Differential DNA Recognition by the Enantiomers of 1-Rh(MGP)$_2$phi: A Combination of Shape Selection and Direct Readout†

Sonya J. Franklin‡ and Jacqueline K. Barton*

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

Received July 27, 1998; Revised Manuscript Received September 10, 1998

ABSTRACT: The enantiomers of the symmetric metallointercalator complex 1-Rh(MGP)$_2$phi^{5+} [MGP = 4-(guanidylmethyl)-1,10-phenanthroline; phi = phenanthrenequinone diimine] bound to DNA decamer duplexes containing their respective 6 bp recognition sequences have been investigated using ^{1}H NMR. Shape selection due to the chirality of the metal center and hydrogen-bonding contacts of ancillary guanidinium groups to 3′-G N7 atoms define the recognition by complexes which bind by intercalation to duplex DNA. The titration of Λ-Rh into the self-complementary decamer containing the recognition sequence (5′-GACATATGTC-3′, L1) resulted in one symmetric bound conformation observed in the ^{1}H NMR spectrum, indicating that the DNA duplex retains its symmetry in the presence of the metal complex. Upfield chemical shifts of duplex imino protons and the disruption of the NOE base−sugar contacts defined the central T5-A6 intercalation site. The downfield shift of the G8 imino proton supports the conclusion that the pendant guanidinium arms make simultaneous H-bonding contacts to the N7 atoms of 3′-G8 bases on either side of the site. A variable-temperature study of a partially titrated sample (2:3 Λ-Rh/L1) showed the exchange rate (k_{obs}) at 298 K to be 68 s^{-1} and the activation barrier to exchange ($\Delta G^{\ddagger}$ of association) to be 2.7 kcal/mol, a value comparable to the stacking energy of one base step. The results presented coupled with biochemical data are therefore consistent with binding models in which Λ-1-Rh(MGP)$_2$phi^{5+} (Λ-Rh) traps the recognition site 5′-CATATG-3′ in an unwound state, permitting intercalation centrally and hydrogen bonding to guanines at the first and sixth base pair positions. The data suggest a different model of binding and recognition by Δ-Rh. The titration of Δ-Rh into a DNA decamer containing the 6 bp recognition site (D1, 5′-CGCATCTGAC-3′; D2, 5′-GTCAGATGCG-3′) resulted in two, distinct conformers, in slow exchange on the NMR time scale. The rate of exchange between the two conformers (k_{obs}) at 298 K is 37 s^{-1}, most likely due to partial dissociation between binding modes. The slower rate relative to Λ-Rh association reflects the relative rigidity of the D1 and/or D2 sequence in comparison to L1. NOE cross-peaks between the intercalating phi ligand and protons of T5-C6, as well as the upfield shifts observed for imino protons at this step, serve to define the central T5-C6 step as the single site of intercalation. The downfield shift of the 3′-G imino protons indicates the complex makes hydrogen bond contacts with these bases. The complex, which is too small to span a 6 bp B-form DNA sequence, nonetheless makes major groove contacts with 3′-G bases to either side of the site. Notably, both 3′-guanine bases are necessary to impart site specificity and slow dissociation kinetics with the 5′-CATCTG-3′ site, as evidenced by the extremely exchange-broadened two-dimensional NOESY spectra of Δ-Rh bound to modified duplexes containing N^7-deazaguanine at either G8 or G18; the loss of one major groove contact completely abolishes specificity for 5′-CATCTG-3′. DNA chemical shifts upon binding and intermolecular NOE contacts therefore support a model in which Δ-Rh intercalates in one of two canted binding conformations. Within this model, each intercalation mode allows one guanidinium−guanine hydrogen bond at a time, while bringing the other arm close to the phosphate backbone.

To mimic nature's ability to distinguish small DNA sequences within an entire genome, we must understand and employ the principles governing that recognition. Proteins are able to perform this remarkable feat by using both direct readout and shape selection mechanisms, the combination of which allows the protein great subtlety in sequence recognition through an ensemble of noncovalent contacts (*1*). Direct readout interactions involve the relative three-dimensional orientation of various contact points in a given sequence. These interactions between amino acid side chains and the array of hydrogen-bonding and van der Waals contacts available on DNA comprise combinations of charge and shape complementarity. Additionally, proteins exploit the inherent flexibility and contour of DNA sequences as an indirect method of differentiating a recognition site. As chemists, we can employ the same geometrical and electronic principles for the rational design of small, selective metal-

† This research was supported by National Institutes of Health Grant GM 33309 and an NIH-NRSA Postdoctoral Fellowship (S.J.F.).

* To whom correspondence should be addressed.

‡ Present address: Department of Chemistry, University of Iowa, Iowa City, IA 52242.

10.1021/bi981798q CCC: $15.00
Published on Web 11/01/1998

lointercalators as models of DNA-binding proteins (*2*, *3*). By using intercalation of an aromatic ligand to provide binding affinity, and the rigid, octahedral coordination sphere as a scaffold on which to append oriented binding functionalities, we can generate specificity with these small metal complexes.

Metallointercalators have proven to be remarkably good mimics for protein recognition of specific DNA sequences, discriminating among sites through a variety of mechanisms (*2*−*8*, *30*). Δ-Metallointercalators have been shown generally to bind more tightly and deeply to right-handed DNA than their enantiomers, due to matching of the shape and symmetry with the major groove (*9*, *31*). The binding sites of these complexes are further differentiated by the depth of intercalation (*10*), the amount of steric clash between DNA and ancillary ligands (*10*, *11*), and complementary three-dimensional arrays of hydrogen-bonding and van der Waals contacts (*2*−*8*, *30*). Thus, these metallointercalators allow us to model the direct and indirect sequence recognition of proteins, such as the TATA-box binding protein, which recognizes an inherently flexible site (*12*, *32*), the Engrailed homeodomain, which recognizes and binds to a set of similar, sometimes overlapping, DNA sequences (*13*), and the *trp* repressor protein, which has been shown to use both protein−base contacts and structural recognition to discriminate among sites and adapt to base substitutions (*14*).

With a combination of several elements of recognition, such as shape selection and hydrogen bond complementarity, metallointercalators can be designed to selectively target larger, more complex sites. We have designed a metallointercalator, $Rh(MGP)_2phi^{5+}$ [MGP = 4-(guanidylmethyl)-1,10-phenanthroline; phi = phenanthrenequinone diimine], which incorporates guanidinium functionalities on the ancillary ligands so flanking G bases can be targeted (*7*, *30*). The guanidinium group of an arginine residue making hydrogen-bonding contacts with G N7 and O6 atoms in the major groove of DNA is a common recognition element of DNA binding proteins (*1*). Because of the asymmetry of the monosubstituted phen ligands (MGP), the synthesis of $Rh(MGP)_2phi^{5+}$ generates three isomers with different arrangements of the guanidinium groups with respect to the intercalating ligand, in the statistically expected ratio of 1:2:1. Isomer 1 (discussed here) has both guanidinium functionalities oriented axially, directed above and below the intercalating ligand, whereas isomer 2 has both oriented equatorially, away from the phi ligand. Isomer 3, the most abundant, has a mixed orientation. In addition, because of the chirality at the metal center, each of the three isomers is produced as a pair of Λ- and Δ-enantiomers, which can be chirally resolved. This family of six compounds allows us to explore the interplay of shape selection and flexibility at the intercalation site, as well as the effect of different orientations imposed by ancillary guanidinium contacts by each enantiomer.

The enantiomers of axially substituted isomer 1 have been shown via DNA photocleavage studies to have substantial sequence selectivity (*7*, *8*, *30*). Λ- and Δ-1-$Rh(MGP)_2phi^{5+}$ each recognize a 6 bp sequence, differing at only one base, adjacent to the intercalation site (Figure 1). Λ-1-$Rh(MGP)_2phi^{5+}$ (Λ-Rh) preferentially recognizes the site 5′-CATATG-3′, whereas Δ-1-$Rh(MGP)_2phi^{5+}$ (Δ-Rh) recognizes 5′-CATCTG-3′ (photocleavage occurs at the underlined bases). These sequences both have centrosymmetric 3′-guanine bases which are the expected targets for hydrogen-bond contacts. However, in normal B-form DNA, modeling suggests that neither enantiomer is capable of spanning the 6 bp site to make both guanine contacts simultaneously.

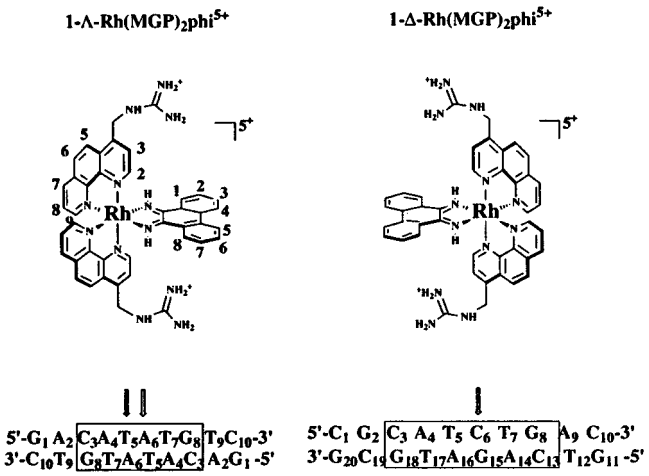

FIGURE 1: Schematic depiction of Λ- and Δ-1-$Rh(MGP)_2phi^{5+}$, with ligand proton numbering schemes indicated. The DNA decamer duplexes L1 and D1−D2 are given below with the numbering schemes used in the text. The recognition sequence is inscribed in a box, and the sites of photocleavage are indicated by arrows.

An explanation for recognition by the Λ-enantiomer has been suggested on the basis of previous biochemical assays, which measure sequence-dependent unwinding (*7*, *30*). Once bound by Λ-Rh, the flexible, AT-rich DNA recognition site has been shown to be unwound by 70° (*7*, *30*). It has not been established, however, whether the binding of the complex causes the unwinding or if the complex captures the DNA site in an unwound conformation. Previous photocleavage studies gave additional evidence for the recognition mechanism, with the dramatic change in site selectivity upon systematic sequence substitution. Varying the center of the site (from 5′-CATATG-3′ to 5′-CACGTG-3′) completely abolishes selective binding, whereas substitutions at the outside of the 6 bp site (3′-G to C or A) attenuate the preference (≈75 and 67%, respectively) (*7*, *30*). Whether promoting or trapping the unwound conformation, the observed unwinding presumably allows the complex to take advantage of the hydrogen-bonding contacts to 3′-G, though this hypothesis needs to be investigated in more detail.

