Solution Structure of a Metallointercalator Bound Site Specifically to DNA

Brian P. Hudson and Jacqueline K. Barton*

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

Received December 5, 1997

Abstract: The solution structure of the complex of the metallointercalator Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ ((*R*,*R*)-Me_2trien = 2*R*,9*R*-diamino-4,7-diazadecane; phi = 9,10-phenanthrenequinone diimine) bound to d(GAGTG-CACTC)$_2$ at 4 °C has been determined using ^{1}H NMR spectroscopy coupled with restrained molecular dynamics calculations. Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ binds specifically to its designed site, 5′-TGCA−3′, with slow exchange kinetics. Using interproton distance restraints (including 70 intermolecular restraints) derived from NOESY spectra in both D_2O and H_2O and torsional restraints derived from ECOSY data, 20 disparate starting structures converged to a set of final structures with an average RMSD (between structures) of 1.0 Å. As shown previously, binding occurs by deep intercalation of the phi ligand at the central 5′-GC-3′ step of the decamer, positioning the ancillary Me_2trien ligand in the major groove. C_2 symmetry is maintained with no kinking or bending of the DNA double helix. The binding site is deformed from canonical B-DNA and appears to be pulled slightly out of the major groove into a conformation which maximizes favorable interactions with the complex. The axial amines of the metallointercalator are positioned to hydrogen bond to the N7 and O6 of the G_5(G_{15}) bases. Van der Waals contacts between the axial methyl groups of Δ-α-[Rh[(*R*,*R*)-Me_2trien]-phi]$^{3+}$ and the methyl groups of T_4(T_{14}) are apparent, with conformational changes in the binding site creating a binding pocket across the major groove face of the T_4 nucleotide. The intercalated 5′-GC-3′ step experiences a positive slide and significant overwinding (helical twist = 53°) which allows the guanine bases to stack exclusively on the phi ligand. The site selectivity and small size of this DNA-binding molecule also permit an examination of how intercalation is propagated through the DNA helix. While the intercalated step is overwound, adjacent steps are unwound, with maximal unwinding (helical twist = 22°) occurring at the 5′-G_3T_4-3′ and 5′-A_7C_8-3′ steps which define both the edge of the 5′-TGCA-3′ binding site and the edge of the conformational changes associated with binding. The flanking trimers are relatively unperturbed from canonical B-DNA, with the duplex experiencing a net unwinding of 39 ± 4°.

Introduction

As the applications of metallointercalators increase, the need for detailed structural information on the nature of the binding interaction becomes paramount. Metallointercalators have been exploited as luminescent probes for DNA,[2] tools to examine the local structural polymorphism of nucleic acids,[3] photoreactive[4] and electrochemically active[5] probes of DNA structure and protein binding, reagents targeted to recognize mismatches[6] or to repair thymine dimers,[7] and probes of long range charge transport through the DNA helix.[8,9] Metallointercalators containing appended peptides have been developed which react with DNA hydrolytically[10] and may form the foundation for new tools in molecular biology. All of these applications depend sensitively upon the intimate structural interactions of the metallointercalator with its nucleic acid binding site.

These many applications require that we consider not only the structure of the site recognized by the metallointercalator but also the structure of the site upon binding. Our ability to predict functional characteristics of the metallointercalator requires an understanding of whether and how the DNA

* To whom correspondence should be addressed.

(1) Hudson, B. P.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1995**, *117*, 9379.

(2) (a) Murphy, C. J.; Barton, J. K. *Methods Enzymol.* 1993, *226*, 576. (b) Friedman, A. E.; Chambron, J. C.; Sauvage, J. P.; Turro, N. J.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 4960. (c) Stoeffler, H. D.; Thornton, N. B.; Temkin, S. L.; Schanze, K. S. *J. Am. Chem. Soc.* **1995**, *117*, 7119. (d) Yam, V. W.-W.; Lo, K. K.-W.; Cheung, K.-K.; R. Y.-C. Kong, *J. Chem. Soc., Chem. Commun.* **1995**, 1191.

(3) (a) Chow, C. S.; Barton, J. K. *Methods Enzymol.* **1992**, *212*, 219. (b) Chow, C. S.; Hartmann, K. M.; Rawlings, S. L.; Huber, P. W.; Barton, J. K. *Biochemistry* **1992**, *31*, 3534. (c) Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972. (d) Lim, A. C.; Barton, J. K. *Biochemistry* **1993**, *32*, 11029.

(4) Uchida, K.; Pyle, A. M.; Morii, T.; Barton, J. K. *Nucleic Acids Res.* **1989**, *17*, 10259.

(5) (a) Gupta, N.; Grover, N.; Neyhart, G. A.; Singh, P.; Thorp, H. H. *Inorg. Chem.* **1993**, *32*, 310. (b) Welch, T. W.; Corbett, A. H.; Thorp, H. H. *J. Phys. Chem.* **1995**, *99*, 7.

(6) Jackson, B. A.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 12986.

(7) Dandliker, P. J.; Holmlin, R. E.; Barton, J. K. *Science* **1997**, *275*, 1465.

(8) (a) Murphy, C. J.; Arkin, M. R.; Ghatlia, N. D.; Bossmann, S. H.; Turro, N. J.; Barton, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5315. (b) Murphy, C. J.; Arkin, M. R.; Jenkins, Y.; Ghatlia, N. D.; Bossmann, S. H.; Turro, N. J.; Barton, J. K. *Science* **1993**, *262*, 1025. (c) Arkin, M. R.; Stemp, E. D. A.; Turro, C.; Turro, N. J.; Barton, J. K. *J. Am. Chem. Soc.* **1996**, *118*, 2267. (d) Arkin, M. R.; Stemp, E. D. A.; Holmlin, R. E.; Barton, J. K.; Hörmann, A.; Olson, E. J. C.; Barbara, P. F. *Science* **1996**, *273*, 475. (e) Hall, D. B.; Holmlin, R. E.; Barton, J. K. *Nature* **1996**, *382*, 731. (f) Hall, D. B.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 5045. (g) Arkin, M. R.; Stemp, E. D. A.; Pulver, S. C.; Barton, J. K. *Chem. Biol.* **1997**, *4*, 389.

(9) Holmlin, R. E.; Dandliker, P. J.: Barton, J. K. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2714.

(10) Fitzsimons, M. P.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 3379.

S0002-7863(97)04134-6 CCC: $15.00
Published on Web 07/07/1998

conformation might be altered. Any application of these complexes as novel chemotherapeutics depends on understanding the structural changes in the nucleic acid that are associated with binding. Intercalation and proper stacking are required for the fast, efficient transfer of electrons through the DNA π stack.[7−9] The detailed understanding of how these metallointercalators are stacked within the helix is critical to developing parameters which describe the efficiency of DNA-mediated electron transfer between intercalated donors and acceptors.

In the past several years, a number of factors governing the recognition and binding of octahedral metallointercalators have been elucidated. NMR studies of complexes bound to DNA without high specificity have indicated the general intercalation of the expansive aromatic heterocyclic ligand.[11−13] These NMR studies have also indicated the preferential binding of right-handed isomers. Data consistent with intercalation from the major groove side have been obtained for most complexes.[12,13] For rhodium complexes containing the 9,10-phenanthrenequinone diimine (phi) ligand, photocleavage experiments have been valuable in assessing DNA site affinity and specificity.[14−17] DNA binding affinity is provided in large measure by stacking of the intercalated phi; affinities have ranged from 10^6 to 10^8 M^{-1}. Site selectivity is determined primarily by the ancillary ligands of the metal complex, which mediate binding on the basis of their interaction with the major groove of the DNA double helix. Large, aromatic ancillary ligands lead to binding at sites where the local conformation of the DNA, as determined by its sequence, can accommodate their steric bulk. An element of this shape-selective recognition is an inherent enantioselectivity: right-handed (Δ) enantiomers fit into the groove of the right-handed DNA double helix, while left-handed (Λ) enantiomers experience clashes with the DNA deoxyribose-phosphate backbone.[14]

The recognition of selected DNA sequences by enantiomers of $[Rh(en)_2phi]^{3+}$ (en = ethylenediamine) illustrates another element of metallointercalator design: the use of the rigid, substitutionally inert, octahedral framework of the metal center as a scaffold for the placement of functionality in the major groove of DNA. Δ-$[Rh(en)_2phi]^{3+}$ binds selectively to 5′-GC-3′ base steps, presumably due to hydrogen bonding by the axial amines of the metallointercalator. Λ-$[Rh(en)_2phi]^{3+}$ binds the same site, but also binds to 5′-TR-3′ (R = purine) sites due to a positive van der Waals interaction between the thymine methyl group and the aliphatic chain of one of the ethylenediamine ligands.[18] The binding characteristics of the enantiomers of $[Rh(en)_2phi]^{3+}$ demonstrate recognition by hydrogen bonding and van der Waals interactions in the absence of shape selection.

The small metallointercalator Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$, shown in Figure 1, was designed to bind to the 4 base pair site 5′-TGCA-3′.[19] While its high affinity for DNA would be provided by its positive charge and the intercalation of the phi ligand into the 5′-GC-3′ step, selectivity would be dictated by an ensemble of hydrogen bonds between the axial amines of the metallointercalator and the O6 of the guanine bases and by favorable van der Waals contacts between the methyl groups of the metallointercalator and the methyl groups of the thymine bases to the 5′ side of the intercalated base step, in analogy to the enantiomers of $Rh(en)_2phi$. Photocleavage studies showed 5′-TGCA-3′ to be a strong binding site for Δ-α-[Rh[(*R*,*R*)-Me_2-trien]phi]$^{3+}$.[19]

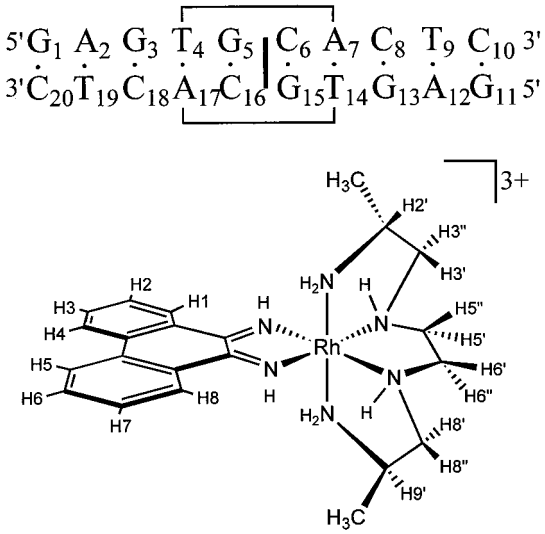

Figure 1. The oligonucleotide duplex, $d(GAGTGCACTC)_2$, enumerated as used in the text, showing the extent of the 5′-TGCA-3′ recognition site in brackets and the location of the intercalation site with a vertical line. Also shown is a diagram of Δ-α-[Rh[(*R*,*R*)-Me_2-trien]phi]$^{3+}$, labeled to show the proton nomenclature used in the text.

