^{1}H NMR Structural Evidence for the Sequence-Specific Design of an Intercalator: Δ-α-[Rh[(*R*,*R*)-Me$_2$trien]phi]$^{3+}$ Bound to d(GAGTGCACTC)$_2$

Brian P. Hudson, Cynthia M. Dupureur, and Jacqueline K. Barton*

Division of Chemistry & Chemical Engineering and the Beckman Institute, California Institute of Technology Pasadena, California 91125

Received May 3, 1995

Through the synthesis, characterization, and application of octahedral metallointercalators, we have clarified many elements essential for molecular recognition of DNA. Among these are aspects of symmetry,[1] shape,[2] and the assembly of noncovalent contacts in the DNA major groove.[3] Our understanding of the elements of DNA recognition has allowed us to prepare[4] and examine[5] Δ-α-[Rh[(*R*,*R*)-Me$_2$trien]phenanthrenequinone diimine]$^{3+}$, a small metallointercalator rationally designed for recognition of 5′-TGCA-3′. This recognition is achieved via a combination of intercalation, hydrogen bonding, and stereospecific van der Waals contacts. Δ-α-[Rh[(*R*,*R*)-Me$_2$trien]phi]$^{3+}$ (RhM$_2$Tphi, phi = phenanthrenequinone diimine) (Figure 1A), like its parent molecule Δ-[Rh(en)$_2$phi]$^{3+}$,[6] targets 5′-GC-3′ steps as a result of hydrogen bonding between the axial amines of the metal complex and the O6 of guanine. In addition, the rhodium complex possesses methyl groups axially disposed to contacts with thymine methyl groups two bases in the 5′ direction from the intercalation site, resulting in specific binding at 5′-TGCA-3′. Diminishing photocleavage intensity is observed at sites 5′-TGCA-3′ > 5′-TGC-3′ > 5′-GC-3′,[5] a hierarchy of sites consistent with the building up of an ensemble of noncovalent interactions to achieve appreciable and predictable site-selectivity.

By combining photocleavage assays with NMR spectroscopy, we continue to develop the structural details of rhodium(III) metallointercalator–DNA interactions.[7] Here we report a high-resolution ^{1}H NMR structural model of RhM$_2$Tphi intercalated into d(GAGTGCACTC)$_2$. These results are important to consider in the general context of understanding intercalator–DNA interactions. This RhM$_2$Tphi–decamer adduct contains the smallest intercalator (~500 amu) bound site-specifically to the largest oligonucleotide (one site in a decamer) yet characterized by NMR spectroscopy.[8]

In one-dimensional ^{1}H NMR studies of RhM$_2$Tphi titrated into the decamer duplex d(GAGTGCACTC)$_2$, one set of shifted resonances grows in as resonances for the free duplex diminish (Figure 1B). This behavior indicates a single bound state and is consistent with binding at the central 5′-GC-3′ base step; the C_2 symmetry of the self-complementary duplex is maintained in the bound form. This site is the 5′-TGCA-3′ site predicted by photocleavage experiments.[5] Protons on the phi ligand