Although the observed unwinding suggests how Λ-Rh may target a 6 bp site, it has been unclear how Δ-Rh recognizes similarly the outer bases of a large, relatively inflexible site, since the central 5′-TC-3′ step makes unwinding less favorable. Base substitution experiments (*7*, *30*) have shown similar reductions in the extent of photocleavage with changes at the outer base pairs of the Δ-Rh site. However, the mechanism of Δ-Rh recognition is still unclear.

Many of these structural details and energetic questions can be addressed by NMR. One-dimensional and two-dimensional NMR experiments have thus been carried out on the Λ-Rh and Δ-Rh complexes bound to short DNA oligonucleotides containing their respective recognition sites. This NMR study elucidates the key elements of recognition of each enantiomer with its site, and suggests structural

models with which to consider previous photocleavage results.

EXPERIMENTAL PROCEDURES

Synthesis of Materials. Λ- and Δ-1-Rh(MGP)$_2$phi^{5+} complexes were made, the isomers separated, and enantiomers resolved with baseline resolution by published methods (*7*, *30*). The single strands L1 (5′-GACATATGTC-3′), D1 (5′-CGCATCTGAC-3′), and D2 (5′-GTCAGATGCG-3′) were each synthesized on the 10 μmol scale on an Applied Biosystems DNA/RNA synthesizer, using standard phosphoramidite chemistry, and retaining the final trityl protecting group (*15*). Each strand was purified by HPLC (C_{18} reversed phase column, 50 mM triethylammonium acetate/CH_3CN eluant) and the trityl protecting group removed (5 M acetic acid) and purified a second time by HPLC. The strands were desalted (Millipore C_{18} Sep-pak), and the cation from the HPLC buffer was exchanged to Na^+ (Sephadex SP-C25 resin, prepared with NaCl). The two non-self-complementary strands (D1 and D2) were dissolved in D_2O and titrated together. The titration was followed by ^{1}H NMR, particularly noting the disappearance of the single strand TMe group signals. The single strands D1Z (5′-CGCATCT**Z**AC-3′) and D2Z (5′-GTCAGAT**Z**CG-3′) containing one N^7-deazaguanine base (**Z**; Glen Research) were similarly synthesized and purified, using the oxidant (1*R*)-(−)-(10-camphorsulfonyl-oxaziridine), as described in Glen Research specifications. Each was titrated with their respective complement to give the duplexes D1Z−D2 and D1−D2Z.

Sample Preparation. Samples for two-dimensional (2D) NMR experiments were prepared in "100%" D_2O (Cambridge Isotope Laboratories) or 90:10 H_2O/D_2O with 10 mM sodium phosphate buffer, in concentrations from 0.8 to 1.1 mM (1:1 metal:DNA ratios). The titrations of Δ-1-Rh(MGP)$_2$phi^{5+} into D1−D2 and of Λ-1-Rh(MGP)$_2$phi^{5+} into L1 were each followed by ^{1}H NMR at 10 °C. Prior to titration, DNA and metal complex solutions were quantitated by UV−visible spectroscopy: Δ- and Λ-Rh, $\epsilon_{360} = 19.4$ mM^{-1} cm^{-1} [based on the parent complex Rh(phen)$_2$phi^{3+}] (*16*); D1−D2 duplex, $\epsilon_{260} = 132.0$ mM^{-1} cm^{-1}; and L1 duplex, $\epsilon_{260} = 134.4$ mM^{-1} cm^{-1}. The Δ-Rh/D1Z−D2 and Δ-Rh/D1−D2Z samples were prepared at 0.45 mM Δ-Rh and 0.5 mM DNA concentrations in a similar manner.

Instrumentation and Experiments. All 2D spectra were recorded on a Varian 600 MHz Unity Plus System, and processed with Felix version 2.3 (BIOSYM/Molecular Simulations) software. 2D-NOESY spectra were collected at 10 °C with a t_{mix} of 300 ms, and water suppression by presaturation during relaxation delay and mixing. Additionally, a 2D-NOESY spectrum of Λ-Rh/L1 (0.8:1) was taken on a 500 MHz Brüker AMX500 spectrometer at 25 °C, and this is the spectrum shown in Figure 3. 2D-NOESY spectra in D_2O of 1:1 Δ-Rh/D1−D2 in D_2O with a t_{mix} of 100 ms, overtitrated 1.5:1 Δ-Rh/D1−D2, and undertitrated 2:3 Λ-Rh/L1 were also recorded. The 2D-NOESY spectra of D1−D2Z and D1Z−D2 were recorded and assigned in D_2O at 15 °C prior to the addition of Δ-1-Rh(MGP)$_2$phi^{5+}. Δ-Rh/D1Z−D2 and Δ-Rh/D1−D2Z 2D-NOESY spectra were taken at 15 °C with a t_{mix} of 200 ms, with water suppression by presaturation.

Structural models were generated using InsightII version 2.3.0 software (BIOSYM/Molecular Simulations). B-form DNA was constructed with the Biopolymer module of the software, with a noncanonical 5′-N_5pN_6-3′ intercalation step (double rise, 20° unwinding), and energy minimized (100 steps, steepest decent). Λ- and Δ-Rh complex coordinates

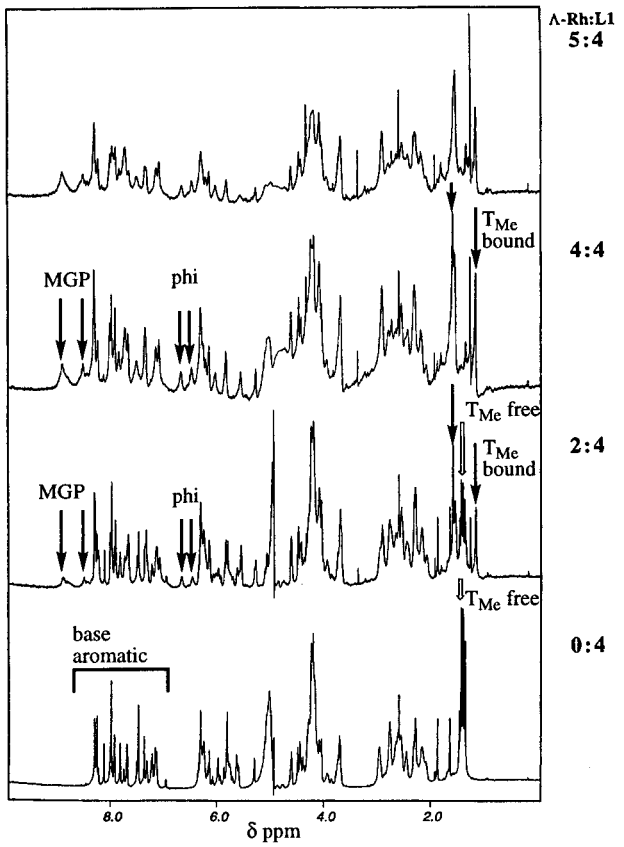

FIGURE 2: ^{1}H NMR titration of Λ-Rh into L1. Representative free (white arrows) and bound (black arrows) peaks are noted, indicating the increasing intensity of peaks due to Λ-Rh ligand peaks and metal-bound DNA species (Λ-Rh/L1), with concomitant loss of free L1 duplex resonances. Overtitration does not further change the chemical shifts of the bound DNA duplex resonances.

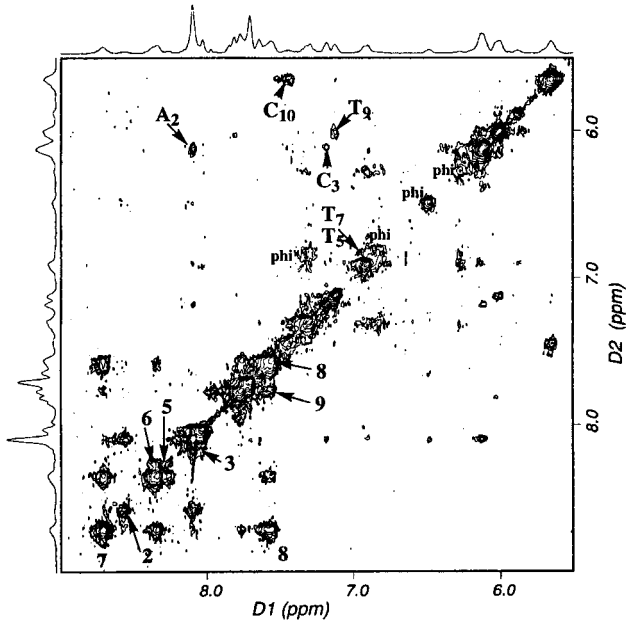

FIGURE 3: Expanded aromatic region of the Λ-Rh/L1 2D-NOESY spectrum in D_2O (0.8:1 Λ-Rh/L1 duplex). The spectrum was collected at 25 °C, with a t_{mix} of 300 ms.

Table 1: Summary of DNA Chemical Shift Data (ppm) for Aromatic Base, Thymine Methyl, and Guanine/Thymine Imino Protons in the Absence and Presence of Λ-1-Rh(MGP)$_2$phi^{5+} [a]

base H6/H8	free	bound	Δδ(ppm)
G1	7.87	7.87	0
A2	8.24	8.25	0.01
C3	7.25	7.31	0.06
A4	8.19	8.25	0.06
T5	7.15	7.04	−0.11
A6	8.19	8.19	0
T7	7.09	7.10	−0.01
G8	7.75	7.96	0.21
T9	7.30	7.27	−0.03
C10	7.63	7.61	−0.02
G/T imine	**free**	**bound**	**Δδ(ppm)**
G1·C10			
A2·T9	13.92		
C3·G8	12.36	12.63	0.27
A4·T7	13.41	12.83	−0.58
T5·A6	13.36	12.31	−1.05
TMe	**free**	**bound**	**Δδ(ppm)**
T5	1.37	1.56	0.19
T7	1.30	1.19	−0.11
T9	1.34	1.55	0.21

[a] Shifts of (5′-GACATATGTC-3′)$_2$.

were based on the parent complex crystal structure (*16*) and the complexes manually docked into the DNA, avoiding van der Waals radius conflicts.