Previous NMR investigations[1] of the bound complex of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ ((*R*,*R*)-Me_2trien = 2*R*,9*R*-diamino-4,7-diazadecane) and the self-complementary decamer $d(GAGTGCACTC)_2$ at 25 °C were made possible by the relatively slow exchange kinetics of this metallointercalator. They show that the rhodium complex binds exclusively to its designed site in a mode that maintains the C_2 symmetry of all components of the system. All intermolecular NOEs were consistent with intercalation of the phi ligand between the G_5 and C_6 bases at the center of the duplex and simultaneous positioning of the ancillary Me_2trien ligand in the major groove with the methyl groups of the ligand and T_4 poised for a favorable van der Waals interaction. The presence of a significant number of NOEs between nonexchangeable protons, both intermolecular and intramolecular, allowed the derivation of a model using restrained molecular dynamics and energy minimization. This model showed that the phi ligand was deeply intercalated into the base stack and that no kinking or bending of the DNA duplex was evident. While functional groups in the major groove at the binding site were in positions *consistent with* the proposed hydrogen bonding and van der Waals interactions, the model was not restrained well enough by the experimental data for conclusions regarding these additional details of binding to be made.

We have since collected larger and more detailed data sets for this system at a lower temperature, 4 °C. A more rigorous treatment of the experimental restraints plus the addition of data for exchangeable protons and deoxyribose sugar conformations has allowed the determination of a full solution structure of the 1:1 Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$/$d(GAGTGCACTC)_2$ complex. This structure is much better resolved than the previous model and remains generally consistent with it. The increase

(11) Greguric, I.; Aldrich-Wright, J. R.; Collins, J. G. *J. Am. Chem. Soc.* **1997**, *119*, 3621.
(12) (a) Collins, J. G.; Shields, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 9840. (b) David, S. S.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 2984. (c) Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 10286. (d) Dupureur, C. M.; Barton, J. K. *Inorg. Chem.* **1997**, *36*, 33.
(13) Shields, T. P.; Barton, J. K. *Biochemistry* **1995**, *34*, 15049.
(14) Barton, J. K. *Science* **1986**, *233*, 727.
(15) Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303.
(16) (a) Sitlani, A.; Barton, J. K. *Biochemistry* **1994**, *33*, 12100. (b) Krotz, A. H.; Kuo, L. Y.; Shields, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 3877.
(17) Terbrueggen, R. H.; Barton, J. K. *Biochemistry* **1995**, *34*, 8227.
(18) Shields, T. P.; Barton, J. K. *Biochemistry* **1995**, *34*, 15037.
(19) Krotz, A. H.; Hudson, B. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12577.

in resolution allows a fuller characterization of the conformation of the DNA decamer and reveals interesting details previously unobserved. The specificity of this small metallointercalator for a single site in the center of a relatively long (10 bp) oligonucleotide allows the observation of how the effects of intercalation are propagated along the DNA helix. In addition, various factors which contribute to both binding affinity and specificity of this synthetic complex can be addressed.

Experimental Section

Materials. Diastereomers of [Rh[(*R*,*R*)-Me$_2$trien]phi]$^{3+}$ were synthesized and separated as described previously for [Rh[(*S*,*S*)-Me$_2$trien]-phi]$^{3+}$ ((*S*,*S*)-Me$_2$trien = 2*S*,9*S*-diamino-4,7-diazadecane),[20] using D-alanine as the starting material for ligand synthesis rather than L-alanine. The right-handed, C_2-symmetric (Δ-α-) diastereomer was exchanged to the chloride salt on a plug of Sephadex SP-C25 ion-exchange resin (Sigma) prepared with HCl. Solution concentration was determined by UV-visible spectroscopy (ϵ_{373} = 13 500 M^{-1} cm^{-1} at pH = 7.0). Phosphoramidites, reagents, and solid supports for DNA synthesis were purchased from Glen Research. The self-complementary tritylated decamer oligodeoxynucleotide 5′-GAGTGCACTC-3′ was synthesized on an Applied Biosystems 392 DNA/RNA synthesizer using phosphoramidite chemistry,[21] purified by HPLC, detritylated with 80% acetic acid, and purified again by HPLC. The purified decamer was desalted on a Millipore C_{18} Sep-Pak and exchanged to the sodium salt on Sephadex CM-C25 ion-exchange resin (Sigma) prepared with NaCl. Solution concentration was determined by UV-visible spectroscopy (ϵ_{260} = 6600 $M_{\text{nucleotides}}{}^{-1}$ cm^{-1}). Sodium 3-trimethylsilyl-[2,2,3,3-D$_4$]-propionate (TMSP) and 99.96% D$_2$O were purchased from Cambridge Isotope Labs (CIL). Deionized H$_2$O was purified with a Millipore Milli-Q Plus system. All other chemicals and biochemicals used were of the highest quality and purity available.

Instrumentation. UV-visible absorbance measurements were made on either a Varian Cary 2200 spectrophotometer or a Beckman DU 7400 diode array spectrophotometer. HPLC was performed on a Waters 600E multi-solvent delivery system and monitored with a model 484 tunable wavelength detector. Separate C_{18} semipreparative columns were used for DNA and metallointercalator samples. ^{1}H NMR spectra were collected either on a 600 MHz Varian Unity Plus spectrometer with variable temperature control or on a 500 MHz Brüker AMX500 spectrometer with variable temperature control and pulsed-field gradients in three dimensions. NMR data processing and molecular mechanics calculations were performed on a Silicon Graphics Iris Indigo workstation with additional data processing performed using a Gateway2000 P5-166 personal computer.

NMR Sample Preparation. All NMR samples were buffered to pH 7.0 with 10 mM sodium phosphate containing 20 mM NaCl. After preparation, the samples and appropriate stocks were lyophilized at least three times with D$_2$O before being reconstituted in 99.96% D$_2$O. For observation of exchangeable protons, samples were reconstituted in 90% H$_2$O, 10% D$_2$O. Samples of free d(GAGTGCACTC)$_2$ were at 1.0 mM (duplex) concentration. Other samples were prepared at a concentration of 1.11 mM duplex and titrated with a 10 mM solution of Δ-α-[Rh[(*R*,*R*)-Me$_2$trien]phi]$^{3+}$ to a [Rh]/[duplex] ratio of 1:1 (titration monitored by 1D ^{1}H NMR) and bound complex concentration of 1.0 mM. Samples of free metallointercalator were either 1.0 mM or 2.5 mM. All sample volumes were 700 μL and contained TMSP as an internal standard (0.00 ppm).

Spectroscopy. NOESY spectra (600 MHz, 10 ppm sweep width, hypercomplex mode, 2048 complex points, 640 t_1 blocks, 32 scans per t_1 block, 1.3 s relaxation delay, water suppression by presaturation during relaxation delay and mixing) of 1:1 Δ-α-[Rh[(*R*,*R*)-Me$_2$trien]-phi]$^{3+}$/d(GAGTGCACTC)$_2$ at 4 °C were collected concurrently at seven different mixing times: 0, 20, 40, 60, 80, 100, and 120 ms. NOESY spectra of d(GAGTGCACTC)$_2$ were collected under identical conditions (but only 512 t_1 blocks) with a mixing time of 120 ms. WATERGATE NOESY spectra (500 MHz, 20 ppm sweep width, TPPI, 2048 complex points, 512 t_1 blocks, 128 scans per t_1 block, 1.3 s relaxation delay) of Δ-α-[Rh[(*R*,*R*)-Me$_2$trien]phi]$^{3+}$/d(GAGTGCACTC)$_2$ and d(GAGTG-CACTC)$_2$ samples were collected at 4 °C with a 125 ms mixing time. ECOSY spectra (500 MHz, 8 ppm sweep width, TPPI, 8192 complex points, 256 t_1 blocks, 160 scans per t_1 block, 1.3 s relaxation delay, water suppression by presaturation during relaxation delay) were collected at 27 °C. Additional ECOSY spectra for d(GAGTGCACTC)$_2$ were also collected with 2048 complex points and 512 t_1 blocks under otherwise identical conditions. Additional ^{1}H NMR spectra—1D (WATERGATE and presaturation) and phase-sensitive COSY—were collected at 4 °C and 27 °C at 500 MHz for the purpose of assigning unbound metallointercalator proton resonances and verifying assignments made for other samples.

NMR Processing. All NMR spectra were processed using *Felix* version 2.3 (BIOSYM/Molecular Simulations) using a 90° skewed sinebell squared apodization in both dimensions. NOESY and ECOSY spectra were zero-filled to 2048 points or 512 points in the t_1 dimension, respectively. No baseline correction was used.

Cross-peaks in NOESY spectra were integrated (rectangular integration areas) on both sides of the diagonal. Mixing time buildup data were fit to a second-order polynomial using Microsoft Excel version 7.0 for Windows95. Fit quality was checked using eq 1

$$\alpha = \frac{\sum |y_i - a - bx_i - cx_i^2|}{\sum |y_i|} \qquad (1)$$

where a, b, and c are the fit parameters, x_i is mixing time, y_i is integration volume, and i ranges from 0 to 120 ms. Peaks with $\alpha > 0.01$ were eliminated from consideration. Interproton distances were calculated using the inverse sixth-power dependence of NOE intensity on interproton distance using eq 2

$$r = (b_{\text{ref}} r_{\text{ref}}^6 b^{-1})^{1/6}$$

In the equation, r is interproton distance, b is the initial slope of the buildup curve or, in the case of WATERGATE data, the integrated volume of the peak. Calculations were referenced using the CH5-H6 interproton distance, r_{ref} = 2.47 Å, and the associated experimental integrated NOE volume, b_{ref} (average for C_6, C_8, and C_{10}). Distances for identical assignments on both sides of the diagonal were averaged. "Trivial" NOEs (for example, geminal protons and CH5-H6) were eliminated from the restraint set. Lower distance restraints were set to 1.80 Å. Upper distance restraints were set to the calculated distance and expanded by 20% to compensate for inherent errors in distance estimation.[22] In addition, upper distance restraints were increased by 1.8 Å for every methylene group in which the chemical shifts of the constituent protons were coincident and 1.0 Å for every methyl group.[23] Assignments for methylene groups whose individual protons could not be stereospecifically assigned were allowed to float during energy minimization and molecular dynamics calculations.