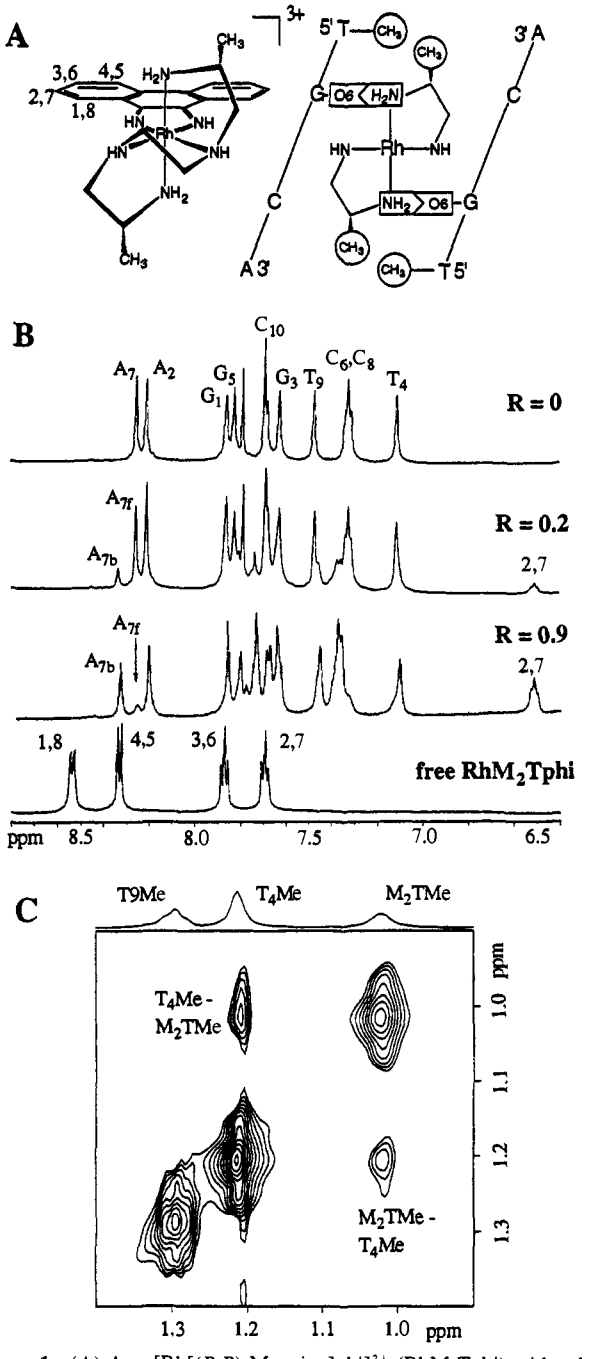

Figure 1. (A) Δ-α-[Rh[(*R*,*R*)-Me$_2$trien]phi]$^{3+}$ (RhM$_2$Tphi) with schematic representation of same bound to 5′-TGCA-3′; proposed hydrogen-bonding and van der Waals contacts are indicated. (B) 1D ^{1}H NMR spectra of RhM$_2$Tphi titrated into d(G$_1$A$_2$G$_3$T$_4$G$_5$C$_6$A$_7$C$_8$T$_9$C$_{10}$)$_2$ at 300 K: R = 0 (1.67 mM free duplex); R = 0.2; R = 0.9; 1 mM free RhM$_2$-Tphi (R = [RhM$_2$Tphi]/[duplex]). Shown is the aromatic proton region. A$_{7f}$ and A$_{7b}$ denote the free and bound forms of the A$_7$H$_8$ proton. (C) Partial NOESY of 1 mM RhM$_2$Tphi–D(GAGTGCACTC)$_2$ (R = 0.9) at 280 K. The NOE between the methyl protons of T$_4$ and RhM$_2$Tphi corresponds to the van der Waals contact designed for preferential recognition of 5′-TGCA-3′. Samples were prepared in buffered D$_2$O (10 mM phosphate, 20 mM NaCl, pH 7.0). Spectra were collected on a Brüker AMX500 spectrometer and referenced to TMSP (δ_c = 0.0 ppm).

* To whom correspondence should be addressed.

(1) Barton, J. K. *Science* **1986**, *233*, 727.
(2) Sitlani, A.; Barton, J. K. *Biochemistry* **1994**, *33*, 12100.
(3) Dupureur, C. M.; Barton, J. K. In *Comprehensive Supramolecular Chemistry*; Lehn, J.-M., Ed.; Pergamon Press: New York, 1995; Vol. 5 (in press).
(4) Krotz, A. H.; Barton, J. K. *Inorg. Chem.* **1994**, *33*, 1940.
(5) Krotz, A. H.; Hudson, B. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12577.
(6) Krotz, A. H.; Kuo, L. Y.; Shields, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 3877.
(7) (a) Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303. (b) David, S. S.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 2984. (c) Collins, J. G.; Shields, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 9840. (d) Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 10286.
(8) (a) Searle, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1993**, *25*, 403. (b) Brown, D. R.; Kurz, M.; Kearns, D. R.; Hsu, V. L. *Biochemistry* **1994**, *33*, 651. (c) Yang, D.; Wang, A. H.-J. *Biochemistry* **1994**, *33*, 6595. (d) Gao, X.; Stassinopoulos, A.; Rice, J. S.; Goldberg, I. H. *Biochemistry* **1995**, *34*, 40. (e) Ikemoto, N.; Kumar, R. A.; Dedon, P. C.; Danishefsky, S. J.; Patel, D. J. *J. Am. Chem. Soc.* **1994**, *116*, 9387. (f) Wu, W.; Vanderwall, D. E.; Stubbe, J.; Kozarich, J. W.; Turner, C. J. *J. Am. Chem. Soc.* **1991**, *116*, 10843. (g) Manderville, R. A.; Ellena, J. F.; Hecht, S. M. *J. Am. Chem. Soc.* **1994**, *116*, 10851.