RESULTS

Interactions of Λ-1-Rh(MGP)$_2$phi^{5+} with (L1)$_2$

Characterization of the DNA Oligomer. The decamer duplex L1 (5′-$G_1A_2C_3A_4T_5A_6T_7G_8T_9C_{10}$-3′) which contains the palindromic 6 bp recognition site for Λ-Rh was first characterized. The 2D-NOESY spectra of L1 were fully assigned in both D_2O and 90:10 H_2O/D_2O (Tables 1 and 2, selected peaks). Due to the C_2 symmetry of the duplex, only one peak is observed for each proton of the 10 nucleotides.

Conformation in Solution after Titration of Λ-Rh into L1. The titration of Λ-Rh into L1 was followed by ^{1}H NMR in D_2O (Figure 2). The binding of the complex is slow on the NMR time scale under these conditions, as evidenced by the smooth disappearance of peaks such as the TMe (indicated), with the concomitant appearance of bound DNA peaks. Both the DNA peaks and the metal complex peaks which grow in (indicated) are sharp and not exchange broadened. As in the spectrum of L1 alone, the addition of an equimolar amount of Λ-1-Rh(MGP)$_2$phi^{5+} to the duplex again results in only one set of peaks per nucleotide. Since the binding is on a slow time scale, this indicates the association of the complex does not disrupt the symmetry of the DNA, and there is only one bound conformation. Peaks between 6 and 7 ppm, upfield from their uncomplexed shifts, are assigned to the intercalated phi ligand. Since intercalation typically results in upfield ^{1}H shifts due to the shielding of the stacked bases' ring current (*17*), this indicates that the metallointercalator is well-stacked with the DNA.

Peak Assignments Using 2D-NMR of Λ-Rh/L1. The solution behavior of Λ-Rh/L1 was investigated using 2D-NMR in D_2O and 90:10 H_2O/D_2O, and compared to that of the free oligomer. As observed in the one-dimensional (1D) spectra, the 2D-NMR data show the symmetry of the duplex is retained with the bound complex (Figure 3). Nearly all peaks in the 1:1 Λ-Rh/L1 NOESY spectrum in D_2O could be assigned. Although the spectrum is well-behaved overall, the inherent flexibility of the system prevents some NOE contacts from being observed, notably T7, for which only the H2′, H2″, and aromatic peaks were observed. At the same step, A4 H2 shifts significantly downfield [Δδ(ppm) = 0.70] from its free position. Additionally, the chemical shift differences between H2′ and H2″ peaks are smaller in the bound form for both A6 and T5 nucleotides [T5 free, Δδ(ppm) = 0.37; T5 bound, Δδ(ppm) = 0.10; A6 free, Δδ(ppm) = 0.33; and A6 bound, Δδ(ppm) = 0.09]. Unlike free L1 duplex, several protons of the terminal bases are doubled in the bound spectrum, indicative of fraying at the duplex ends. In the 90:10 H_2O/D_2O NOESY spectrum, neither the labile protons of the guanidinium groups nor the phi NH protons could be observed, presumably due to solvent accessibility.

Intercalation Site. The site of intercalation was determined from the shifts of the base imino protons relative to the free DNA positions. Chemical shift changes of single-strand DNA imino and other base protons upon the formation of a duplex are influenced by hydrogen-bonding (downfield shifts) and base ring-current magnetic anisotropy effects (upfield shifts) (*17*). Since a ligand ring system inserting itself between base pairs typically has a large ring current, by observing the upfield shift of imino peaks with intercalation, we can determine the binding sites of the metallointercalators (*18*).

Table 2: Sugar Proton Chemical Shifts and Shift Perturbations (ppm) of (5′-GACATATGTC-3′)$_2$ in the Absence and Presence of Λ-1-Rh(MGP)$_2$phi^{5+}

	H1′			H2′			H2″		
residue	δ(free) (ppm)	δ(bound) (ppm)	Δδ (ppm)	δ(free) (ppm)	δ(bound) (ppm)	Δδ (ppm)	δ(free) (ppm)	δ(bound) (ppm)	Δδ (ppm)
G1	5.56			2.54	2.51	−0.03	2.72	2.67	−0.05
A2	6.23	6.22	−0.01	2.72	2.76	0.04	2.92	2.88	−0.04
C3	5.53			2.06	2.06	0.0	2.40	2.40	0.0
A4	6.19	6.25	0.06	2.60	2.59	−0.01	2.93	2.89	−0.04
T5	5.70	5.98	0.28	2.11	2.14	0.03	2.48	2.24	−0.24
A6	6.13			2.58	2.42	−0.16	2.91	2.51	−0.40
T7	5.74			2.08	2.18	0.10	2.41	2.44	0.03
G8	5.92	6.16	0.24	2.58	2.63	0.05	2.73	2.84	0.11
T9	6.08	6.11	0.03	2.13	2.13	0.0	2.52	2.52	0.0
C10	6.26	6.26	0.0	2.23	2.27	0.04	2.23	2.27	0.04

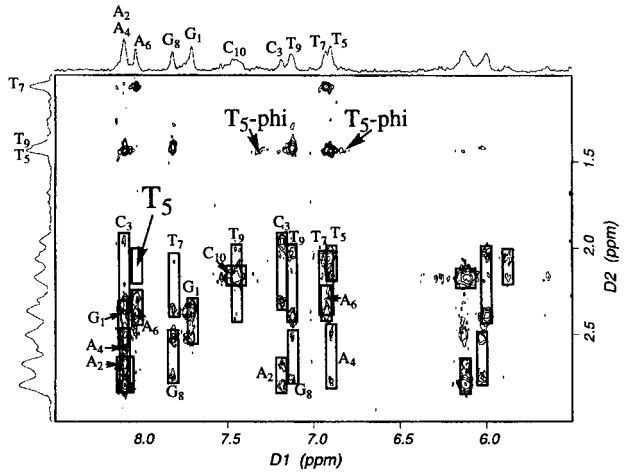

FIGURE 4: Expanded H2′ and H2″ aromatic region of the Λ-Rh/L1 NOESY spectrum in D_2O. Cross-peaks between T5Me and phi peaks, and the missing cross-peaks between T5 H2′ and H2″ and A6 H8, are indicated in a larger font.

In keeping with the photocleavage results, in which comparable amounts of cleavage are observed at T5 and A6 in the center of the site (*7*, *30*), we find the 2D-NMR results indicate the complex intercalates symmetrically and exclusively between these two bases. The NMR spectrum exhibits a notable lack of complexity, as the centrosymmetry is retained upon complex binding (Figure 3). With the addition of Λ-Rh to L1, the T5 imino peak moves dramatically upfield [$\Delta\delta$(ppm) = −1.05] and is somewhat broadened (Figure S1 of the Supporting Information). This shift indicates that the symmetric association of the metal complex at the center of the duplex is intercalative, with the phi ligand lying between T5 and A6. Additionally, the phi ligand peaks shift upfield upon stacking between base pairs, and phi(4,5), phi(3,6), and phi(2,7) exhibit NOE cross-peaks to T5Me. The phi(2,7) peak, which is furthest upfield at 6.39 ppm, also has a cross-peak to T5 H2′ and H2″. The smaller upfield shift of the T7 imino proton [$\Delta\delta$(ppm) = −0.62] is consistent with the attenuated effect of intercalation one base pair away. The notable absence of the cross-peaks between T5 H2′ and H2″ and A6 H8 (Figure 4), disrupting the NOE walk along the duplex (*19*), gives further credence to this base step being the intercalation site.

Orientation of the Guanidinium Arms. The importance of the guanidinium functionalities as recognition elements was also examined. Two peaks due to the methylene protons of the guanidinium arms occur at 4.92 and 5.44 ppm, and are related by a cross-peak in the 2D-NOESY spectrum. Although each methylene proton is related through C_2 symmetry to a proton of the other MGP methylene arm, the geminal protons are magnetically inequivalent once they are bound to the complex. These protons have intramolecular NOE peaks to MGP H5 and MGP H6 in both the D_2O and 90:10 H_2O/D_2O spectra. In the latter case, there are weak intermolecular cross-peaks from the 4.92 ppm peak to T7 H6, T7Me, and T5Me. There are no intermolecular cross-peaks that can be observed in the D_2O spectrum. Cross-peaks to sugar protons were not observed. The few DNA−metal cross-peaks serve to establish the proximity of the ancillary arm to the purported G8 H-bonding site, though the predominant observation is the lability of the guanidinium arm.

The greatest perturbation in base aromatic peaks (H6 and H8) occurs not at the intercalation site, but at the presumed site of contact with the guanidinium arms (G8) (Table 1). This shift, consistent with hydrogen bonding into the ring system, is mirrored in the significant downfield shift of the hydrogen-bonded G8 imino proton at the center of the C3·G8 base pair. Upon complex binding, the G8 imino proton moves downfield 0.27 ppm to 12.63 ppm.

The labile guanidinium protons were not observed in the 90:10 H_2O/D_2O Λ-Rh/L1 NOESY spectrum. The hydrogen-bonding interaction between the guanidinium group and G8 is unlikely to be rigid, and may even be mediated by water molecules. Thus, the rapid equilibrium with solvent and motion at the H-bonding site are expected to wash out these peaks. It is not uncommon, even with larger DNA-binding motifs such as Zn-finger proteins, for the exchanging guanidinium proton resonances of arginine residues to be missing in the NMR spectra (*20*, *21*). Similarly, the guanidinium protons are also not observed in the 90:10 H_2O/D_2O spectrum of the free Λ-Rh metal complex.

Groove Approach. The NMR evidence corroborates an earlier systematic study of the binding site preference of Λ-Rh which confirmed the importance of major groove contacts between the complex and G8 (*7*, *30*). The guanidinium methylene cross-peaks to T5Me and T7Me offer direct spectral evidence for a major groove approach. However, there are a few coincidences of shift (at both 10 and 25 °C) which make several other key peaks, such as MGP H2 and MGP H3 on the DNA-binding face of the complex, uninformative in determining groove approach. MGP H3 has a shift equivalent to both A2 H8 and A4 H8. Even at the intercalation site, the shift of A6 H2 is identical to phi(1,8), which requires that the phi(1,8) to MGP H2 and MGP H3 NOE cross-peaks not be mistaken for minor groove contacts between A6 H2 and the inner face of the metal complex. Also, both C3 amine proton shifts are coincident with metal complex peaks [C3NH with phi(2,7) and C3NH with MGP H9]. These peak assignments are internally consistent and have been confirmed by studying exchange cross-peaks between free and bound DNA with a 2:3 Λ-Rh/L1 sample.

