soc. 114*, 3157-3159.

2. Hwang, T.-L., Kadkhodaei, M., Mohebbi, A., and Shaka, A. J. (1992) Coherent and Incoherent Magnetization Transfer in the Rotating-Frame. *Mag. Res. Chem. 30*, S24-S34.

3. Rucker, S. P. and Shaka, A. J. (1989) Broad-Band homonuclear cross polarization in 2D NMR using DIPSI-2. *Mol. Phys. 68*, 509-517.

4. Piotto, M., Saudek, V., Sklenar, V., (1992) Gradient-tailored excitation for single-quantum NMR-spectroscopy of aqueous solutions. *J. Biomol. NMR 2*, 661-665.

5. Delaglio, F., Wu Z., and Bax A. (2001) Measurement of homonuclear proton couplings from regular 2D COSY spectra. *J. Magn. Reson. 149*, 276-281.

6. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., and Bax, A. (1986) Assignment of the ^{31}P and ^{1}H resonances in oligonucleotides by two-dimensional NMR spectroscopy, *FEBS Lett.,* 208, 94-98.

7. Wu, Z., Tjandra, N., and Bax, A. (2001) ^{31}P chemical shift anisotropy as an aid in determining nucleic acid structure in liquid crystals, *J. Am. Chem. Soc. 123*, 3617-3618.

8. Wu, Z., Tjandra, N., and Bax, A. (2001) Measurement of ^{1}H3'-^{31}P dipolar couplings in a DNA oligonucleotide by constant-time NOESY difference spectroscopy. *J. Biomol. NMR 19*, 367-370.

9. Smallcombe, S. H., Patt, S. L., and Keifer, P. A. (1995) WET solvent suppression and its applications to LC NMR and high-resolution NMR, *J. Magn. Reson. A*, 117, 295-303.

10. Delaglio F, Grzesiek, G., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol. NMR 6*, 277-293.

11. Johnson BA, Blevins, R.A. (1994) NMR View - A computer-program for the visualization and analysis of NMR data. *J Biolmol. NMR 4*, 603-614.

12. Kuszewski, J., Schweiters, C., and Clore, G. M. (2001) Improving the accuracy of NMR structures of DNA by means of a database potential of mean force describing base-base positional interactions. *J. Am. Chem. Soc. 123*, 3903-3918.

13. Schwieters, C. D., Kuszewski, J. J., Tjandra, M. and Clore, G. M. (2003) the Xplor-NIH NMR molecular structure determination package. *J. Magn. Res.* 160, 66-74.

14. Walters, K. J., and Russu, I. M. Sequence Dependence of Purine C8H Exchange Kinetics in the Dodecamer 5'-d(CCCCAATTCCCC)-3'. *Biopolymers* (1993) 33, 943-951.

15. Tjandra, N., Tate, S., Ono, A., Kainosho, M., and Bax, A. (2000) The NMR structure of a DNA dodecamer in a n aqueous dilute liquid crystalline phase. *J. Am. Chem. Soc.,* 122, 6190-6200.

16. Delaglio F, Wu, Z., Bax, A. (2001) Measurement of Homonuclear Proton Couplings from Regular 2D COSY Spectra. *J. Magn.Reson.* 149, 276-281.

17. Wu, Z., Delaglio, F., Tjandra, N., Zhurkin, V.B., and Bax, A. (2003) Overall structure and sugar dynamics of a DNA dodecamer from homo- and heteronuclear dipolar coupling and ^{31}P chemical shift anisotropy. *J. Biomol.* NMR 26, 297-315.

18. Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C., and Altona, C. J. (1984) Nucleic-acid constituents .36. C-13 NMR in conformational analysis of nucleic-acid fragments .2. A reparametrization of the karplus equation for vicinal NMR coupling-constants in CCOP and HCOP fragments. *J. Biomol. Struct. Dyn.* 1, 1387-1405.