0002-7863/95/1517-9379$09.00/0

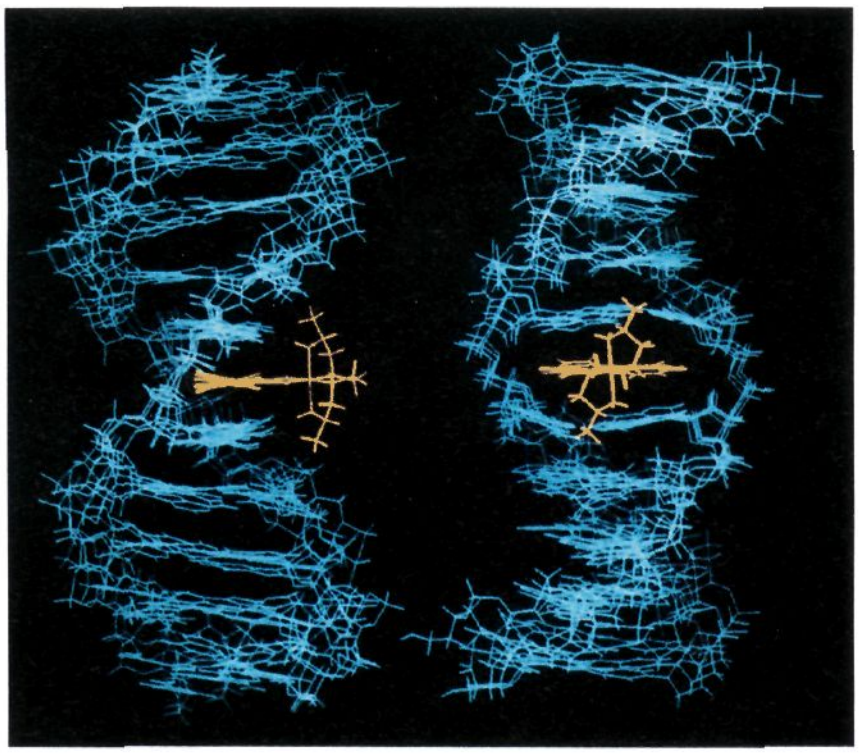

Figure 2. Model of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ bound to d(GAGTGCACTC)$_2$. Side and front views of four overlayed low-energy structures. The starting structures included a centrally intercalated A-form helix and three B-form structures with the metal complex intercalated in the center of the site, skewed to the side, and pulled out of the site.

experience strong upfield shifts of 0.5−1.2 ppm, characteristic of intercalation.[9] Variable temperature ^{1}H NMR studies indicate an exchange rate between free and bound forms of 90 s^{-1} at the 325 K coalescence point.[10] Chemical shift changes for the duplex also reflect the site-specificity of the intercalation; large shifts occur only at the central 5′-TGCA-3′ recognition site.

Two-dimensional NOESY data provide the basis for a high-resolution structural model which is fully consistent with the designed binding of RhM_2Tphi to 5′-TGCA-3′. In the diagnostic base−H2′2″ region of NOESY spectra of 0.9:1 RhM_2Tphi-d(GAGTGCACTC)$_2$, a distinct break is observed in the sequential connectivities at the central 5′-G_5C_6-3′, supporting site-specific intercalation of the metal complex even at millimolar concentrations. NOESY spectra yield 13 distinct intermolecular and 156 distinct intramolecular NOEs.[11] Included among these are 11 through-space connectivities between the intercalative phi ligand and either G_5 or G_6 nucleotide residues. Strong evidence for the success of our design of this complex comes from the detection of intermolecular NOEs between the methyl group of the Me_2trien ligand and the methyl group of T_4 (Figure 1C), precisely as designed.

Given the C_2 symmetry of the DNA−intercalator adduct, a total of 364 unambiguous distance restraints were applied with molecular dynamics to develop structural models for RhM_2Tphi intercalation into the DNA decamer.[12] Figure 2 illustrates these resulting structures. The duplex is essentially B-DNA, except at the intercalation site, and no bend in the helix is apparent. The models demonstrate deep intercalation of RhM_2Tphi into the DNA duplex from the major groove. The metal complex is positioned to maximize stacking of the phi ligand between guanines to the 5′ side of the intercalation site. This intercalative geometry poises the axial amines of Me_2trien for hydrogen bonding to the guanine O6 atoms of G_5. Also evident are intermolecular contacts between the T_4 thymine methyl groups and methyl groups of the slightly bowed RhM_2Tphi. The results obtained from this system demonstrate that elaboration of ligand functionality on octahedral metallointercalators provides a sensible strategy for the rational construction of site-specific DNA-binding molecules.

