Recognition of Base Mismatches in DNA by 5,6-Chrysenequinone Diimine Complexes of Rhodium(III): A Proposed Mechanism for Preferential Binding in Destabilized Regions of the Double Helix†

Brian A. Jackson and Jacqueline K. Barton*

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

Received November 24, 1999; Revised Manuscript Received March 16, 2000

ABSTRACT: 5,6-Chrysenequinone diimine (chrysi) complexes of rhodium(III) have been shown to be versatile and specific recognition agents for mismatched base pairs in DNA. The design of these compounds was based on the hypothesis that the sterically expansive chrysi ligand, which should be too wide to readily intercalate into B-DNA, would bind preferentially in the destabilized regions of the DNA helix near base mismatches. In this work, this recognition hypothesis is comprehensively explored. Comparison of the recognition patterns of the complex $[Rh(bpy)_2(chrysi)]^{3+}$ with a nonsterically demanding analogue, $[Rh(bpy)_2(phi)]^{3+}$ (phi = 9,10-phenanthrenequinone diimine), demonstrates that the chrysi ligand does indeed disfavor binding to B-DNA and generate mismatch selectivity. Examination of mismatch recognition by $[Rh(bpy)_2(chrysi)]^{3+}$ in both constant and variable sequence contexts using photocleavage assays indicates that the recognition of base mismatches is influenced by the amount that a mismatch thermodynamically destabilizes the DNA helix. Thermodynamic binding constants for the rhodium complex at a range of mismatch sites have been determined by quantitative photocleavage titration and yield values which vary from 1×10^6 to 20×10^6 M^{-1}. These mismatch-specific binding affinities correlate with independent measurements of thermodynamic destabilization, supporting the hypothesis that helix destabilization is a factor determining the binding affinity of the metal complex for the mismatched site. Although not the only factor involved in the binding of $[Rh(bpy)_2(chrysi)]^{3+}$ to mismatch sites, a model is proposed where helix destabilization acts as the "door" which permits access of the sterically demanding intercalator to the base stack.

Octahedral metallointercalators containing the phenanthrenequinone diimine (phi)[1] intercalating ligand have proven to be versatile complexes in the design of molecular recognition agents for DNA (Figure 1) (*1*−*6*). The binding affinity provided by the intercalative interaction efficiently delivers these complexes to the major groove of DNA where shape-selective interactions, direct sequence readout, or combinations thereof can modulate the specificity of the complex for specific base sequences. In addition, the photoreactivity of the complexes, their relative ease of modification, and their rigid octahedral shape are all characteristics that can be used advantageously (*1*, *2*).

In an effort to broaden the applicability of the rhodium intercalators beyond the recognition of specific base sequences or DNA structures, our laboratory recently became interested in the design of small molecules to bind base-mismatch sites in DNA (*7*−*10*). Mismatch sites in DNA present a unique challenge in molecular recognition. Because they can involve any of the four possible DNA bases, a general mismatch recognition agent cannot rely on the direct readout interactions which have been applied by many small-molecule and protein systems to identify unique base sequences in a long nucleic acid polymer. Although the many possible mismatch sites which can exist in DNA introduce a significant amount of heterogeneity to this recognition design problem, extensive study of DNA mismatches themselves has indicated that many such sites do have something in common: the presence of mispaired bases often destabilizes the DNA helix. This destabilization is reflected both in the thermodynamics of DNA structure (*11*−*15*), which can often be observed as a change in the temperature at which two DNA strands thermally denature, and in the kinetics of base-pair opening at and near the mismatched site (*16*−*23*). It was hypothesized that this helix destabilization, which

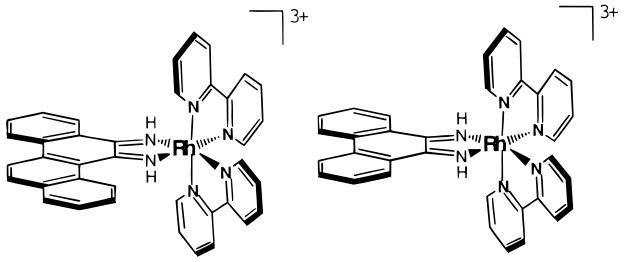

FIGURE 1: Structures of the complexes Δ-$[Rh(bpy)_2(chrysi)]^{3+}$ (left) and Δ-$[Rh(bpy)_2(phi)]^{3+}$ (right).