19. Schwieters, C. D. and Clore, G. M. (2007) A physical picture of atomic motions within the dickerson dna dodecamer in solution derived from joint ensemble refinement against NMR and large-angle x-ray scattering data. *Biochemistry* 46, 1152-1166.

20. Clore, G. M., Oschkinat, H., McLaughlin, L. W., Benseler, F., Happ, C. S., Happ, E., and Gronenborn, A. M. (1988) Refinement of the solution structure of the DNA dodecamer 5'd(CGCGPATTCGCG)$_2$ containing a stable purine-thymine base pair: combined use of nuclear magnetic resonance and restrained molecular dynamics. *Biochemistry* 27, 4185-4197.

21. Saenger, W. (1984) Principles of Nucleic Acid Structures, Springer Verlag, New York.

22. Butcher, S. E., Dieckmann, and Feigon, J. (1997) Solution structure of a GAAA tetraloop receptor. *EMBO J. 16*, 7490-7499.

23. Rajagopal P., Gilbert D. E., van der Marel, G. A., Jacques H. van Boom J. H., and Feigon J., (1988) Observation of Exchangeable Proton Resonances of DNA in Two-Dimensional NOE Spectra Using a Presaturation Pulse; Application to d(CGCGAATTCGCG) and d(CGCGAm6ATTCGCG). *J. Magn. Reson.* 78, 526-537.

24. Lu, X.-J. and Olson, W. K. (2003) 3DNA a software package for the analysis, rebuilding and visualization three-dimensional nucleic acid structures. *Nucleic Acids Res. 31*, 5108-5121.

25. Lu, X.-J. and Olson, W. K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures *Nature Protocols 3*, 1213-1227.

S1

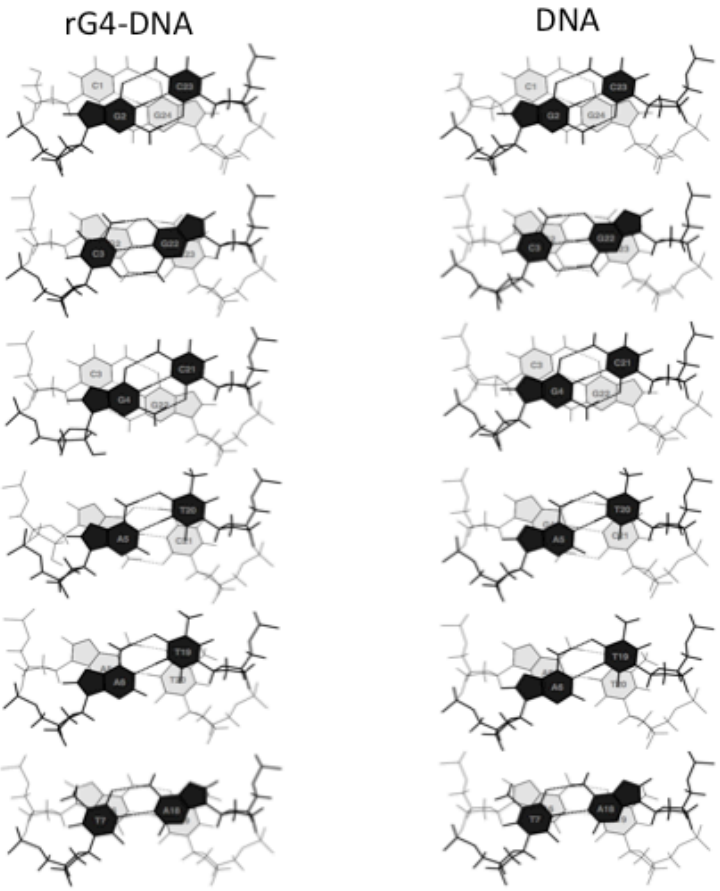

Figure S1. The six unique base-pair stacking diagrams for the best rG4-DNA and dd-DNA structures. The diagrams were generated using 3DNA v2.0.[24,25]