Spin−spin coupling constants were measured directly from ECOSY spectra.[24] Using known values of $J_{\text{H-H}}$ for the C2′-endo and C3′-endo configurations,[25] dihedral angles could be estimated from experimental data.[26] The upper and lower torsional restraints were set to 20% greater and less than the calculated angles, respectively. After calculation and manipulation, the entire set of distance and torsional restraints was doubled for application to both halves of the C_2-symmetric system.

Molecular Mechanics. All calculations were performed using the *InsightII* version 2.3.0 software package (BIOSYM/Molecular Simula-

(20) Krotz, A. H.; Barton, J. K. *Inorg. Chem.* **1994**, *33*, 1940.
(21) Caruthers, M. H.; Barone, A. D.; Beaucage, S. L.; Dodds, D. R.; Fisher, E. F.; McBride, L. J.; Matteucci, M.; Stabinsky, Z.; Tang, J.-Y. *Methods Enzymol.* **1987**, *154*, 287.
(22) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1988**, *79*, 493.
(23) Zhang, X.; Patel, D. J. *Biochemistry* **1990**, *29*, 9451.
(24) (a) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1985**, *107*, 6394. (b) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Magn. Reson.* **1987**, *75*, 474.
(25) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(26) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11.

tions) with an AMBER force field[27] modified for the specialized modeling of nucleic acids with cationic ligands[28] and further modified to maintain the octahedral rhodium center and planar 9,10-phenanthrenequinone diimine (phi) ligand of Δ-α-[Rh[(*R,R*)-Me_2trien]phi]$^{3+}$. Solvent was simulated with a distance-dependent dielectric ($\epsilon = 4r_{ij}$). A Lennard−Jones potential was used to model nonbonded interactions with the nonbonded cutoff distance set to 10 Å.

Starting models were constructed using the *Biopolymer* module of *InsightII* (BIOSYM/Molecular Simulations). Canonical B-DNA and A-DNA were capped with 5′-OH groups and used without further modification. B-form and A-form models of d(GAGTGCACTC)$_2$ were also constructed with engineered 5′-G_5pC_6-3′ intercalation sites (doubled rise, 20° unwinding) and energy minimized briefly (100 steps steepest descents) to eliminate unrealistic bond lengths and angles. Additional starting structures were constructed with substantial variation of helical parameters (see Supporting Information Table S15). Δ-α-[Rh[(*R,R*)-Me_2trien]phi]$^{3+}$ was inserted into the models as the mirror image of the crystal structure[20] of Λ-α-[Rh[(*S,S*)-Me_2trien]phi]$^{3+}$. In many cases, the metal complex was inserted from the major groove between the base pairs of the central 5′-GC-3′ base step, a position consistent with qualitative examination of intermolecular NOEs.

Starting structures were subjected initially to 500 steps of conjugate gradient minimization before molecular dynamics simulation (time step = 0.5 fs), which consisted of 1 ps at 5 K followed by heating over 5 ps to 1000 K. The system was maintained at 1000 K for 15 ps, then cooled over 6 ps to 300 K, where it was maintained for 30 ps. Coordinates were collected every 0.5 ps for the last 5 ps of the simulation and averaged. The average structure was subjected to 100 steps of steepest descents and 20 000 steps of conjugate gradients energy minimization to yield the final structure. Several iterations of this procedure were required to identify improper assignments and erroneous restraints. Unscaled restraints were applied throughout the entire procedure using a square-well potential with an $E = (50\ \text{kcal/mol Å}^2)(r - r_{+/-})^2$ energy penalty applied outside the upper (r_+) and lower (r_-) restraint boundaries and no energy penalty (beyond those applied by the force field) within the restraint boundaries. Final structures were analyzed in *InsightII*, and helical parameters were calculated using the *rnew*8 program[29] obtained from the FTP archive of the Nucleic Acid Database (NDB).

Results

Assignments. Assignment of the 120 ms D_2O NOESY of 1:1 Δ-α-[Rh[(*R,R*)-Me_2trien]phi]$^{3+}$/d(GAGTGCACTC)$_2$ yielded a total of 725 cross-peak assignments for nonexchangeable protons. Assignment of the WATERGATE NOESY of the same system reconstituted in 90% H_2O, 10% D_2O yielded 97 additional cross-peak assignments involving exchangeable protons. Elimination of trivial, poorly fit, and overlapped cross-peaks and averaging of identical cross-peaks found on both sides of the diagonal returned a total of 336 (304 nonexchangeable, 32 exchangeable) cross-peaks available for distance restraint calculation. Application of these restraints to the two C_2-symmetrically related halves of the system resulted in a total of 672 interproton distance restraints. Due to overlap and poor resolution, ECOSY yielded only 12 spin−spin coupling constants ($6 \times J_{H1'-H2'}$, $6 \times J_{H1'-H2''}$), which resulted in a total of 24 torsional restraints for use in subsequent modeling.

Intermolecular Restraints. The final restraint set yielded 70 intermolecular restraints. Forty-six of those restraints (22 G_5/G_{15} to phi, 24 C_6/C_{16} to phi) place the phi ligand firmly between the G_5 and C_6 bases, consistent with the observed loss of G_5H1′−C_6H6, G_5H2′−C_6H6, and G_5H2″−C_6H6 NOEs (this interruption of the "NOE walk" occurs when DNA bases are separated by the insertion of an intercalator). The only assignment ambiguities that arose from the 2-fold symmetry of the DNA-metallointercalator complex involved NOEs between protons of the intercalated phi ligand and the G_5H1 protons at the center of the GC base pairs at the intercalation site. The remaining intermolecular NOEs (24) occur between protons on the Me_2trien ancillary ligand and protons typically found exclusively in the major groove of B-form DNA. Strongest among these, as noted previously,[1] is an NOE between the methyl group (1′) of Me_2trien and the methyl group of T_4, entirely consistent with the proposed methyl−methyl van der Waals interactions. Also strong are contacts between the 1′-methyl and the H6, H2′, and H2″ of T_4. These restraints, in concert with NOEs between protons typically in the minor groove and protons on the leading edge of the intercalating ligand (G_5H1′ to phi H5, H6, H7 for example), demonstrate an interaction between the metallointercalator and DNA with the ancillary ligand in the major groove and the intercalating ligand thrust deeply into the base stack.

Table 1. Measured Spin−Spin Coupling Constants (Hz) from ECOSY for d(GAGTGCACTC)$_2$ at 27 °C, Free and Bound to Δ-α-[Rh[(*R,R*)-Me_2trien]phi]$^{3+}$ [a]

	free		bound	
	$3J_{H1'-H2'}$	$3J_{H1'-H2''}$	$3J_{H1'-H2'}$	$3J_{H1'-H2''}$
G_1/G_{11}	9.5	5.5	10	5
A_2/A_{12}	9	6	*	*
G_3/G_{13}	8.5	6	9.5	5
T_4/T_{14}	*	6	*	*
G_5/G_{15}	9	7	8.5	6
C_6/C_{16}	8	6.5	−	−
A_7/A_{17}	8	7	*	*
C_8/C_{18}	*	7	7	6
T_9/T_{19}	7.5	7	7	7
C_{10}/C_{20}	*	*	*	*
C2′-endo[b]	10	6	10	6
C3′-endo[b]	2	8	2	8

[a] Some coupling constants could not be measured (−), while others were unavailable due to peak overlap (*). [b] Values for canonical DNA[25].

Interstrand Restraints. The NMR data define a total of 38 restraints which serve to hold the complementary strands of d(GAGTGCACTC)$_2$ together. Between one and five restraints connect each base to bases on the opposite strand. Restraints between G_5H1 and both protons on C_6N4 as well as a number of intermolecular NOEs between metallointercalator protons and protons on the C_6 *base* (reference Table 1) demonstrate that the cytosine base at the intercalation site (C_6) remains in the base stack and is not extrahelical as has been seen in other structures.[30,31]

Interduplex Restraints. Removed from the final restraint set but still of interest are nine relatively weak NOEs between protons on G_1 and C_{10} (G_1H8−C_{10}H5; G_1H1′−C_{10}H6; G_1H1′−C_{10}H1′; G_1H1′−C_{10}H2′/2″; G_1H1′−C_{10}H3′; G_1H5′/5″−C_{10}H6; G_1H5′/5″−C_{10}H5; G_1H5′/5″−C_{10}H3′). These protons can either be on opposite ends of the duplex or on nucleotides involved in base pairing. In either case, the interactions cannot be intraduplex and still be consistent with the restraint set. It is

(27) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(28) Veal, J. M.; Wilson, W. D. *J. Biomol. Struct. Dyn.* **1991**, *8*, 1119.

(29) (a) Babcock, M. S.; Olson, W. K. *Computation of Biomolecular Structures: Achievements, Problems, and Perspectives*; Soumpasis, D. M., Jovin, T. M., Eds.; Springer-Verlag: Heidelberg, 1993; pp 65−85. (b) Babcock, M. S.; Pednault, E. P. D.; Olson, W. K. *J. Biomol. Struct. Dyn.* **1993**, *11*, 597. (c) Babcock, M. S.; Pednault, E. P. D.; Olson, W. K. *J. Mol. Biol.* **1994**, *237*, 125.

(30) Lipscomb, L. A.; Zhou, F. X.; Presnell, S. R.; Woo, R. J.; Peek, M. E.; Plaskon, R. R.; Williams, L. D. *Biochemistry* **1996**, *35*, 2818.

(31) (a) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357. (b) Reinisch, K. M.; Chen, L.; Verdine, G. L.; Lipscomb, W. N. *Cell* **1995**, *82*, 143.