† We gratefully acknowledge the financial support of the National Institutes of Health (GM33309). We also thank the Parsons Foundation for fellowship support to B.A.J.

* To whom correspondence should be addressed. Email: jkbarton@caltech.edu.

[1] Abbreviations: bpy, 2,2′-bipyridine; chrysi, 5,6-chrysenequinone diimine; phi, 9,10-phenanthrenequinone diimine; EDTA, ethylenediaminetetraacetic acid.

10.1021/bi9927033 CCC: $19.00
Published on Web 04/26/2000

affects a majority of mismatch sites, could be used as a molecular recognition element.

The strategy adopted to recognize destabilized regions of the DNA helix exploited a sterically demanding DNA intercalator. Building on experience with 9,10-phenanthrenequinone diimine (phi) complexes of rhodium(III), a derivative of this intercalating ligand was designed. Phi intercalators have been demonstrated to bind readily to B-form DNA by insertion of their planar aromatic surface between two DNA base pairs (*5*). This interaction has been characterized structurally both by NMR spectroscopy (*6*) and by X-ray crystallography (*24*). With a width complementary to that of a standard DNA base pair, the phi ligand fits snugly into the nucleic acid. The 5,6-chrysenequinone diimine (chrysi) intercalating ligand, like phi, is a planar aromatic surface for insertion into DNA, but, unlike phi, contains an additional ring on one side of the molecule (Figure 1). Molecular modeling studies indicated that this ring, which increases the width of the ligand by approximately 2.1 Å, would make the molecule too large to bind readily to B-DNA. It was hypothesized, however, that the destabilized regions near mismatch sites might provide access for the sterically demanding ligand. Thus, first generation complexes incorporated the tetracyclic chrysenequinone diimine (chrysi) ligand (Figure 1). It is noteworthy that the solution structure of a somewhat analogous bulky benzochrysene−dA adduct covalently attached and intercalated within a DNA duplex was recently characterized using NMR methods (*25*).

Chrysi complexes have now been demonstrated to be versatile recognition agents that bind and, upon photoactivation, cleave at over 80% of the possible single base mismatch sites in DNA, and with specificity sufficient to cleave a single mismatch site in a 2.7 kilobase DNA plasmid (*8*). Although the utility of these complexes for mismatch recognition has already been demonstrated, a more comprehensive and quantitative understanding of this recognition was needed As a result, the studies described here examine the binding and photocleavage properties of the molecules more comprehensively to inform the development of a model for their binding specificity.

MATERIALS AND METHODS

Reagents, Instrumentation, and General Methods. All reagents were the highest purity commercially available and, unless otherwise noted, were used as obtained without further purification. $[Rh(bpy)_2(chrysi)]Cl_3$ was synthesized and purified, and its enantiomers were resolved as described previously (*7*, *9*). DNA synthesis was performed either on an ABI 392 DNA/RNA synthesizer (reagents from Glen Research, Sterling, VA) or at the Caltech Biopolymer Synthesis Facility. Oligonucleotides were purified on Poly-Pak II purification cartridges (Glen Research) according to the manufacturer's instructions or by reverse-phase HPLC. HPLC purification was performed on either Waters 660E or Hewlett-Packard 1050 or 1100 instruments using a Dynamax C_{18} reverse-phase column. Absorption spectra were recorded on a Beckman DU 7400 spectrophotometer.