Acknowledgment. We are grateful to the NIH (GM33309) for their financial support and to the NSF for postdoctoral (C.M.D.) and predoctoral (B.P.H.) fellowships. We also thank Silvia Cavagnero of the Caltech NMR facility for excellent technical assistance and J. Grant Collins for preliminary results.

Supporting Information Available: Figure depicting 1D variable temperature spectra of RhM_2Tphi−d(GAGTGCACTC)$_2$; graph of chemical shift perturbations of the decamer upon binding of RhM_2-Tphi; contour plot of the base−H2′2″ region of a 120 ms NOESY spectrum of RhM_2Tphi−d(GAGTGCACTC)$_2$; list of inter- and intramolecular NOEs; views of starting structures; stereoview of one final structure (9 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA951437N

(9) (a) Feigon, J.; Denny, W. A.; Leupin, W.; Kearns, D. R. *J. Med. Chem.* **1984**, *27*, 450. (b) Patel, D. J.; Canuel, L. L. *Biopolymers* **1977**, *16*, 857. (c) Patel, D. J.; Shen, C. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2553. (d) Wilson, W. D.; Krishnamoorthy, C. R.; Wang, Y. H.; Smith, J. C. *Biopolymers* **1985**, *24*, 1941. (e) Chandrasekaran, S.; Jones, R. L.; Wilson, W. D. *Biopolymers* **1985**, *24*, 1963. (f) Delbarre, A.; Gourevitch, M. I.; Gaugain, B.; Le Pecq, J. B.; Roques, B. P. *Nucleic Acids Res.* **1983**, *11*, 4467.

(10) (a) Fede, A.; Labhardt, A.; Bannwarth, W.; Leupin, W. *Biochemistry* **1991**, *30*, 11377. (b) The exchange rate was estimated by observation of the A_7H8 proton in a sample containing a 1:1 mixture of free and bound duplex. The melting points of the free and bound duplex are approximately 329 and 338 K, respectively.

(11) ^{1}H NMR data (D_2O, 10 mM phosphate, 20 mM NaCl, pH 7.0, referenced to TMSP) were collected on a Varian Unity Plus 600 MHz spectrometer (298 K, hypercomplex mode, 2048 × 256 × 64 scans) and processed using Felix (v. 2.3, Biosym Technologies). Assignments were made only in cases involving minimal overlap and unambiguous identity.

(12) Interproton distances involving nonexchangeable protons were calculated using linear least-squares fitting of NOE buildups (40, 80, and 120 ms mixing times), averaged across the diagonal, and classified as strong (<2.5 Å), medium (2.5−3.5 Å), or weak (>3.5 Å) to generate molecular distance restraints with overlapping limits (strong, 1.8−2.8 Å; medium, 2.3−3.8 Å; weak, 3.3−5.0 Å). Every restraint involving a complete methyl or methylene group as a member was moved to the next-weaker category. Distance restraints were applied as square-well potentials with upper- and lower-limit force constants of 30 kcal/mol·Å^2. Interstrand hydrogen bonds were added as 26 restraints with force constants of 120 kcal/mol·Å^2 to maintain Watson−Crick base pairing. All molecular simulations were performed using a modified AMBER2 force field[13] with a distance-dependent dielectric ($\epsilon = 4r_{ij}$), distance restraints, and a 12 Å nonbonded cutoff distance using the the NMRchitect interface to the Discover module of InsightII (v. 2.3, Biosym Technologies). Initial starting structures were generated using the mirror image of the crystal structure[4] of Λ-α-[Rh[(*S*,*S*)-Me_2trien]phi]$^{3+}$ and DNA models created using the biopolymer module of InsightII. Starting structures were subjected to 500 steps of conjugate gradient minimization followed by molecular dynamics with a time step of 0.5 fs. Dynamics calculations were started at 5 K and heated to 1000 K over 6 ps. The system was maintained at 1000 K for 16 ps, gradually cooled to 300 K over 7 ps, and maintained at 300 K for 30 ps. Coordinates were stored every 0.5 ps for the last 5 ps and then averaged. The resulting structure was then subjected to 100 steps of steepest descents minimization followed by conjugate gradient minimization until the RMS derivative for the system was $<10^{-4}$ kcal/mol·Å^2. All calculations were performed on a Silicon Graphics Iris Indigo workstation. Average pairwise RMS deviation among the final structures is 1.48 Å and, among the bound recognition sites, is 1.09Å. All intermolecular distances are within 0.1 Å of defined constraints.

(13) (a) Veal, J. W.; Wilson, W. D. *J. Biomol. Struct. Dyn.* **1991**, *8*, 1119. (b) AMBER2 has been modified to maintain phi ligand planarity and octahedral coordination geometry about rhodium(III) during simulation.