NOE cross-peaks between backbone sugar protons and ancillary ligand protons on the DNA binding face of the complex are not necessarily expected for a major groove approach. None were observed in the major groove binding of either Δ-Ru(phen)$_2$dppz^{2+} or Rh(phen)$_2$phi^{3+} to the hexamer duplex (5′-GTCGAC-3′) (*22*, *23*). In those cases, NOE cross-peaks between the ancillary phen ligand H2 and H3 were observed to protons of the bases adjacent to the intercalation site, but no sugar contacts were evident. Conversely, contacts by the face of the minor groove binder Δ-[Ru(phen)$_2$DPQ]$^{2+}$ (DPQ = dipyrido[3,2-*d*:2′,3′-*f*]-quinoxaline) (*24*) to the sugar protons of the much narrower minor groove were observed.

Geometry of the Recognition Site. There are several spectral changes upon Λ-Rh complexation which support the conclusion that the binding site is significantly unwound. The decreased shift difference between H2′ and H2″ for the central two bases suggests these geminal protons are in an environment less asymmetric than that in native B-form DNA. This may reflect a large unwinding at the intercalation site. Additionally, the large downfield shift of A4 H2

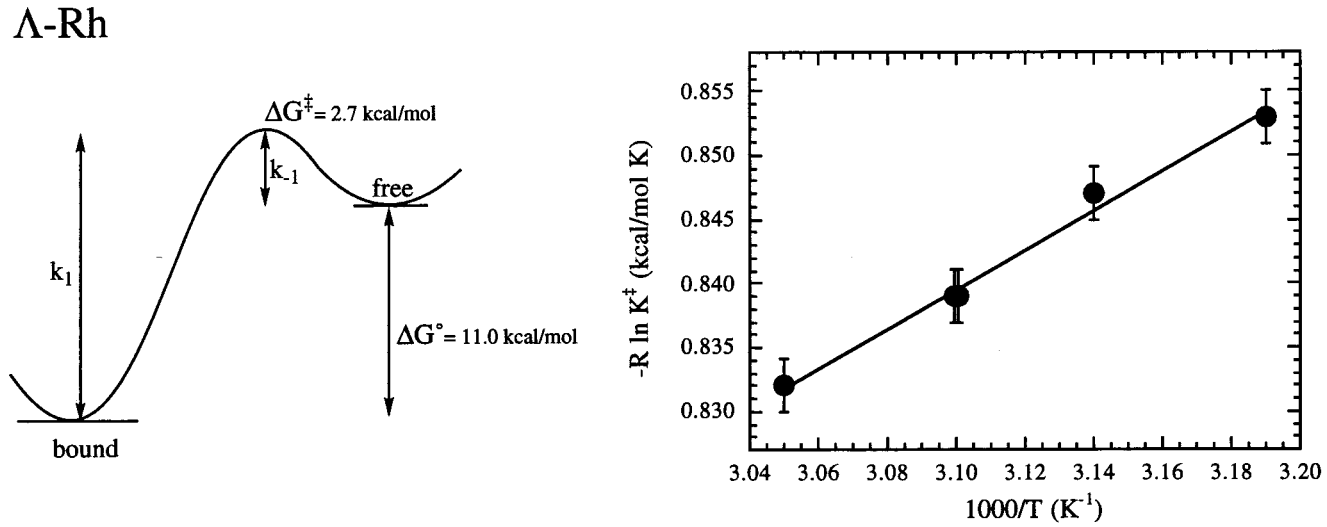

FIGURE 5: Reaction coordinate and Eyring plot of exchange data from a variable-temperature study of Λ-Rh/L1 (2:3 ratio). $\Delta G^{\ddagger}_{298}$ was calculated from the Eyring plot, and $\Delta G°$ was calculated from the previously determined binding constant ($K \approx 10^8$ M^{-1}) (*7*, *30*). The fit shown is the line $y = 0.360 + (0.155)x$, with an R of 0.995.

suggests significant unwinding, which would cause this proton to move out of the base stack.

NOE contacts between T5Me and phi(1,8) were not observed, consistent with these phi peaks experiencing greater solvent accessibility. A more ladder-like base stack, and thus an extremely wide major groove, would not shield the complex from solvent as effectively as a narrower groove with tight shape complementarity. Also, the difficulty in distinguishing the broadened peaks of T7 seems to suggest that T7 has particular flexibility, as the recognition of twistability at this site would predict (*7*, *30*). The fraying at the end of the duplex is also consistent with flexibility.

Exchange Kinetics. The kinetics of binding were investigated by following the coalescence of free and bound DNA resonances with temperature in a partially titrated Λ-Rh/L1 (2:3) sample. In this system, these pairs of peaks exhibit simple two-site exchange, so the rate constant at coalescence, k_{ex}, for the forward (and reverse) complexation reaction can be calculated (*25*) from the difference in chemical shift at coalescence ($\Delta\delta_{Hz}$) and the coalescence temperature (T_c): $k_{ex} = (\pi\Delta\delta_{Hz})/\sqrt{2}$.[1]

From k_{ex}, activation parameters for this exchange can be determined from an Eyring plot of $-R \ln K^{\ddagger}$ versus $1/T$ (Figure 5). The equilibrium rate constant is calculated with the equation $K^{\ddagger} = (k_{ex}h)/(k_bT)$, where k_b is Boltzmann's constant, h is Plank's constant, and R is the ideal gas constant. Although there are inherent difficulties in the measurement of the coalescence temperature and $\Delta\delta_{Hz}$ among the many peaks in this complicated system, five well-separated pairs of peaks were identified in the aromatic and methyl regions. The error in $-R \ln K^{\ddagger}$ due to coalescence and $\Delta\delta_{Hz}$ measurements is estimated to be ± 0.002 kcal mol^{-1} K^{-1}.

The association rate (k_{ex}) at 298 K was found to be 68 s^{-1}, corresponding to an activation energy ($\Delta G^{\ddagger}_{298}$) of 2.7 kcal/mol. The free energy ($\Delta G°$) for intercalation into this site can be calculated from the binding constant to be 11.0 (± 0.3) kcal/mol (*7*, *30*), but the activation barrier to binding is approximately the same as the stacking energy of one base pair ($\approx$2.4 kcal mol^{-1} $base^{-1}$) (*26*). This relatively small barrier appears to be insufficient to promote the unwinding and significant rearrangement of the site, including intercalation. Thus, the activation barrier for intercalation appears to be due almost entirely to unstacking the DNA helix. Very little energetic price is paid for unwinding this site by 70°, consistent with the inherent flexibility of AT-rich sequences. The complex can, therefore, be considered to trap the sequence in its unwound state for the relatively small cost of disrupting the base stack to insert the phi ligand.

Interactions of Δ-1-Rh(MGP)$_2$phi^{5+} with D1−D2

Characterization of the DNA Oligomer. The decamer duplex D1−D2 5′-$C_1G_2C_3A_4T_5C_6T_7G_8A_9C_{10}$-3′−5′-$G_{11}T_{12}$-$C_{13}A_{14}G_{15}A_{16}T_{17}G_{18}C_{19}G_{20}$-3′ containing the asymmetric 6 bp recognition site for Δ-Rh in the center of the decamer was initially characterized. The single strands were titrated together to give the 1:1 duplex, following the titration by ^{1}H NMR. The 2D-NOESY spectra of D1−D2 were fully assigned in both D_2O and 90:10 H_2O/D_2O (Tables 3−5). Unlike the palindromic L1 duplex, D1−D2 lacks C_2 symmetry, and thus has unique shifts for each proton of the 20 nucleotides.

Conformation in Solution after Titration of Δ-Rh into D1−D2. The titration of Δ-Rh into the non-self-complementary duplex D1−D2 shows a binding pattern dramatically different than that seen for the Λ-enantiomer bound to its site (Figure 6). An equimolar amount of Δ-Rh was titrated into the D1−D2 duplex. In the presence of Δ-Rh, each of the 20 nucleotides has at least two distinct chemical environments, which are in slow exchange on the NMR time scale. These two conformers have approximately equal intensity. As the metal complex is added, broad phi peaks grow between 6 and 7 ppm. These peaks, shifted upfield from free metal phi peaks, indicate the ligand is well-intercalated into the base stack.

Peak Assignments Using 2D-NMR of Δ-Rh/D1−D2. Both the D_2O and 90:10 H_2O/D_2O NOESY spectra were exten-

[1] Note that at coalescence, $\Delta\delta_{Hz}$ is defined as 0 Hz. However, an estimate of the true individual peak positions, taking into consideration changes with temperature, allows us to calculate the coalescence equilibrium rate, k_{ex}.