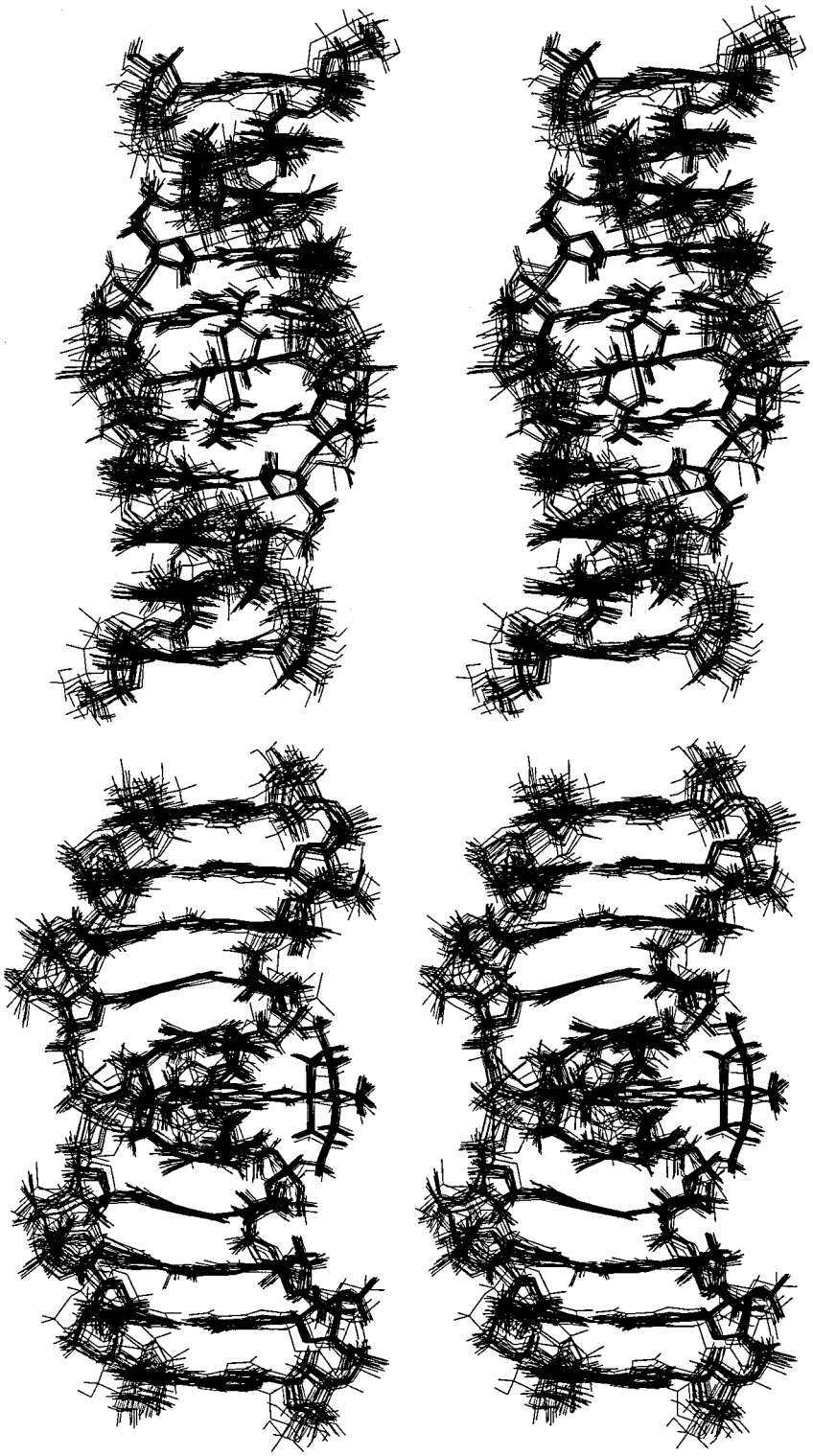

Figure 2. Two stereoviews, one into major groove, the other rotated 90° about the helix axis, of the solution structure of the bound Δ-α-[Rh-[(*R*,*R*)-Me_2trien]phi]$^{3+}$/d(GAGTGCACTC)$_2$ complex.

therefore necessary to describe these data in terms of an *interduplex* interaction, where the GC base pair at the 5′ end of one duplex stacks upon the CG base pair at the 3′ end of the other. Interduplex cross-peaks have not been noted either in spectra of identical samples at higher temperature (25 °C) or in spectra of d(GAGTGCACTC)$_2$ free (no metallointercalator) in solution at 4 °C.

Coupling Constants. Spin−spin coupling constants measured at 27 °C for d(GAGTGCACTC)$_2$, both free in solution and bound to Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$, are shown in Table 1. These spectra were collected at a higher temperature than those in the NOESY experiments to obtain high enough signal intensity for measurement. Even then, sufficient intensity was only observed for H1′−H2′/2″ cross-peaks. Some peaks were unavailable due to peak overlap, while no peaks at all were observed for C_6(bound). Also included in Table 1 are literature values for these coupling constants for C2′-endo- and C3′-endo-deoxyribose sugars.[25] Where they are not obscured by overlap, values measured for free and bound DNA appear to be approximately the same. G_1, G_3, and G_5 have coupling

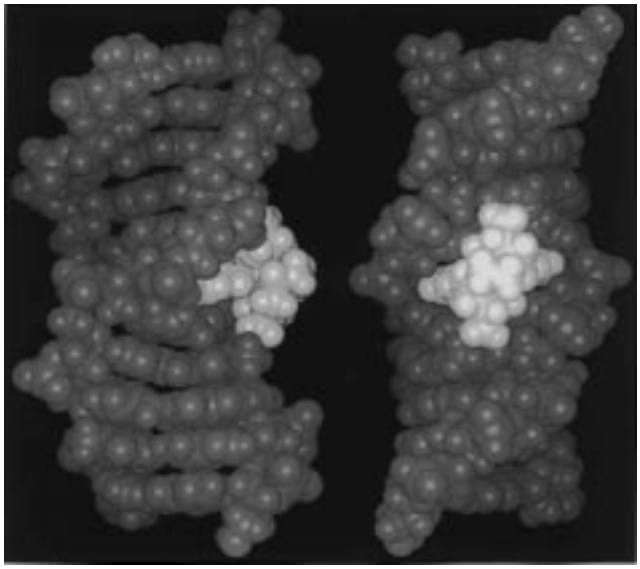

Figure 3. Space-filling views, one into the major groove, the other rotated 90° about the helix axis, of one representative structure of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ (white) bound to d(GAGTGCACTC)$_2$ (blue).

constants qualitatively consistent with C2′-endo conformation, while those for C_8 are T_9 are ambiguous.

Determination of Structure. Molecular dynamics and energy minimization calculations performed on 20 disparate starting structures (vide supra) result in 20 final structures, shown superimposed in Figure 2 (with a space-filling view of a representative structure shown in Figure 3) The average rms deviation among the 20 structures (all atoms) is 1.0 ± 0.2 Å. On average, each structure violates only 3 restraints by more than 5% of their upper limits, with a maximum average distance violation of 0.3 Å. On average, only 1 torsional restraint is violated by more than 5%, with a maximum average violation of 3°. The average rmsd between structures for the metallointercalator and the interacting DNA residues (i.e., T_4, G_5, T_{14}, and G_{15}) is 0.7 ± 0.1 Å. In general, residues at or near the intercalation site (5′-G_5C_6-3′) are more easily superimposed than those toward the ends of the duplex. The average rmsd for the metallointercalator alone is 0.5 ± 0.2 Å. The rmsd of the metallointercalator from its crystal structure is 0.6 ± 0.1 Å. Differences between the crystal structure of the metal complex and its solution structure are the result of two effects, (1) the ancillary Me_2trien is conformationally flexible, and its axial methyl groups are displaced in the direction of the DNA major groove by 0.6 Å; and (2) the modified force field used for calculations maintains a planar phi ligand and an octahedral rhodium center, removing the small distortions in these moieties that are found in the crystal structure.

As described previously,[1] the metallointercalator Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ is shown to bind d(GAGTGCACTC)$_2$ by tight, deep intercalation at the center 5′-TGCA-3′ binding site. Average helical parameters for the final structures are shown in Figure 4. In general, helical parameters and qualitative examination show the oligonucleotide to be the B-form. As expected, intercalation at 5′-G_5C_6-3′ is characterized by a doubled rise at that particular step. The net unwinding across the decamer relative to canonical B-DNA is 39 ± 4°. This is consistent with the initial model and with standard models of intercalation,[32,33] where increased rise is accompanied by a concomitant unwinding at the intercalation site. However, in this case, the intercalated base step displays a winding angle of 53 ± 3°, an *overwinding* of approximately 17°. The net unwinding is the result of unwinding at flanking base steps (primarily 5′-G_3T_4-3′ and, by symmetry, 5′-A_7C_8-3′). The intercalation site is also characterized by a large positive slide at the 5′-T_4G_5-3′, 5′-$G_5C_6$3′, and 5′$C_6A_7$3′ steps. These effects are displayed in Figure 5, which shows the 5′-T_4G_5-3′ and 5′-G_5C_6-3′ steps and illustrates these changes. The 5′-T_4G_5-3′ and 5′-C_6A_7-3′ steps also experience shifting which pulls the TA base pairs farther into the major groove relative to the 5′-G_5C_6-3′ step.

Changes in other helical parameters may compensate for some of the significant perturbations seen at the intercalation site. These include a tilting of the intercalation site (a negative tilt at 5′-T_4G_5-3′, a positive tilt at 5′-C_6A_7-3′, and no significant tilt at 5′-G_5C_6-3′), buckling of base pairs on either side of the intercalator (diminishing with distance from the intercalator), and positive stagger of the GC base pairs of the intercalation site. Average propeller twist is −14°, which aids in maximizing hydrophobic stacking area.[34] Other helical parameters are not altered significantly at the metallointercalator binding site but do show perturbations from canonical values toward the duplex ends, possibly indicative of fraying. The intercalation site itself appears to be *pulled* into the major groove by its interaction with the metallointercalator. This observation has also been made for Flexi-Di, a derivative of ditercalinium, a bisintercalator which also binds in the major groove.[35]

Intermolecular Interactions. Groups involved in intermolecular hydrogen bonding and van der Waals interactions are shown in Figure 6. The distance between the methyl carbon of the metallointercalator and the thymine methyl carbon atom is 3.0 ± 0.1 Å (averaged between two positions on twenty structures), consistent with the predicted van der Waals interaction and the observed NOE cross-peak between the associated protons. The relative positioning of the two methyl groups, however, is significantly different from that in the original model,[1] in which the methyl group of the metallointercalator approached the thymine methyl group from the direction of the interstrand hydrogen bonds. In the solution structure, the thymine base has experienced a large positive slide which, in combination with other factors, has pulled the T_4 methyl group past the methyl group of the Me_2trien ligand, placing this methyl group into what appears to be a hydrophobic "pocket" formed by the C2′ methylene, the CH6 group, and the methyl group of the T_4 base. Possible van der Waals interactions with the T_4 ring (distance to C6 = 3.5 Å) and the CH2′ methylene group of the T_4 deoxyribose sugar (4.6 ± 0.2 Å C−C distance) are consistent with the NOESY data.

Also seen in the structures are four possible intermolecular hydrogen bonds. The overwinding and positive slide at the intercalation site places the axial amine of the metallointercalator close to the N7 of G_5 with an N−N distance of 3.4 ± 0.1 Å. The O6 of G_5, while more distant (N−O distance 4.1 ± 0.1 Å), may also be involved in a hydrogen bond, suggesting that binding of the metallointercalator and the conformational changes of the intercalation site are stabilized by hydrogen bonds from the axial amines of the metallointercalator to both the N7 and O6 of G_5.

Correlations of Chemical Shifts with Structure. Chemical shift changes between free and bound DNA provide an "internal

(32) Sobell, H. M. *Prog. Nucleic Acid Res. Mol. Biol.* **1973**, *13*, 153.
(33) Neidel, S.; Berman, H. M. *Prog. Biophys. Mol. Biol.* **1983**, *41*, 43.
(34) Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025.
(35) Peek, M. E.; Lipscomb, L. A.; Bertrand, J. A.; Gao, Q.; Roques, B. P.; Garbay-Jaureguiberry, C.; Williams, L. D. *Biochemistry* **1994**, *33*, 3794.

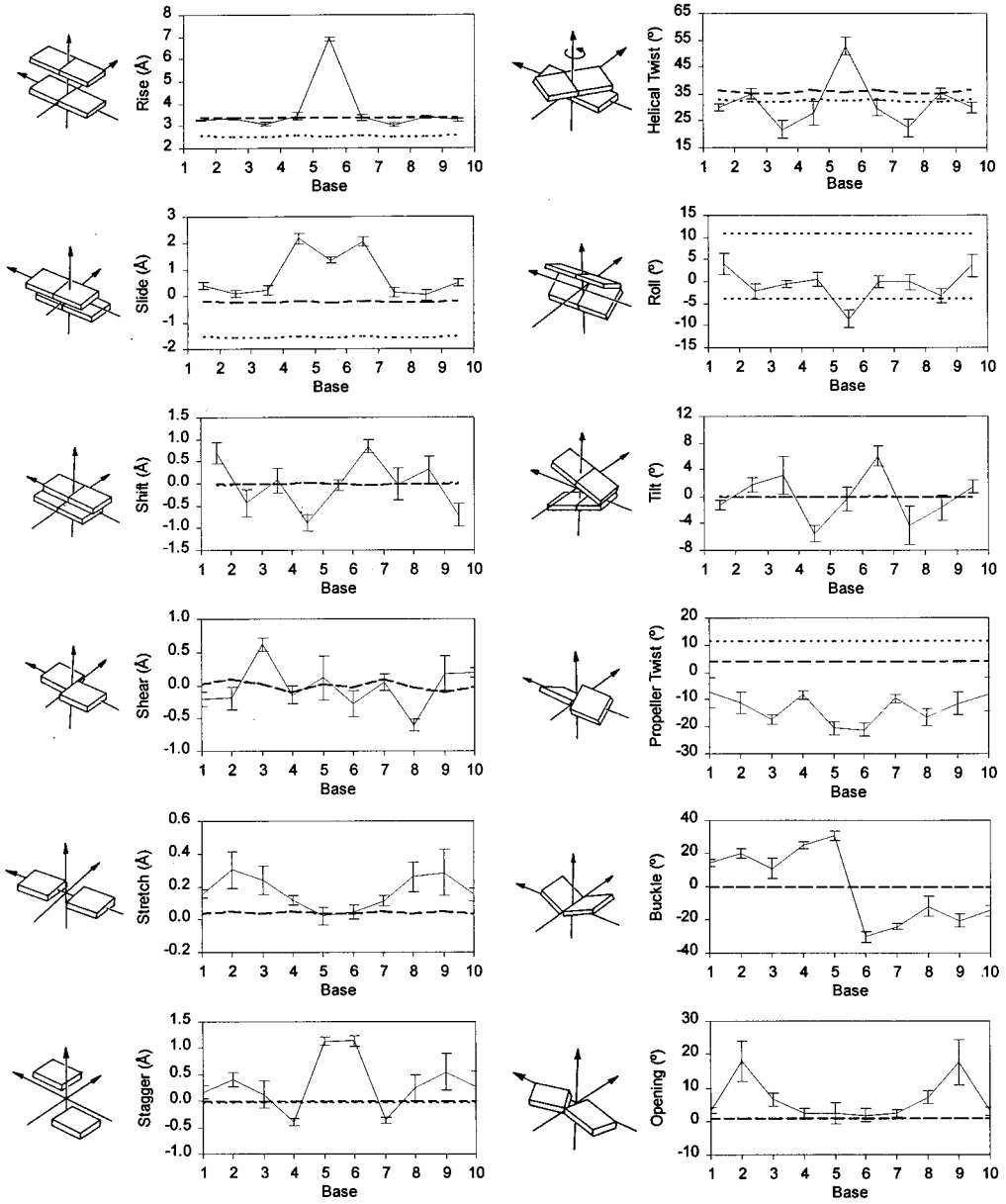

Figure 4. Selected average helical parameters for base pairs and double-stranded base steps in the solution structure of the bound Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$/d(GAGTGCACTC)$_2$ complex (−) versus canonical B-DNA (- - -) and canonical A-DNA (···). Error bars represent standard deviation.

control" for a structure derived from NOE distance restraints. The DNA base stack and metallointercalator phi ligand provide an environment dominated by aromatic ring currents. Any significant perturbation in nucleic acid conformation, especially upon binding of an intercalator, is likely to be accompanied by an ensemble of chemical shift changes.

Chemical Shifts of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$. Table 2 shows chemical shifts changes for Δ-α-[Rh[(*R*,*R*)-Me_2trien]-phi]$^{3+}$ upon binding to d(GAGTGCACTC)$_2$. As has been noted previously, both for this metallointercalator and for others,[12,13] the nonexchangeable protons of the phi ligand experience large ($\Delta\delta = -0.5$ to -1.2 ppm) upfield chemical shift changes upon intercalation into DNA. The phi imine protons show an upfield chemical shift change as a result of the DNA base ring currents, but of lesser magnitude due to their position slightly removed from the base stack. It is of note that the phi imine, while observed in the DNA-bound sample, is not observed in a sample of free metallointercalator under identical conditions. Acidification of the free rhodium complex was required for the imine resonance to be seen, suggesting that the pK_a of the imine protons is increased slightly upon intercalation of the phi ligand into DNA. The nonexchangeable protons on the ancillary Me_2-trien ligand (H3″ and methyl group 1′ protons) which are positioned to interact with the floor of the DNA major groove experience greater chemical shift perturbations than those that are not so positioned. The chemical shift changes of the Me_2-trien amine protons are difficult to interpret due to the complicated dependence of their signal position on pH and other solution conditions (data not shown).

DNA Chemical Shift Changes. The features of the 2D NOESY spectrum of d(GAGTGCACTC)$_2$ are consistent with B-form DNA and, as such, the structure of the decamer free in solution is assumed to be similar to canonical B-DNA for the purposes of these analyses. Table 3 shows the changes in chemical shift that occur for protons in d(GAGTGCACTC)$_2$ upon binding of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$. Several protons are affected directly by the intercalation of the phi ligand into the base stack. G_5H1 is located in the center of the base stack directly above the phenanthrene of the phi ligand and therefore experiences greater shielding ($\Delta\delta = -0.54$ ppm).

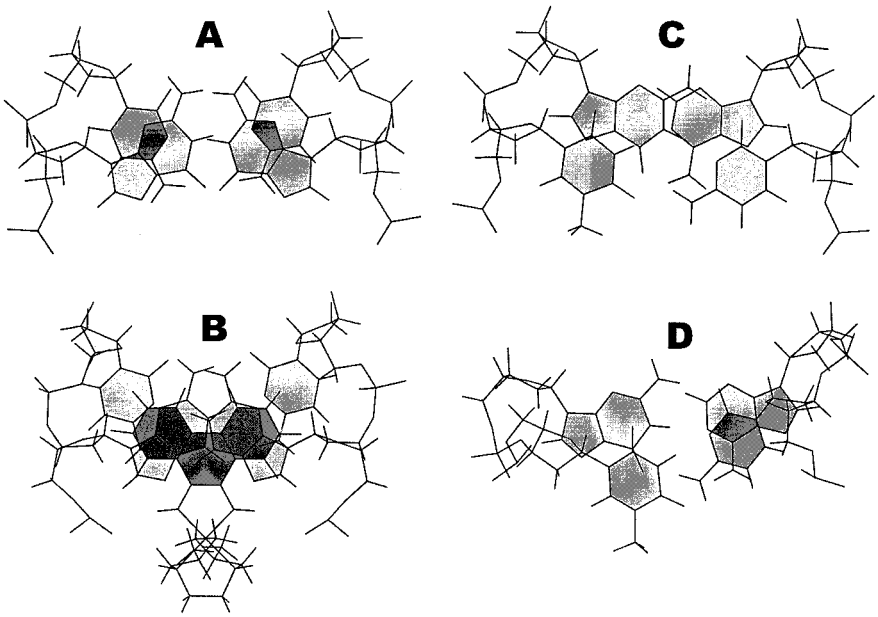

Figure 5. Excerpted base steps from representative structures viewed down the local helix axis: (A) $(G_5C_6)\cdot(G_{15}C_{16})$ from canonical B-DNA; (B) $(G_5C_6)\cdot(G_{15}C_{16})$ from solution structure; (C) $(T_4G_5)\cdot(C_{16}A_{17})$ from canonical B-DNA; (D) $(T_4G_5)\cdot(C_{16}A_{17})$ from solution structure. Darker shading indicates areas of base overlap.