Table 3: Base Aromatic Proton Chemical Shifts and Shift Perturbations (parts per million) of 5′-CGCATCTGAC-3′ and 5′-GTCAGATGCG-3′ in the Absence and Presence of Δ-1-Rh(MGP)$_2$phi^{5+} [a]

	base H6/H8							A H2 and C H5				
	δ(D1−D2)	δ(D1Z−D2)	δ(D1−D2Z)	δ(bound)		Δδ(D1−D2-bound)		δ(free)	δ(bound)		Δδ(D1−D2-bound)	
residue	(ppm)	(ppm)	(ppm)	A	B	A	B	(ppm)	A	B	A	B
C1	7.65	7.73	7.73	7.65	7.65	0.0	0.0	5.92	5.78	5.91	−0.14	−0.01
G2	7.97	7.99	7.99	7.97	7.97	0.0	0.0	−	−	−	−	−
C3	7.41	7.41	7.41	7.36	7.41	−0.05	0.0	5.45	5.47	5.47	0.02	0.02
A4	8.35	8.33	8.31	8.27	8.35	−0.08	0.0	7.64	7.64	7.89	0.0	0.25
T5	7.19	7.19	7.17	7.12	7.20	−0.07	0.01	−	−	−	−	−
C6	7.56	7.55	7.55	7.48	7.58	−0.08	0.00	5.52	5.54	5.77	0.02	0.25
T7	7.29	7.29	7.28	7.31	7.56	0.02	0.27	−	−	−	−	−
G8	7.94	**6.87**	7.90	7.86	7.96	−0.08	0.02	−	−	−	−	−
A9	8.18	8.18	8.19	8.19	8.24	0.01	0.06	7.94	7.95	7.95	0.01	0.01
G11	8.00	7.92	7.93	8.00	8.00	0.0	0.0	−	−	−	−	−
T12	7.50	7.49	7.49	7.50	7.54	0.0	0.04	−	−	−	−	−
C13	7.54	7.54	7.55	7.51	7.58	−0.03	0.04	5.69	5.69	5.69	0.0	0.0
A14	8.21	8.18	8.19	8.05	8.22	−0.16	0.01	7.48	7.48	7.76	0.0	0.28
G15	7.71	7.71	7.72	7.71	7.88	0.0	0.17	−	−	−	−	−
A16	8.09	8.08	8.07	8.11	8.39	0.02	0.30	7.70	7.38	7.71	−0.32	0.01
T17	7.07	7.06	7.08	7.08	7.26	0.01	0.19	−	−	−	−	−
G18	7.85	7.83	**6.87**	7.72	7.87	−0.13	0.02	−	−	−	−	−
C19	7.35	7.36	7.34	7.36	7.41	0.01	0.06	5.39	5.42	5.52	0.03	0.13
G20	7.97	7.90	7.90	7.97	7.97	0.0	0.0	−	−	−	−	−

[a] Base aromatic H6 and H8 proton shifts are reported for the standard D1−D2 duplex and for the N^7-deazaguanine-modified duplexes D1Z−D2 and D1−D2Z. The H8 chemical shifts of the modified bases (G8 and G18) are bold. The reported bound chemical shifts are for Δ-Rh/Dl, and chemical shift differences given are between the free and bound D1−D2 duplex. The two conformational isomers (Δ-Rh/DNA-bound forms A and B) cannot be distinguished from one another. They are here arbitrarily reported in increasing chemical shift order, with no further assignments implied. Three bases (A14, A16, and T17) have an additional minor component peak in the aromatic region, which is not reported.

Table 4: Summary of DNA Chemical Shift Data (ppm) for the Guanine/Thymine Imino Protons[a]

	imine	A		B		C	
	δ(free)	δ(bound)	Δδ	δ(bound)	Δδ	δ(bound)	Δδ
C1·G20		13.15					
G2·C19	13.07	13.07	0.0				
C3·G18	12.63	12.63	0.0	12.82	0.19		
A4·T17	13.56	13.31	−0.25	13.42	−0.14	13.56	0.0
T5·A16	13.56	12.74	−0.82	13.04	−0.52	13.56	0.0
C6·G15	12.54	11.99	−0.55	12.54	0.0	12.67	0.13
T7·A14	13.83	13.83	0.0	13.96	0.13	14.09	0.26
G8·C13	12.68	12.68	0.0	12.92	0.24		
A9·T12	13.92	13.92	0.0	13.97	0.05		
C10·G11		12.92					

[a] Shifts of 5′-CGCATCTGAC-3′/5′-GTCAGATGCG-3′ in the absence and presence of Δ-1-Rh(MGP)$_2$phi^{5+}. For the central base steps, a third peak (C) is observed in the imino region only. As one of the peaks consistently corresponds to the free DNA position, this may reflect a population of partially dissociated Δ-Rh/DNA, in which the intercalator has moved away from the imino proton in the center of the base pair.

sively assigned, utilizing exchange peaks between conformationally related pairs, as well as NOE peaks. A ROESY spectrum aided in distinguishing the exchange and NOE cross-peaks (Figure S2 of the Supporting Information). The differences in shifts between the two conformers are most obvious in the base aromatic region (Figure 7); for example, δ(ppm) for A16 H2 = 7.38 and 7.71, and δ(ppm) for A16 H8 = 8.11 and 8.39 (Table 3). Interestingly, in all cases, one of the two exchange-related peaks has a shift close to or the same as the free proton shift. An overtitrated sample (1.5:1 Δ-Rh/D1−D2) also exhibited both sets of peaks in exchange, showing the two conformations to be both metal-associated forms, and not excess free DNA. This conclusion is corroborated by considering the constant of binding of this complex to DNA (*K*), which has been determined under similar conditions to be 10^7 M^{-1} (*7*, *30*). As the spectra were recorded at millimolar concentrations, the complex must be fully bound under these conditions.

The phi NH protons were assigned in the H_2O Δ-Rh/D1−D2 NOESY spectrum as well. Four sharp peaks occur between 9 and 12 ppm, presumably from the two phi NH protons of each different conformation, though this shift is dramatically upfield of that expected from analogies with other phi-containing metallointercalators (*4*−*6*, *22*, *23*). Although they are in exchange at the same rate as the rest of the complex, exchange peaks are not observed between peaks so far apart in chemical shift, as coalescence is directly related to both rate and Hz separation. As a consequence, pairs of phi NH peaks exchanging between conformations cannot be assigned. Two of the phi NH peaks, however, have additional cross-peaks. One [δ(ppm) = 12.15] has an intermolecular proximity to A16 H2, whereas the peak at 10.80 ppm has intramolecular NOEs to both a nonlabile phi peak and MGP H2.

Unlike the Λ-Rh/L1 system, the spectrum of Δ-Rh/D1−D2 is well-defined for resonances at the ends of the duplex, but more difficult to assign for resonances in the middle. However, there are distinct differences in the environments of the terminal bases between the two conformers. This is particularly obvious in the shifts of C1, G2, and C3 H2′ and H2″, and in the C1 H5 shift (Tables 3 and 5).

Intercalation Site. Δ-Rh, like its enantiomer, binds to DNA by intercalation in the center of the site. Though the complexity and overlap of the region prevent distinguishing a missing step in the NOE walk through the aromatic−sugar H2′ and H2″ peaks, NOE cross-peaks between the upfield shifted phi protons and T5Me, C6 H5, and C6 H1′ define the intercalation site. Changes in shifts of the CH5−H6 cross-peaks and the labile CNH_2 cross-peaks can be described as minimal except for those of C6 protons, at the intercalation

Table 5: Sugar Proton Chemical Shifts and Shift Perturbations (ppm) of 5′-CGCATCTGAC-3′/5′-GTCAGATGCG-3′ in the Absence and Presence of Δ-1-Rh(MGP)$_2$phi^{5+}[a]

	H1′			H2′			H2″		
residue	δ(free) (ppm)	δ(bound) (ppm)	Δδ (ppm)	δ(free) (ppm)	δ(bound) (ppm)	Δδ (ppm)	δ(free) (ppm)	δ(bound) (ppm)	Δδ (ppm)
C1	5.75	5.78	0.03	2.00	2.01	0.01	2.00	2.42	0.42
G2	5.92			2.62	2.68	0.06	2.62	2.74	0.12
C3	5.64			2.12	1.95	−0.17	2.48	2.16	−0.32
A4	6.29	6.11, 6.30	−0.18, 0.01	2.68	2.70	0.02	2.98	2.99	0.01
T5	5.89			2.10	2.00	−0.10	2.48	2.10	−0.38
C6	5.98	6.07, 6.26	0.09, 0.28	2.05	2.07	0.02	2.47	2.49	0.02
T7	5.78			2.01	2.03	0.02	2.36	2.38	0.02
G8	5.55			2.71	2.77	0.06	2.75	2.89	0.14
A9	6.25	6.26	0.01	2.62	2.65	0.03	2.87	2.91	0.04
C10	6.02	6.01	−0.01	2.05	2.05	0.0	2.05	2.07	0.02
G11	6.20	6.03	−0.17	2.71	2.74	0.03	2.80	2.81	0.01
T12	6.20	6.18	−0.02	2.19	2.21	0.02	2.55	2.55	0.0
C13	5.46			2.04	2.03, 2.22	−0.01, 0.18	2.33	2.37, 2.55	0.04, 0.22
A14	5.97	5.82, 5.96	−0.01, −0.15	2.72	2.57	−0.15	2.90	2.90	0.0
G15	5.58			2.56	2.59, 2.35	0.03, −0.21	2.71	2.66, 2.35	−0.05, −0.36
A16	6.18			2.55	2.57	0.02	2.90	2.92	0.020
T17	5.76			1.97	1.99, 2.14	0.02, 0.17	2.39	2.37, 2.14	−0.02, −0.25
G18	5.81			2.57			2.64		
C19	5.74	5.78	0.04	1.91	1.96	0.05	2.33	2.36	0.03
G20	6.16	6.16	0.0	2.33	2.36	0.03	2.62	2.64	0.02

[a] The two conformational isomers (Δ-Rh/DNA-bound form) cannot be distinguished from one another. Shifts for the two conformers are thus reported together, with no further assignments implied. In some instances, H2′ and H2″ for a given nucleotide could not be distinguished because of the ambiguity of the two conformational iomers.

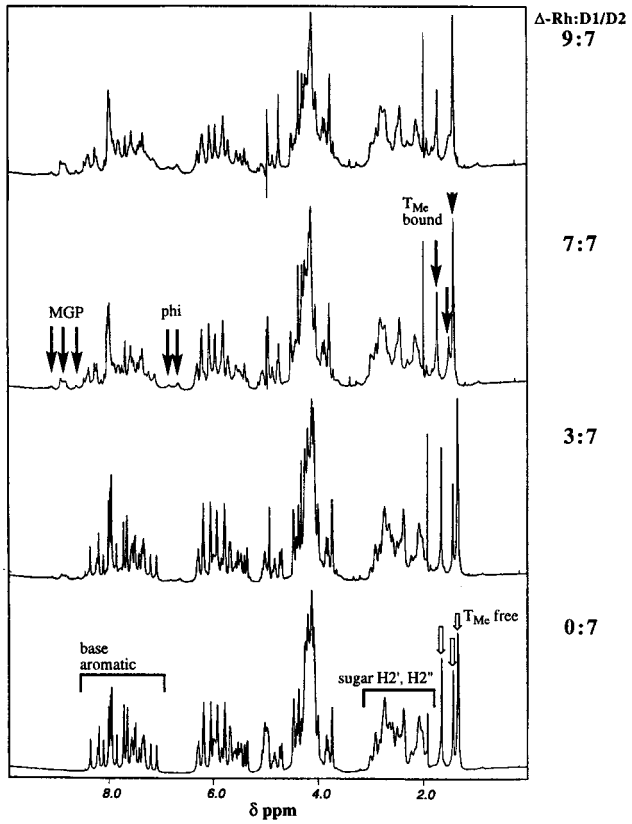

FIGURE 6: ^{1}H NMR titration of Δ-Rh into D1−D2. Representative free (white arrows) and bound (black arrows) peaks are noted. As the Δ-Rh:D1−D2 ratio increases, the intensity of peaks due to Δ-Rh ligand peaks and metal-bound DNA species (Δ-Rh:D1−D2) increases, with concomitant loss of free D1−D2 duplex resonances. Overtitration does not further change the chemical shifts of the bound DNA duplex resonances.

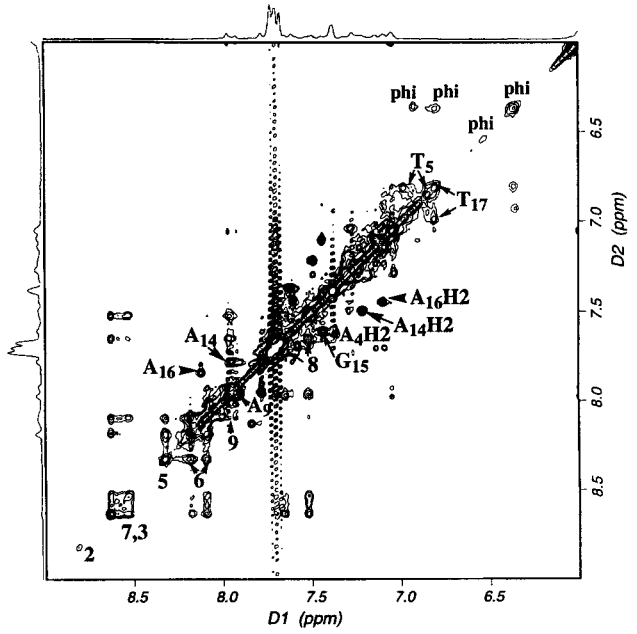

FIGURE 7: Expanded aromatic region of the Δ-Rh/D1−D2 2D-NOESY spectrum in D_2O (1:1 Δ-Rh/D1−D2 duplex). The spectrum was collected at 10 °C, with a t_{mix} of 300 ms. Exchange peaks between resonances of the two bound DNA forms are indicated (base number), as are sets of Δ-Rh phi ligand peaks (phi) and MGP ligand peaks (numbers).

site, which are badly exchange broadened. Additionally, the most notable changes in the complicated imino region occur for G15 and T5, which are consistent in direction and magnitude with adjacent intercalation (Table 5; see also Figure S3 of the Supporting Information). Only these two bases experience significant upfield shifts.

Groove Approach. The unusual shifts and flexibility of protons near the center of the D1−D2 duplex prevented assignments of many specific major groove interactions. There were very few cross-peaks between ancillary MGP ligands and the duplex, generally. As noted above, cross-

peaks to sugar protons of the major groove are uncommon for this type of complex (*22*, *23*). The NOE contacts which were observed included those between metal complex phi and duplex A16 H2, which lies in the minor groove. However, these NOEs are also consistent with severely canted intercalation into the major groove. A computer-generated model suggests a canted orientation for intercalation from the major groove would bring this minor groove A16 H2 proton within 5 Å of the phi NH. Apparent weak NOE cross-peaks between MGP H2 and MGP H9 protons and one (but not the other) of the A16 H2 chemical shifts [δ(ppm) = 7.71] are also observed. However, these MGP contacts are in fact with G15 H8 protons of the major groove [contacts to both conformer shifts are observed; G15 H8 δ(ppm) = 7.71 and 7.88]. The peaks of the intercalating ligand have NOE contacts to the sugars only at C6 H1′, rather than to the 5′ H2′ and H2″ protons as more centrally oriented intercalators, including Λ-Rh, exhibit. The conclusion that the complex binds from the major groove was confirmed quite clearly by the significant effect on specificity and kinetic exchange of removing a single major groove contact, namely, the N7 atom of either G8 or G18 (vide infra).

Interactions with a Modified Site for Δ-Rh/D1Z−D2 and Δ-Rh/D1−D2Z. To test the importance of the hydrogen bonding interactions between Δ-Rh guanidinium groups and the major groove side of 3′-G to either end of the recognition site, modified duplexes were synthesized in which either G8 or G18 was replaced with a N^7-deazaguanine base.

The 2D-NOESY spectra of the modified duplexes alone were initially assigned. With the exception of the aromatic protons of the deaza-G bases, there were only minor differences in shift, and therefore conformation, between these duplexes and D1−D2 (Table 3). The deaza-G H8 protons, adjacent to the missing N7 atom, were shifted significantly upfield, as expected [D1Z−D2, G8 δ(ppm) = 6.87; D1−D2Z, G18 δ(ppm) = 6.87].

Given that two conformers were observed in slow exchange in the Δ-Rh/D1−D2 spectra, it was anticipated that removing one H-bonding contact might drive the equilibrium between conformers to give unique bound states (A and B). The Δ-Rh/deaza-DNA spectra were recorded under conditions similar to those of the Δ-Rh/D1−D2 spectra, and at concentrations where all metal complex is expected to be bound, at the very least, nonspecifically. In fact, unlike the original spectrum, neither Δ-Rh/D1Z−D2 nor Δ-Rh/D1−D2Z spectra were in the NMR slow exchange regime.

NOESY cross-peaks were extremely exchange broadened, to the extent that most peaks were lost in noise. The resonance of T5Me, adjacent to the binding site, was indistinguishable in both data sets. A few peaks may be considered consistent with a preference for one conformer, or binding mode, over the other, but specific assignments would represent an overinterpretation of the data. The predominant effect of adding the complex to the modified oligomer was extreme exchange broadening of the modified duplex resonances.

Orientation of the Guanidinium Arms. The orientation of the ancillary groups is less obvious for Δ-Rh binding to its underivatized site than in the Λ-Rh/L1 system. The methylene protons of the guanidinium arms are more difficult to distinguish; however, at least one [δ(ppm) = 5.29] appears to have a cross-peak to C13 H2′ and 2″. Analogous to the

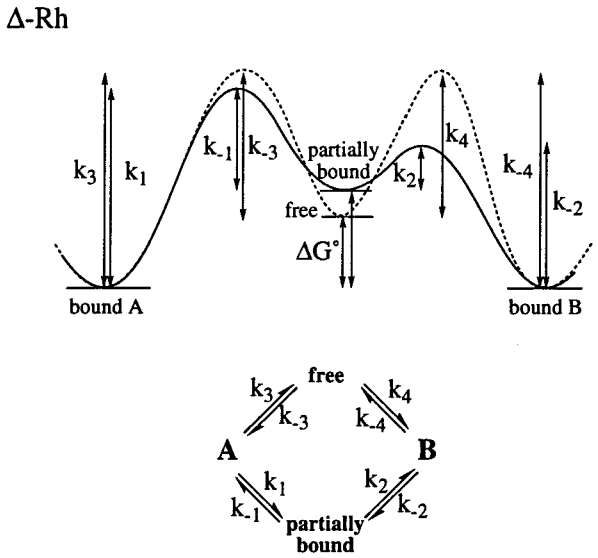

FIGURE 8: Possible reaction coordinate for the exchange between the two bound Δ-Rh conformers. The dotted line depicts the full dissociation of the complex prior to reintercalation, and the solid line depicts the partially intercalated intermediate.

G8 imino shifts in the Λ-Rh/L1 spectrum, downfield shifts are observed for the imino protons of T7, G8, and G18, which suggest either hydrogen bonding into these ring systems or a decrease in local stacking.

Geometry of the Recognition Site. Consistent with the photocleavage results which show Δ-Rh cuts predominantly at the 3′ C6 (*7*, *30*), these NMR spectral changes support the conclusion that the Δ-complex is severely canted in its recognition site. The minimal difference in chemical shift that one bound conformer exhibits relative to the free DNA position indicates that the bound orientation for one conformer perturbs the DNA very little. This is fully consistent with two substantially canted orientations, in which bound form A changes the chemical environment to one side of the intercalation site, but has little contact with the other, and bound form B has an analogous effect to the other side of the site.

Exchange Kinetics. A variable-temperature study of the exchange between bound conformers was undertaken for the Δ-Rh system. Due to the already extreme complexity of the spectrum, a fully titrated 1:1 sample of Δ-Rh/D1−D2 was employed rather than the partially titrated sample as in the 2:3 Λ-Rh/L1 system studied above. Thus, the rate of exchange between A and B isomers could be measured, but not the complex−DNA association rate as described for Λ-Rh/L1. The rate of exchange between bound forms (k_{obs}) was found by this method to be 37 s^{-1} at 298 K. An Eyring plot of the data was linear over this range, but gave activation parameters which were internally inconsistent, indicating that this exchange is not a simple two-species process. This is not unexpected, as the possibility of full dissociation to free DNA and Δ-Rh exists, as well as direct conformation change between bound forms without disintercalation.

One possible reaction coordinate is given in Figure 8, describing exchange between two bound conformers (A and B) through two different pathways. Full dissociation of the metal complex (dotted line) is presumed to be a somewhat more energetic process than rotating between canted position while the phi ligand remains partially intercalated (solid line).

Although the nearly equal amounts of conformers observed in the spectra at both 10 and 25 °C allow us to suggest that $\Delta G°$ between the two bound forms is nearly zero, we cannot determine if the activation barriers $\Delta G^{‡}$ between the canted modes are identical. Since the site is asymmetric, which is particularly relevant to the intercalating ligand's overlap with the base stack, it is likely the barriers are not equal, and thus have been depicted as such in Figure 8 ($k_1 \neq k_{-2}$).

Since the two bound forms are in slow exchange on the NMR time scale, only the fastest exchange pathway between A and B can be observed (Figure 8). However, along this pathway, only the rate-limiting barrier can be measured. As this reaction hypothesis suggests that full disintercalation is a higher-energy process than canting, the latter pathway would be the exchange pathway observed. The rate measured, then, would be k_1, or the rate of canting from bound form A to a more symmetric, partially intercalated intermediate. However, whether the k_{obs} value of 37 s^{-1} represents full dissociation or exchange between canted positions without disintercalating, it is a slower process than the dissociation of Λ-Rh from L1 (68 s^{-1} at 298 K). This is consistent with the greater inherent flexibility of the Λ-Rh recognition site.