Table 2. Chemical Shifts[a] (ppm) of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$, Free in Solution and Bound to $d(GAGTGCACTC)_2$ at 4 °C in 10 mM Sodium Phosphate (pH 7.0), 20 mM NaCl

	δ_{bound}	δ_{free}	$\Delta\delta^b$
Me(1′,10′)	1.02	1.24	−0.22
2′,9′	3.48	3.52	−0.04
3′,8′	3.19	3.27	−0.08
3″,8″	3.27	3.52	−0.25
5″,6″	3.19	3.21	−0.02
5′,6′	3.70	3.74	−0.04
N(4,7)H	7.16	(7.99)	(−0.83)
$N(2,9)H_2$	6.03	(5.94)	(+0.09)
phi imine	12.89	13.13[c]	−0.24
phi(4,5)	7.76	8.33	−0.57
phi(3,6)	7.41	7.86	−0.45
phi(2,7)	6.50	7.69	−1.19
phi(1,8)	7.40	8.55	−1.15

[a] Proton chemical shifts are relative to TMSP (0.00 ppm). [b] $\Delta\delta = (\delta_{bound} - \delta_{free})$. [c] Sample required mild acidification for detection of this resonance.

Shielding of imine protons internal to the base stack is considered a useful indicator of intercalative interactions.[36] G_5-H1′ is shielded to a lesser extent ($\Delta\delta = -0.27$ ppm) due to its position above the outer edge of the intercalating ligand. G_5-H2′ is deshielded ($\Delta\delta = +0.20$ ppm) due to conformational changes which draw it toward a position in the plane of the phi ligand. C_6H4 (the proton involved in interstrand hydrogen bonding) is shifted upfield ($\Delta\delta = -0.47$ ppm) due to its position over the intercalated phi ligand. It is also possible that the high propeller twist (−21°) of the G_5C_{16} base pair(s) increases the length of the interstrand hydrogen bond, shifting the proton resonance upfield.[37] The C_6H4 proton not normally involved in an interstrand hydrogen bond experiences a large downfield chemical shift ($\Delta\delta = +0.73$ ppm). This, along with a C_6N4−T_4O4 distance of 3.0 ± 0.1 Å (4.6 Å in canonical B-DNA), suggests an additional interstrand hydrogen bond made possible by the positive slide of the T_4pG_5 step and the propeller twisting of the G_5C_{16} base pair.

Figure 5 shows the intercalated 5′-G_5C_6-3′ step as viewed down the local helix axis. A significant feature of this site is the nearly exclusive stacking of the phi ligand upon the G_5 and G_{15} bases. This purine stacking is enhanced by the positive slide and large helical twist at this site. The C_6 and C_{16} bases are excluded from interaction with the phi ligand, though the stacking area of C_6 with the adjacent A_7 base is increased. Both protons on the major groove edge of the C_6 base experience large downfield changes (C_6H6 $\Delta\delta = +0.32$ ppm; C_6H5 $\Delta\delta = +0.45$ ppm) due partially to failure of the phi ligand to replace G_5 as a stacking partner and partially to the major groove edge of C_6 being drawn toward the plane of the phi ligand by the large propeller twist and buckle at the intercalation site. The displacement of C_6 may account for the absence of signals associated with C_6H1′, possibly due to exchange broadening.

The largest chemical shift changes that occur in the decamer upon metallointercalator binding are those of T_4H1′ ($\Delta\delta = -0.79$ ppm) and T_4H2″ ($\Delta\delta = -0.77$ ppm), two protons significantly removed in space (5−6 Å) from the intercalated phi ligand. The large positive slide and major groove-directed shift of the T_4G_5 base pair step moves the deoxyribose sugar of T_4 from its normal position to one over the G_5 purine ring. T_4-H1′ and T_4H2″ point directly down at G_5. T_4H3′ ($\Delta\delta = -0.22$ ppm) and T_4H2′ ($\Delta\delta = -0.16$ ppm), pointed away from G_5, are farther away and less shielded. A small upfield change for G_3H2″ ($\Delta\delta = -0.20$ ppm) demonstrates the effect of the T_4 shift/slide on the 5′ side.

In addition to G_5H1 ($\Delta\delta = -0.54$ ppm), other protons internal to the base stack experience significant chemical shift changes (T_4H3 $\Delta\delta = +0.26$ ppm; G_3H1 $\Delta\delta = -0.21$ ppm; A_7H2 $\Delta\delta = +0.18$ ppm) upon binding of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ at the G_5C_6 base step. The chemical shift change of G_1H1 is negligible (−0.03 ppm). Shifting and sliding of base pairs moves the T_4H3 from a position over the G_5 base to a point over the interstrand hydrogen bonds without altering its position

(36) (a) Patel, D. J.; Shen, C. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2553. (b) Feigon, J.; Denny, W. A.; Leupin, W.; Kearns, D. R. *J. Med. Chem.* **1984**, *27*, 450. (c) Searle, M. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **1993**, *25*, 403.

(37) Wagner, G.; Pardi, A.; Wüthrich, K. *J. Am. Chem. Soc.* **1983**, *105*, 5948.

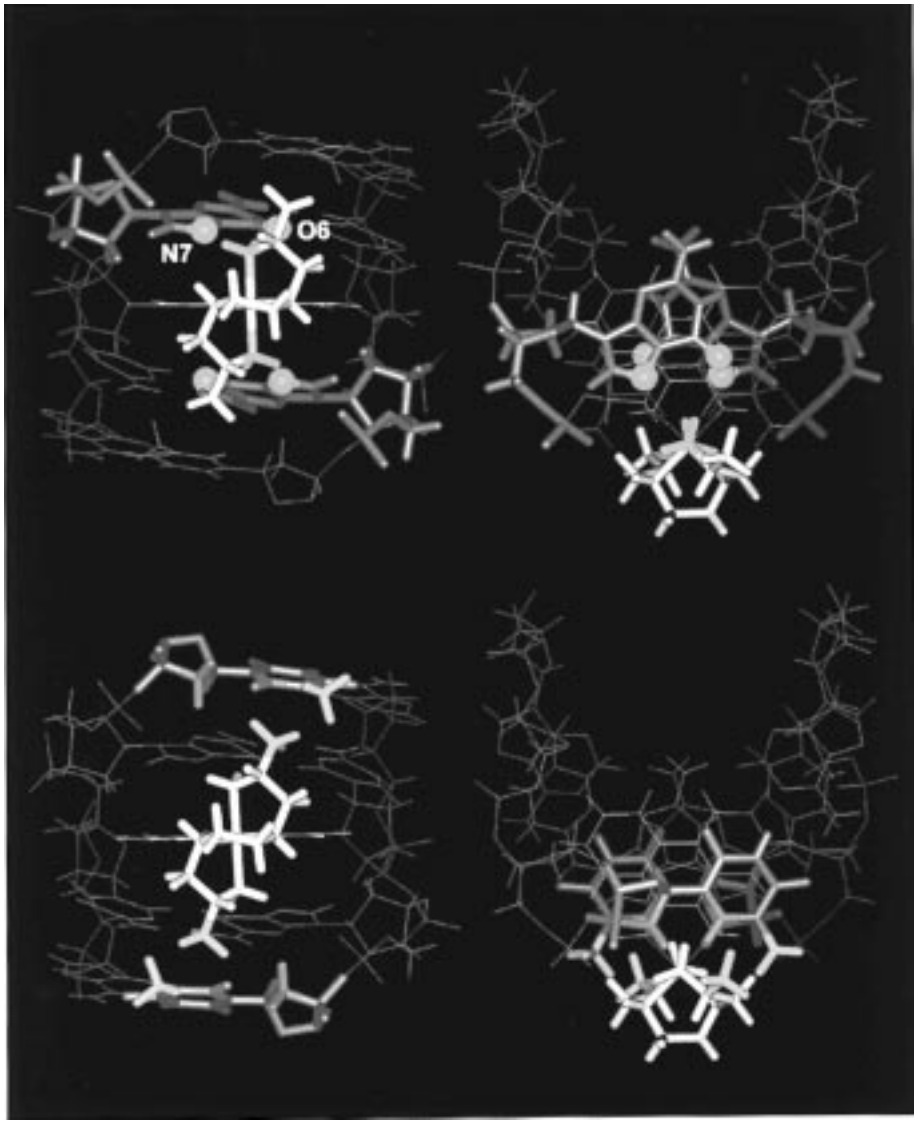

Figure 6. Detail of the binding site of representative structure. At the top are views into the major groove and down the helix axis, emphasizing the Me_2trien ligand and the G_5 and G_{15} residues. Functional groups with proposed involvement in hydrogen bonding (Me_2trien axial amines and guanine N7 and O6) are shown in green to illustrate their relative orientation and proximity ($d_{NN} = 3.4 \pm 0.1$ Å; $d_{NO} = 4.1 \pm 0.1$ Å). Shown at the bottom are identical views, but emphasizing the T_4 and T_{14} residues instead. Shown in red are the methyl groups of the thymine bases and the metallointercalator. These groups are involved in intermolecular van der Waals contacts ($d_{CC} = 3.0 \pm 0.1$ Å). Also shown in magenta are the CH6 and CH2′/2″ groups of the thymine bases, which, along with the C5 methyl group, form possible binding pockets for the axial methyl groups of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$.