DISCUSSION

Similarities and Differences between Λ-Rh/L1 and Δ-Rh/D1−D2. On the basis of these data coupled with biochemical studies (*7*, *30*), we have generated models (Figure 9) for binding by each enantiomer with its recognition site. These models do not constitute structures based upon NOE constraints. Instead, they provide a low-resolution framework through which to bring together the structural and biochemical data obtained. The 6 bp sites which the two enantiomers of 1-Rh$(MGP)_2$$phi^{5+}$ target are identical except at the central base step, suggesting the recognition elements each uses at the outside of the sequence, namely hydrogen-bonding contacts with the ancillary guanidinium groups, are similar. We have shown that both Δ-Rh and Λ-Rh complexes intercalate centrally into their respective sites from the major groove, and contact 3′-guanine bases to either side of the recognition sequences. However, modeling of the complexes with B-form DNA shows neither enantiomer is geometrically able to make these hydrogen-bonding contacts with straight intercalation into standard DNA.

The Λ-enantiomer binds centrally and symmetrically to a DNA decamer containing the 6 bp site 5′-CATATG-3′. The intercalation of the centrosymmetric complex into a centrosymmetric duplex again results in only one set of bound DNA peaks. Because this symmetry is retained, the complex must intercalate at the T5-A6 step. Further evidence for the site of intercalation comes from the loss of the NOE contacts between T5 H2′ and H2″ and A6 H8, and the large upfield shift of the T5 imino proton, due to the ring current of the neighboring phi ligand. The ancillary guanidinium groups make simultaneous symmetric contacts to 3′-G8, as evidenced by the downfield shift of G8 imino protons. This shift is consistent with hydrogen bonding into the base ring system.

Consistent with these data and previous photochemical assays (*7*, *30*), the model in Figure 9 (top) shows Λ-Rh intercalated symmetrically into its recognition site, which is unwound by 70°. With this geometry, the guanidinium arms are poised to make hydrogen bonding contacts with N7 and O6 atoms of G8 to either 3′-side of the intercalation site. The favorable contacts are estimated to be about 0.5 kcal/mol for an amino proton hydrogen bond to purine N7 or N3 (*27*). These contacts add to the stability of this bound conformer, and hence increase the relative preference for this 6 bp site. Additionally, in B-form DNA, negative slide (*28*) for both base pairs of the 5′-pyrimidine−purine-3′ intercalation step stacks the purines on top of one another, and gives better overlap of these bases with the phi intercalator. At the same time, this negative slide necessarily results in unwinding of the 5′-TA-3′ step. This is analogous to the similar but opposite winding observed in the solution structure of a similar phi-containing metallointercalator bound to DNA (*6*). The full NMR solution structure of Δ-α-[Rh-[(*R*,*R*)-Me_2trien]phi]$^{3+}$ in a duplex containing its recognition site, 5′-TGCA-3′, showed the bases to be dramatically overwound (53°), aligning the purines of opposite strands with each other and the large phi intercalator. In this case, the winding due to positive slide at a 5′-purine−pyrimidine-3′ intercalation step also allowed greater purine−intercalator overlap.

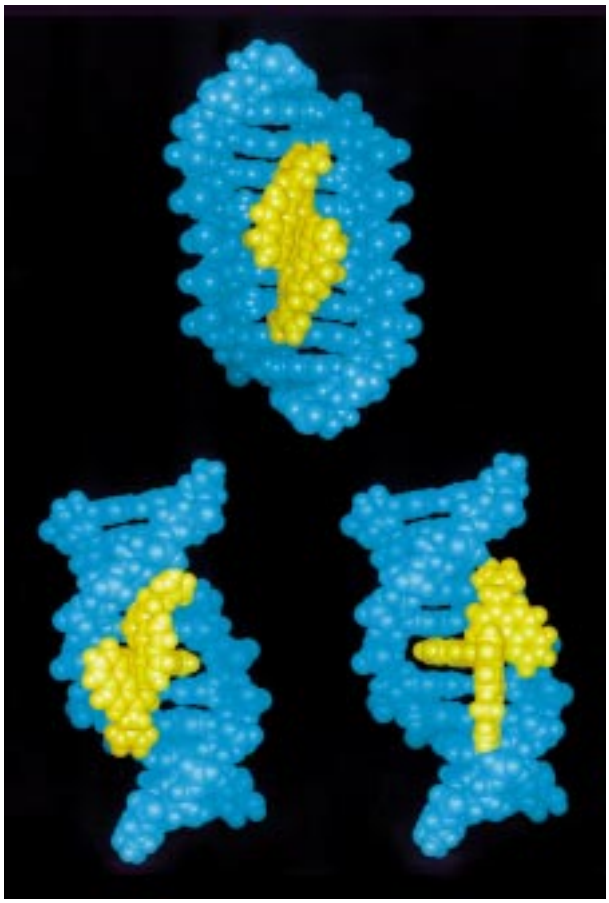

FIGURE 9: Computer-generated models of Λ- and Δ-1-Rh-$(MGP)_2$$phi^{5+}$ intercalated into their respective DNA sites of recognition (Λ-Rh, top; Δ-Rh, bottom). Δ-1-Rh$(MGP)_2$$phi^{5+}$ is shown in two canted conformations in its site in D1−D2. Models were generated by docking the respective octahedral metal complexes into B-form DNA with a generated intercalation step (unwound 20°, 6.8 Å base step). For L1, an additional 12.5° per step of unwinding was included so that the total unwinding over the 6 bp site was 70°.

The kinetics and thermodynamics of Λ-Rh binding to its site show the intercalation to be relatively facile. A variable-temperature NMR study of a partially titrated Λ-Rh/L1 sample showed the rate of association ($k_{obs-298}$) to be 68 s^{-1}, and the $\Delta G^{\ddagger}_{298}$ of binding to be 2.7 kcal/mol. Though it has been shown that the Λ-Rh-bound DNA site is significantly unwound (*7*, *30*), the 5′-ATAT-3′ sequence is known to be quite flexible, and substantial unwinding does not involve a large energetic penalty. In fact, the activation barrier for Λ-Rh binding to its DNA site is remarkably small considering the large conformational change the DNA undergoes relative to B-form DNA. As such, the binding event cannot be inducing the 70° unwinding, but must instead be trapping the flexible sequence in an unwound state upon Λ-Rh intercalation.

Δ-Rh recognizes a 6 bp sequence similar to that of its enantiomer, but through a very different mechanism. Whereas the inherent flexibility of its target sequence allows Λ-Rh to bind the unwound DNA site and thus reach the guanidinium−guanine contacts simultaneously (Figure 9, top), the data show that the Δ-Rh complex instead intercalates very asymmetrically into the central T5-C6 step. Upfield shifts for T5 and G15 imino protons define the intercalation step. Doubled peaks of the duplex upon Δ-Rh binding, however, show that there are two binding modes in slow exchange. The chemical shifts of the two bound forms indicate the complex perturbs one of the duplex strands relatively little; several DNA peaks have shifts for one bound conformer which are very similar to free duplex shifts (Figure 7 and Table 3). This suggests the complex, while intercalated at only one step, approaches the strands asymmetrically in each of two bound conformations. A significantly canted major groove approach is further indicated by the proximity of phi protons to selected DNA base protons, including A16 H2. In contrast to Δ-Ru(phen)$_2$DPQ^{2+}, which was shown to intercalate from the minor groove (*24*), neither Δ-Rh nor Λ-Rh exhibits contacts between ancillary MGP ligands and minor groove sugar protons (H1′ or H4′) adjacent to the intercalation site. Δ-Ru(phen)$_2$DPQ^{2+} has NOE cross-peaks between phen H2, H3, and H4 protons and these sugar protons, and the outside edge of the DPQ intercalating ligand (H13) is close to major groove TMe protons, indicating deep intercalation from the minor groove. The opposite orientation is indicated here for Δ-Rh, with deep, canted intercalation from the major groove placing the phi protons close to the minor groove A H2 protons.

The rate of exchange between bound Δ-Rh conformers ($k_{obs-298} = 37$ s^{-1}), presumably through partial dissociation of the complex, is a slower process than the intercalation of Λ-Rh into its site. This is consistent with the idea that the DNA sequence is more rigid than the Λ-Rh recognition site, and unlikely to be significantly unwound on binding.

A binding approach from the major groove, as well as the importance of guanidinium−guanine contacts on the outside of the site, is further supported by the dramatic loss of specificity upon 3′-N^7-deazaguanine substitution. The N7 atom in the major groove is a key element of recognition, indicating the ancillary guanine groups make contacts to 3′-G bases on both sides of the 5′-CATCTG-3′ 5′-CAGATG-3′ sequence.

These data suggest a model for the binding of Δ-Rh to its site in which the complex binds from the major groove to rigid, B-form DNA via two canted orientations (Figure 9, bottom). The model suggests that a severely canted Δ-Rh major groove approach allows the complex to make only one of the guanidinium−guanine hydrogen bonds at a time, in each of two binding modes. In each case, this necessarily brings the other arm close to the phosphate backbone, which may act as an alternative hydrogen bond acceptor for the guanidinium group. Both guanidinium−guanine contacts are necessary, however, to give Δ-Rh affinity and specificity for its site. It is clear that removing one hydrogen-bonding interaction from the major groove (3′-G N7) dramatically affects the site specificity and the kinetics of binding of Δ-Rh with its site. Rather than merely shifting the binding equilibrium between two local energy minima, the loss of this contact destabilizes the overall binding affinity of this complex for 5′-CATCTG-3′. The presence of two 3′-guanine contacts may facilitate the local affinity of Δ-Rh for 6 bp sequences of this type generally. Within this subset, further refinements in guanine−guanidinium proximity, base−phi stacking, and shape complementarity contribute to sequence preference.

Unlike the Λ-Rh recognition site, for which unwinding the central TA step promotes purine−purine and thus purine−phi overlap, in the Δ-Rh recognition site the purines are already well-stacked in standard B-form DNA. The phi ligand can overlap effectively with the purine bases by intercalating toward one edge of the major groove. Although two canting directions are still possible, this additional asymmetry suggests one canting mode may be somewhat closer to the phosphate backbone than the other. The binding asymmetry of the model also suggests that at least one, and perhaps both, Δ-Rh/D1−D2 conformers have the edge of the phi ligand directed toward the C6 sugar, poised for photocleavage, and may explain the dramatic selectivity for 3′-cleavage on one strand.

It is evident that the two enantiomers of 1-Rh(MGP)$_2$phi^{5+} utilize both direct readout and shape selection mechanisms in refining sequence selectivity, and the interplay of these subtle influences makes these complexes excellent models for the interactions of proteins with DNA. The TATA-box binding protein (TBP) recognizes and binds to a significantly bent, unwound AT-rich sequence (5′-TATATAAA-3′) (*12*). The inherent plasticity of this type of sequence is utilized by the TBP protein in distinguishing its binding site from normal B-form DNA. In an analogous manner, Λ-Rh capitalizes on the flexibility of the 5′-ATAT-3′ run to trap its recognition sequence in a conformation for which the energetically favorable hydrogen bond contacts can be made. Here shape selection depends not on a static structure, but on recognizing the inherent flexibility of a given sequence.

The recognition modes of DNA-binding proteins such as Cro, λ, and *trp* repressors and the Engrailed homeodomain may also be modeled by these metallointercalators. Both Cro and λ repressors bind to multiple, similar DNA sites, sometimes with affinities greater than that for their respective operator sequences. Albright and Matthews suggest that DNA cannot be considered simply either "operator" or "non-operator", and point out that the types of directed, base-specific hydrogen bond interactions used to select a given sequence from random DNA are supplemented by subtler, indirect effects, such as the local steric clashes and conformation (*29*). Distinguishing DNA sequences is a tradeoff

among these factors. Recent work by Haran and co-workers indicates that sequence substitutions in the operator sequence of the *trp* repressor are compensated by an increased bend angle of the DNA (*14*). Thus, indirect readout, or structural recognition, is complementing direct readout in determining the *trp* binding sites.

The Δ-Rh complex, with its equilibrium between binding states, may be considered a repressor protein or homeodomain protein mimic. Like the proteins, Δ-Rh discriminates a general recognition site, 5′-GXTCXC-3′, from random DNA. This recognition is accomplished by complementing the shape of the major groove, well-stacked intercalation into the 5′-TC-3′ step, and sequential hydrogen bonds to the 3′-G bases. The loss of one terminal 3′-G base abolishes the specificity for the site. However, within the context of this general recognition sequence, the ability of the intercalator to favorably overlap the purine bases in two distinct canted orientations and bring the guanidinium arms close to 3′-G in each of two binding conformers allows the recognition of a site which would otherwise be inaccessible to the small complex.

Δ-Rh recognition of a large site via an equilibrium of binding conformations may be considered a model for the binding to overlapped sites by the Engrailed repressor protein, as well (*13*). This protein was crystallized with the operator sequence 5′-TAATTA-3′, but NMR evidence suggests this sequence is recognized by the protein as two overlapping 5′-TAAT-3′ sites. Canting in two conformations to maximize guanidinium−guanine hydrogen bonds is a similar type of recognition of a site otherwise somewhat large for the metallointercalator. Though both 3′-G bases are necessary for Δ-Rh recognition of the full site, the two binding modes in slow exchange are similar to the solution behavior of Engrailed with the two subsets of 5′-TAATTA-3′.

CONCLUSION

The NMR spectroscopic studies presented here have corroborated earlier photocleavage results (*7*, *30*) and allow binding models of each isomer to its recognition site to be established. Consistent with the earlier studies which found that the 6 bp consensus sequence preferred by Λ-Rh was unwound by 70° upon complexation, the NMR results support a very flexible, symmetric Λ-Rh/L1 complex. The decamer retains its centrosymmetry in the presence of metal complex, and Λ-Rh makes simultaneous contacts with the major groove N7 atoms of both guanine bases flanking the site. The favorable interactions at the edge of a rather large sequence are made with little energetic cost; the complex traps the duplex during the course of its inherently fluid motion.

These NMR results also allow us to understand the binding of Δ-Rh to its sequence, nearly identical to that recognized by Λ-Rh, and to define the important contacts necessary for selectivity. The complex was expected to be somewhat small to contact 6 bp. This nonsymmetric, less flexible sequence contains the same flanking guanidine bases, and the central 5′-TC-3′ step which the parent complex, Δ-Rh(phen)$_2$phi^{3+}, marginally prefers. However, in contrast to that of the flexible 5′-ATAT-3′ site, the greater rigidity of this base stack means the ancillary guanidinium arms cannot hydrogen bond to both flanking G N7 atoms at once. The NMR spectral evidence suggests that this complex instead binds to its site in two different conformations, each of which allows one guanidinium−guanine hydrogen bond at a time, and brings the other arm close to the phosphate backbone. This involves two different substantially canted modes, probably with stacking preferentially between the purines of the intercalation site (G15-A16). Although there are two preferred intercalation modes, the shifts of the imino protons upon binding indicate there is only one intercalation site. Both 3′-guanine bases are necessary, however, to impart site specificity and slow dissociation kinetics with the 6 bp site.

These results show the shape of the metallointercalator complex, which directs the enantiomeric preference for intercalation step, and the hydrogen-bonding contacts which add binding stability at base steps flanking the intercalation, as well as the inherent structure and flexibility of the DNA sequence, combine to generate the exquisite differentiation between nearly identical recognition sequences by enantiomers of the same metallointercalator. This understanding should be exploited in the rational design of metallointercalators which target DNA sites with sequence selectivity and in the further investigation of protein−DNA interactions using small molecule mimics.

ACKNOWLEDGMENT

We thank Dr. Masako Kato for synthesizing and resolving the 1-Rh(MGP)$_2$phi^{5+} complexes and Duncan T. Odom for resolution of additional complex. We also thank Brian A. Jackson for many helpful suggestions and discussions.

SUPPORTING INFORMATION AVAILABLE

Expanded aromatic region of the ROESY spectrum of Δ-Rh with D1−D2, expanded downfield (imino) region of L1 and of Λ-Rh/L1 90:10 H_2O/D_2O NOESY spectra, and expanded downfield (imino) region of D1−D2 and of Δ-Rh/D1−D2 90:10 H_2O/D_2O NOESY spectra (4 pages). Ordering information is given on any current masthead page.

REFERENCES

1. Pabo, C. O., and Sauer, R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095.
2. Dupureur, C. M., and Barton, J. K. (1995) in *Comprehensive Supramolecular Chemistry* (Lehn, J.-M., Ed.) Vol. 5, pp 295−315, Pergamon Press, New York.
3. Johann, T. W., and Barton, J. K. (1996) *Philos. Trans. R. Soc. London, Ser. A 354*, 299−324.
4. Krotz, A. H., Kuo, L. Y., Shields, T. P., and Barton, J. K. (1993) *J. Am. Chem. Soc. 115*, 3877−3882.
5. Hudson, B. P., Dupureur, C. M., and Barton, J. K. (1995) *J. Am. Chem. Soc. 117*, 9379−9380.
6. Hudson, B. P., and Barton, J. K. (1998) *J. Am. Chem. Soc.* (in press).
7. Terbrueggen, R. H., and Barton, J. K. (1995) *Biochemistry 34*, 8227−8234.
8. Sitlani, A., Dupureur, C. M., and Barton, J. K. (1993) *J. Am. Chem. Soc. 115*, 12589−12590.
9. Barton, J. K. (1986) *Science 233*, 727−734.
10. Pyle, A. M., Rehmann, J. P., Meshoyrer, R., Kumar, C. V., Turro, N. J., and Barton, J. K. (1989) *J. Am. Chem. Soc. 111*, 3051−3058.
11. Sitlani, A., Long, E. C., Pyle, A. M., and Barton, J. K. (1992) *J. Am. Chem. Soc. 114*, 2303−2312.
12. Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) *Nature 365*, 512−519.

13. Draganescu, A., and Tullius, T. D. (1998) *J. Mol. Biol. 276*, 529−536.
14. Bereket-Samish, A., Cohen, I., and Haran, T. E. (1998) *J. Mol. Biol. 277*, 1071−1080.
15. Caruthers, M. H., Barone, A. D., Beaucage, S. L., Dodds, D. R., Fisher, E. F., McBride, L. J., Matteucci, M., Stabinsky, Z., and Tang, J. Y. (1987) *Methods Enzymol. 154*, 287−313.
16. Pyle, A. M., Long, E. C., and Barton, J. K. (1989) *J. Am. Chem. Soc. 111*, 4520−4522.
17. Wilson, W. D., Li, Y., and Veal, J. M. (1992) in *Advances in DNA Sequence Specific Agents*, Vol. 1, pp 89−165, JAI Press, Inc., Greenwich, CT.
18. Franklin, S. J., Treadway, C. R., and Barton, J. K. (1998) *Inorg. Chem.* (in press).
19. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, Inc., New York.
20. Omichinski, J. G., Clore, G. M., Robien, M., Sakaguchi, K., Appella, E., and Gronenborn, A. M. (1992) *Biochemistry 31*, 3907−3917.
21. Neuhaus, D., Nakaseko, Y., Nagai, K., and Klug, A. (1990) *FEBS Lett. 262*, 179−184.
22. Dupureur, C. M., and Barton, J. K. (1994) *J. Am. Chem. Soc. 116*, 10286−10287.
23. David, S. S., and Barton, J. K. (1993) *J. Am. Chem. Soc. 115*, 2984−2985.
24. Greguric, I., Aldrich-Wright, J. R., and Collins, J. G. (1997) *J. Am. Chem. Soc. 119*, 3621−3622.
25. Perrin, C. L., and Dwyer, T. J. (1990) *Chem. Rev. 90*, 935−967.
26. Searle, M. S., and Williams, D. H. (1993) *Nucleic Acids Res. 21*, 2051−2056.
27. SantaLucia, J., Kierzek, R., and Turner, D. H. (1992) *Science 256*, 217−219.
28. Dickerson, R. E. (1989) *J. Biomol. Struct. Dyn. 6*, 627−634.
29. Albright, R. A., and Matthews, B. W. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3431−3436.
30. Terbrueggen, R. H., Johann, T., and Barton, J. K. (1998) (submitted for publication).
31. Haq, I., Lincoln, P., Suh, D., Norden, B., Chowdhry, B. Z., and Chaires, J. B. (1995) *J. Am. Chem. Soc. 117*, 4788−4796.
32. Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) *Nature 365*, 520−527.

BI981798Q