Table 3. Chemical Shift Changes (ppm) upon Binding of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ to 1.0 mM[a] d((GAGTGCACTC)$_2$ at 4 °C in 10 mM Sodium Phosphate (pH 7.0)/20 mM NaCl

	H8/6	H5/2/Me	H1′	H2′	H2″	H3′	H4′	H5′/5″	imino	amino
G_1	+0.01	0	+0.01	0	+0.01	−0.01	+0.01	0	−0.03	
A_2	0	−0.05	−0.08	−0.02	−0.05	−0.02	−0.03			
G_3	+0.03		−0.06	−0.08	**−0.20**	+0.01	−0.07		**−0.21**	
T_4	+0.04	−0.09	**−0.79**	−0.16	**−0.77**	**−0.22**			**+0.26**	
G_5	−0.04	0	**−0.27**	**+0.20**	−0.14				**−0.54**	
C_6	**+0.32**	**+0.45**		**+0.23**	−0.13	−0.08				**−0.47**[b]/**+0.73**
A_7	+0.07	+0.18	+0.12	+0.18	+0.07	+0.10	+0.09			
C_8	+0.05	+0.04	+0.08	+0.04	+0.03	+0.03				−0.01[b]/+0.03
T_9	−0.05	+0.02	−0.06	−0.04	−0.02	−0.02			+0.07	
C_{10}	−0.03	−0.04	−0.04	−0.01	−0.01	−0.02				−0.01[b]/+0.02

[a] Duplex concentration. [b] The downfield-shifted CH4 proton is involved in the interstrand hydrogen bond. Large, significant chemical shift changes are in boldface.

relative to G_3. The H1 proton of G_3 moves from a point over the interstrand hydrogen bonds to a point over the edge of the A_{17} base but is not displaced relative to the A_2 base. This would deshield T_4H3 and shield G_3H1 relative to canonical B-DNA.

Discussion

^{1}H NMR combined with restrained molecular dynamics simulation and energy minimization has provided us with

an atomic-level view of a metallointercalator bound site specifically to DNA. This view provides us with information regarding several aspects of molecular recognition, including the general characteristics of the intercalation, specific metal−nucleic acid and stacking interactions as determinants of recognition, and implications for the design of new site-specific metallointercalators.

Structural Characteristics of Intercalation. Of the many structural determinations of intercalators bound to nucleic acids by either NMR or X-ray crystallography, only a few[38,39] are available for a single monointercalator bound to a relatively large oligonucleotide. There are many other monointercalators bound in the center of dinucleotide dimers or intercalated one base pair from the edge of larger oligonucleotides.[33] Structures with a single intercalating molecule bound toward the center of a long oligonucleotide are more relevant to the conditions under which many of these drugs have their activity, that is, in genomic DNA at low drug concentrations. Δ-α-[Rh[(*R*,*R*)-Me_2-trien]phi]$^{3+}$, though not presently known to possess any therapeutic properties, fulfills these requirements when bound exclusively at 5′-TGCA-3′ in a 1:1 complex with the decamer d(GAGTGCACTC)$_2$. As such, Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ bound to a long oligonucleotide might be considered a model system for examining the effects of intercalation upon the DNA duplex.

One of the only properties of intercalation that appears to be applicable to nearly all systems is that which defines the interaction: the intercalator is inserted between base pairs, separating them by an additional distance of approximately 3.4 Å. An additional feature that generally appears is a buckling of base pairs on either side of the intercalation site. This allows the base pairs to separate at the center of the base stack without much distortion of the backbone. We see these features with the rhodium complex.

Local Overwinding and Net Unwinding. Of particular interest is the fact that the helix need not be unwound at the intercalation site, despite the many tests for intercalation that employ unwinding as an assay. Crystal structures of dinucleotides containing intercalators show unwinding angles of 4° to 32°.[33] Structures of daunomycin,[40] nogalamycin,[41,42] and idarubicin[43] at 5′-pyrimidine-purine-3′ sites in longer DNA sequences show small amounts of unwinding and in some cases overwinding, but in general show little change from the canonical B-DNA helical twist of 36° at the intercalation site. Interestingly, as we see with the rhodium complex, the DNA duplex in these structures is typically unwound at the base step *adjacent* to the intercalation site, such that a net unwinding is observed for these oligonucleotides regardless of the helical twist at the intercalated base step.

Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ intercalates at the central 5′-GC-3′ step of d(GAGTGCACTC)$_2$, and both overwinding (winding angle of 53°) and positive slide (>1 Å) at the intercalation site allow the phi ligand to stack entirely on the guanine bases. There is overall unwinding of the decamer, however, despite the appreciable overwinding at the intercalation site. As in other structures of intercalators bound to DNA, the unwinding occurs at the base steps adjacent to the intercalation site, with the most unwinding occurring at the 5′-G_3T_4-3′ base step, which forms the apparent boundary between the site of interaction (5′-TGCA-3′) and the ends of the decamer, which are closer to canonical B-form DNA.

The structural features of this complex show parallels to general features of actinomycin D, which also binds at a 5′-GC-3′ (5′-purine-pyrimidine-3′) step, but from the minor groove. Though X-ray analysis shows multiple forms with varying helical parameters,[39] the solution structure of actinomycin D bound to the central step of d(AAAGCTTT)$_2$ is observed to have a slight overwinding (39° total winding angle) at the intercalation site with stacking occurring entirely between the guanine bases and the intercalated phenoxazone moiety. In that case, adjacent base steps are unwound to accommodate the bulky depsipeptide moieties in the minor groove.[44] The perturbations seen in actinomycin D solution structure are therefore similar to those of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ at the same 5′-GC-3′ base step, though the overwinding for the metallointercalator is more extreme. Large helical twist angles dictated by unusual stacking arrangements are not unprecedented: the solution structure of a decamer containing tandem GA mismatches maximizes stacking area by adopting a 98° helical twist angle at that base step.[45] In that case also, flanking base steps experience significant unwinding (16° helical twist angles).

Deoxyribose Sugar Conformation. Structures of intercalated dinucleotides[33] and the only other structure of a DNA-bound metallointercalator, Pt(terpy) intercalated into 5′-CpG-3′,[46] provided an early model for deoxyribose conformation at intercalation sites. These structures showed a C2′-endo (B-DNA-like) conformation at the 3′ nucleotide and a C3′-endo (A-DNA-like) conformation at the 5′ nucleotide. Structures of monointercalators daunomycin,[40] nogalamycin,[42] idarubicin,[43] and actinomycin D[44] in longer oligonucleotides appear to favor C2′-endo or C1′-exo conformations for sugars on both the 5′ and 3′ sides of intercalation sites for nonend residues. 5′- or 3′-end residues show no discernible pattern, suggesting that data for end residues cannot necessarily be compared to data for internal residues. The structure of d(GAGTGCACTC)$_2$ bound to Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ shows the 5′-nucleotide G_5 to exist consistently in the C2′-endo and closely related C3′-exo conformations, similar to other monointercalators intercalated into long sequences.

Propagation of Distortion along the Helix. End effects in DNA structure can be quite pronounced, as typified in this structure by disorder and the tendency toward A-DNA-like deoxyribose conformations in 3′-end residues. Structures of other monointercalators[40,42,43] have demonstrated that end residues next to intercalators show little order in terms of deoxyribose conformation. It is important that an intercalation site be placed far into the interior of an oligonucleotide for the observed effects of intercalation to be generally meaningful in terms of longer sequences of DNA and to give a realistic picture

(38) (a) Kumar, R. A.; Ikemoto, N.; Patel, D. J. *J. Mol. Biol.* **1997**, *265*, 173. (b) Vanderwall, D. E.; Lui, S. M.; Wu, W.; Turner, C. J.; Kozarich, J. W.; Stubbe, J. *Chem. Biol.* **1997**, *4*, 373. (c) Wu, W.; Vanderwall, D. E.; Turner, C. J.; Kozarich, J. W.; Stubbe, J. *J. Am. Chem. Soc.* **1996**, *118*, 1281. (d) Pavlopoulos, S.; Bicknell, W.; Craik, D. J.; Wickham, G. *Biochemistry* **1996**, *35*, 9314. (e) Lian, C.; Robinson, H.; Wang, A. H.-J. *J. Am. Chem. Soc.* **1996**, *118*, 8791. (f) Gao, X.; Stassinopoulos, A.; Rice, J. S.; Goldberg, I. H. *Biochemistry* **1995**, *34*, 40. (g) Brown, D. R.; Kurz, M.; Kearns, D. R.; Hsu, V. L. *Biochemistry* **1994**, *33*, 651.

(39) (a) Kamitori, S.; Takusagawa, F. *J. Am. Chem. Soc.* **1994**, *116*, 4154. (b) Kamitori, S.; Takusagawa, F. *J. Mol. Biol.* **1992**, *225*, 445.

(40) Nunn, C. M.; Van Meervelt, L.; Zhang, S.; Moore, M. H.; Kennard, O. *J. Mol. Biol.* **1991**, *222*, 167.

(41) Egli, M.; Williams, L. D.; Frederick, C. A.; Rich, A. *Biochemistry* **1991**, *30*, 1364.

(42) (a) Smith, C. K.; Brannigan, J. A.; Moore, M. H. *J. Mol. Biol.* **1996**, *263*, 237. (b) Liaw, Y.-C.; Gao, Y.-G.; Robinson, H.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H.-J. *Biochemistry* **1989**, *28*, 9913.

(43) Gallois, B.; d'Estaintot, B. L.; Brown, T.; Hunter, W. N. *Acta Crystallogr.* **1993**, *D49*, 311.

(44) Liu, X.; Chen, H.; Patel, D. J. *J. Biomol. NMR* **1991**, *1*, 323.

(45) Greene, K. L.; Jones, R. L.; Li, Y.; Robinson, H.; Wang, A. H.-J.; Zon, G.; Wilson, W. D. *Biochemistry* **1994**, *33*, 1053.

(46) Wang, A. H.-J.; Nathans, J.; van der Marel, G.; van Boom, J. H.; Rich, A. *Nature* **1978**, *276*, 471.

of how the effects of binding propagate through the helix. In this system, 5′-TGCA-3′ is in the exact center of the decamer, which allows for two trimer flanking sequences. The solution structure shows that though the 4 base pairs of the binding site are distorted from B-form, the flanking sequences are not significantly different from B-DNA. However, as significant unwinding does occur at the 5′-G_3T_4-3′ and 5′-A_7C_8-3′ base steps which define the edges of the recognition site, it is possible that different flanking bases, by virtue of different stacking interactions, may have an effect on binding that is not observed here.

Determinants of Recognition. The most easily identified determinants of site specificity and conformation are the intermolecular hydrogen bonds and van der Waals contacts that occur in the major groove. These are, of course, the elements of recognition that were originally designed to give Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ its specificity for 5′-TGCA-3′. The metallointercalator was designed[19] such that the amino groups in the 2- and 9- positions of the 2*R*,9*R*-diamino-4,7-diazadecane ((*R*,*R*)-Me_2trien) ligand were located in positions to form hydrogen bonds with the O6 of the guanine residues at the intercalated base step, while the methyl groups (the 1 and 10 positions) were positioned for favorable van der Waals contacts with the methyl groups of the thymine bases located to the 5′ side of the intercalated base step. The binding site model was canonical B-form $d(TGCA)_2$ with the central 5′-GC-3′ base step opened an additional 3.4 Å and unwound by 20°. The metallointercalator was docked into the structure to optimize the designed interactions. Though it is known that this class of metallointercalators binds from the major groove,[12,13] it was assumed that both the conformations of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ and its binding site would remain relatively unchanged from those in the model. As shown here, that is not the case, and indeed deviations from the model arise which enhance specificity.

Major Groove Contacts. Hydrogen bonding appears to be stronger and more specific than anticipated during design. The results show that the axial amine of the metallointercalator interacts with both the N7 and O6 of G_5 at the intercalation site. It is possible that preferential binding to 5′-GC-3′ steps occurs because pyrimidine bases lack a functional group analogous to N7 and the presence of the amino group at the N6 position of adenine constitutes an unfavorable interaction with the axial amine of the metallointercalator. A 5′-TX-3′ step would present a favorable hydrogen bonding interaction in the major groove (axial amine to thymine O4), but the site is occluded by the methyl group at the thymine 5 position. Cytosine presents no favorable major groove interactions.

The methyl−methyl van der Waals contact is also more complex than originally anticipated. Sliding and shifting of the T_4 base relative to the canonical B-DNA starting models, though, appears to have several favorable results. The first is that it brings the T_4 and metallointercalator methyl groups close to one another. The second advantage is that the base has been drawn *past* the metallointercalator's methyl group such that this group appears to lie within a hydrophobic pocket defined by the major groove edge of the T_4 base and deoxyribose sugar. In addition, the A_{17} base (paired to T_4) increases its stacking overlap with the C_{16} base adjacent to it. These results show that the axially placed functional groups (in this case methyl) may have a much larger effect on the binding site conformation and, therefore, sequence selectivity of the metallointercalator than originally predicted. Moreover, further optimization of the interaction may be possible.

Stacking of the Metallointercalator in DNA. It is known that purines stack effectively.[47] In addition, phenanthrene has been found experimentally to stack more completely on nucleic acid bases than do other natural nucleic acid bases, primarily due to the exclusion of aromatic surface area from solvent.[48] As shown in Figure 5, total aromatic stacking overlap area in the intercalation site of the bound structure is approximately 3 times that seen in the unbound 5′-G_5pC_6-3′ step of canonical B-form DNA. In addition, the stacking area at the 5′-T_4pG_5-3′ step is increased by local conformational changes. However, the area of overlap in a model with Δ-α-[Rh[(*R*,*R*)-Me_2trien]-phi]$^{3+}$ bound to canonical B-DNA with an engineered intercalation site (doubled rise, 20° unwinding at 5′-G_5C_6-3′) is roughly equivalent to that in the solution structure (data not shown). This suggests that even though the intercalation of the phi ligand may be driven in part by the favorable van der Waals contacts that occur with aromatic stacking, this effect does not necessarily cause the intercalation site to favor an overwound positive slide conformation relative to an underwound one.

The successful application of metallointercalators depends in many respects upon this ability to stack well within the helix. Electron transfer depends on electronic coupling arising from close stacking of donors and acceptors within the helix. The structure of CuTMPyP4 [copper(II) *meso*-tetra(*N*-methyl-4-pyridyl)porphyrin] bound to DNA,[30] for example, showed poor stacking due to the extrahelical position of the cytosine residue adjacent to the site of pseudointercalation, and suggests that this system would constitute a poor one for DNA-mediated electron transfer relative to octahedral Rh(III) and Ru(II) systems. It is interesting that in our structure of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$, because of the complete stacking of the phi between the guanines across the strand, the complementary cytosine residues are not stacked with the intercalator, despite remaining hydrogen bonded to the guanines; instead stacking of C_6 (and C_{16}) with the adjacent adenine residues is increased.

Correlation with Photocleavage. It has been observed previously from NMR and molecular modeling studies of Rh-$(en)_2$phi bound to DNA that the H2′ proton of the 5′-nucleotide is the deoxyribose sugar proton in closest proximity to the intercalated phi ligand.[13] The observation was made that though abstraction of this proton or the H2″ proton is energetically disfavored by 2 kcal/mol relative to other deoxyribose protons,[49] its removal followed by fast atom migration to C3′ could be involved in the initiation of the observed UV-induced photocleavage reaction.[15] The results presented here show that the H2′ and H2″ protons of G_5 are those that are closest to the phi ligand of Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ (distance H2′ to C7 = 2.7 ± 0.1 Å; distance H2″ to C7 = 2.4 ± 0.1 Å). Though there is no intermolecular restraint for G_5H2″-phi(2,7), there is a large NOE (the restraint was rejected from the data set because of poor fit to the second-order polynomial). The H3′ proton of the G_5 deoxyribose sugar is 4.5 ± 0.1 Å distant from the intercalated phi ligand. In addition, no proton on the C_6 deoxyribose sugar is located within 4.5 Å of the phi ligand. Previous photocleavage data show that it is the 5′-G of the intercalation site that is cleaved by Δ-α-[Rh[(*R*,*R*)-Me_2trien]-phi]$^{3+}$ under UV irradiation.[19] These new results further suggest that it is the H2′ or H2″ of the 5′-G that is abstracted during photocleavage.

(47) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1983.
(48) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X.-F.; Sheils, C. J.; Paris, P. L.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **1996**, *118*, 8182.
(49) Miaskiewicz, K.; Osman, R. *J. Am. Chem. Soc.* **1994**, *116*, 232.

Site Selectivity. Implications for Minimalist Design. Finally, it is important to note the significance of the site selectivity of this small metallointercalator with regard to structural determination. Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ recognizes 5′-TGCA-3′, a 4 base pair site, yet has a molecular weight (for the complex cation) of 483.4 and an end-to-end distance across the Me_2trien ligand of only 9 Å, actually less than the length of the DNA site to which it binds. The combination of its limited bulk and major groove targeting means that the binding site need not be distorted (beyond admitting the intercalating ligand) to accommodate it, as must occur for the minor groove to admit the bulky depsipeptides of actinomycin D.[39,44] As such, the changes that occur in the DNA duplex upon binding are purely a consequence of the intercalative interaction and the intermolecular major groove contacts that determine site specificity. The altered conformation of the DNA binding site demonstrates that the relative rigidity of the rhodium(III) complex has, in and of itself, proved to be a useful tool for design. Since DNA has demonstrated a large degree of conformational plasticity both here and elsewhere,[17,50] rigidity in the DNA-binding molecule simplifies the molecular recognition design problem by defining at least one-half of the interaction. Because of the site selectivity, the recognition site can be nested in the interior of a much longer sequence, as has been done here with d(GAGTGCACTC)$_2$, so that the results are analogous and more relevant with regard to interactions with bulk DNA. One can see that the interaction of the octahedral metallointercalator with DNA, oriented by intercalation, permits the minimalist predictive design of a site-selective binding agent, one not unlike a much larger DNA-binding protein in affinity and specificity. The platform for intercalation defines the coordinate system for an ensemble of interactions between the metallointercalator and the DNA bases in a system that one day might conceivably be used to define a library of DNA−binding agents.

Summary

The solution structure of the metallointercalator Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$ has been determined using ^{1}H NMR techniques. Distance restraints from NOESY experiments in D_2O and H_2O and torsional restraints from ECOSY experiments have been used to define a family of structures whose structural perturbations from B-DNA are consistent with chemical shift changes that occur in the DNA decamer upon binding. Though the oligonucleotide experiences an overall unwinding, the structural changes at the binding site itself include overwinding of the intercalated base step with some unwinding at the flanking base steps and the most unwinding at the base steps that define the edge of the 5′-TGCA-3′ binding site. The shifted, tilted, and highly slid binding site maximizes the number of possible intermolecular interactions, including up to four intermolecular hydrogen bonds, two definite and several possible van der Waals contacts, and extremely favorable stacking interactions. The intercalating ligand stacks entirely upon guanine bases at the intercalation site, an interaction which may be relevant from the standpoint of both affinity and recognition. The sequences flanking the 5′-TGCA-3′ binding site are essentially B-form, demonstrating that the effects of intercalation do not extend significantly beyond the binding site. This minimalist structure of a synthetic complex intercalated site selectively with a DNA duplex, with the intercalating ligand providing a platform to orient a range of noncovalent contacts of the octahedral complex within the DNA major groove, provides a useful perspective for the predictive design of metallointercalators which bind specific DNA sequences.

Acknowledgment. We are grateful to the NIH (GM33309) for their financial support. B.P.H. thanks the NSF, NIH, and Ralph M. Parsons Foundation for predoctoral fellowships. We also thank Cynthia Dupureur and Sonya Franklin for helpful discussions and Silvia Cavagnero of the Caltech NMR facility for technical assistance.

Supporting Information Available: Tables of proton chemical shifts for d(GAGTGCACTC)$_2$, free in solution and bound to Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$. Tables (12) of experimental distances and torsional restraints with the corresponding average structural distances and angles. Table containing brief descriptions of all starting structures. Full contour plot of 120 ms NOESY spectrum of 1:1 Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$/d(GAGTGCACTC)$_2$ in D_2O at 4 °C with additional figure depicting partial aliphatic region of same. Full contour plot of 125 ms WATERGATE NOESY spectrum of 1:1 Δ-α-[Rh[(*R*,*R*)-Me_2trien]phi]$^{3+}$/d(GAGTGCACTC)$_2$ in 90% H_2O, 10% D_2O at 4 °C with two additional figures depicting the imine−imine and imine−amine/aromatic regions of same. Figure depicting superimposition of Δ-α-[Rh[(*R*,*R*)-Me_2trien]-phi]$^{3+}$ molecules from solution structures with mirror-image of Λ-α-[Rh[(*S*,*S*)-Me_2trien]phi]$^{3+}$ crystal structure (23 pages, print/PDF). See any current masthead page for ordering information and Internet access instructions.

JA974134X

(50) Kim, Y. C.; Hahn, S.; Sigler, P. B. *Nature* **1993**, *365*, 512.