Oligonucleotides were 5′-end-labeled with ^{32}P using [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and purified by denaturing polyacrylamide gel electrophoresis. Oligonucleotides were isolated from excised polyacrylamide slices by elution into ammonium acetate/EDTA buffer (500 mM NH_4OAc, 1 mM EDTA, pH 9) and ethanol-precipitated. DNA photocleavage reactions were performed on samples of end-labeled DNA with unlabeled, identical sequence carrier DNA in running buffer (50 mM Tris Base, 20 mM sodium acetate, 18 mM NaCl, pH 7.0) or in Tris buffer alone (10 mM Tris Base, pH 8.5). DNA samples were annealed by heating to 90 °C and controlled-cooling to room temperature using a Perkin-Elmer Cetus DNA thermocycler. Concentrations of metal complex solutions were determined spectrophotometrically using $\epsilon_{271} = 63\,800\ M^{-1}\ cm^{-1}$ for $[Rh(bpy)_2(chrysi)]^{3+}$ and $\epsilon_{358} = 19\,400\ M^{-1}\ cm^{-1}$ for $[Rh(bpy)_2(phi)]^{3+}$. Concentrations of oligonucleotides were determined spectrophotometrically using per-base extinction coefficients of A: 8600 $M^{-1}\ cm^{-1}$; C: 6800 $M^{-1}\ cm^{-1}$; G: 9700 $M^{-1}\ cm^{-1}$; T: 8400 $M^{-1}\ cm^{-1}$ which are based on values from single-stranded DNA homopolymers. Extinction coefficients for the double-stranded regions of DNA hairpins were systematically reduced 15% from the calculated values based on temperature-dependent hypochromicity measurements made on a representative sample of these hairpins. Unless otherwise noted, DNA concentrations are given per oligonucleotide. Samples were prepared between 5 and 30 min before irradiating on an Oriel Hg/Xe lamp system equipped with a monochromator, IR filter, and 300 nm cutoff filter. The irradiations were performed in open, horizontally positioned 1.7 mL microcentrifuge tubes. Dried irradiation samples were resuspended in denaturing formamide loading dye and electrophoresed on 20% denaturing polyacrylamide gels unless indicated otherwise. Images of the gels were obtained by phosphoimagery (Molecular Dynamics) and quantitated using the Image Quant software package; phosphoimagery provides at least 5 orders of magnitude in linear range for quantitation.

Comparison of DNA Photocleavage by $[Rh(bpy)_2(chrysi)]^{3+}$ and $[Rh(bpy)_2(phi)]^{3+}$. Solutions of $[Rh(bpy)_2(chrysi)]^{3+}$ or its phi analogue, $[Rh(bpy)_2(phi)]^{3+}$, bound to the 5′-end-labeled oligonucleotide, 5′-GCCCTACGCA***C***GACATGAT-GCTGCGG-3′, annealed to one of two complementary strands were prepared in 1× running buffer. The first complement (with a G opposite ***C***) produced a properly matched duplex; the second (with a C opposite) produced a CC mismatch at the site. Solutions were prepared at concentrations increasing from 1×10^{-9} to 1×10^{-5} M in 10-fold increments with equimolar amounts of the metal complex and the DNA duplex. The samples were prepared 4 min before irradiation at 365 nm for 15 min on the lamp system described above. Dark control samples were prepared containing 1×10^{-5} M metal and DNA but were not irradiated; light control samples were prepared containing 1×10^{-5} M DNA and irradiated in the absence of metal complex. Samples were dried after irradiation and brought up in formamide loading dye before analysis by gel electrophoresis.

Mismatch-Specific DNA Photocleavage by Δ-$[Rh(bpy)_2(chrysi)]^{3+}$: Single Sequence Contexts. Mismatch-specific photocleavage with a constant sequence context surrounding the mispair site was examined on a set of hairpin oligonucleotides of the general sequence PQ/TC shown in Figure 2. In this nomenclature, PQ refers to the mismatch site [systematically substituted with each DNA base mismatch

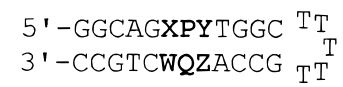

FIGURE 2: General schematic of mismatch sequence context hairpins. The base pairs XW and YZ flank the mismatch site PQ.

(CC, AC, CA, ..., etc.) and the two Watson−Crick base pairs] and TC refer to the bases on the top strand occupying the X and Y sites, respectively. Photocleavage samples (6.7 μM DNA polymers, 1.7 μM Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$, 1× running buffer) were prepared 4 min before irradiation at 313 nm for 5 min. Samples were dried and electrophoresed as described. The fraction of DNA cleaved at the mismatch site was determined by phosphorimagery and the Imagequant software package (Molecular Dynamics). This data set was combined with that derived from an analogous experiment reported previously (*7*).

Mismatch-Specific DNA Photocleavage by Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$: Multiple Sequence Contexts. Individual primer scale (0.04 μM) preparations of 29 base pair oligonucleotide hairpins containing one of the eight possible DNA mispairs in a unique single base sequence context were synthesized as described. The general structure of these oligonucleotides, containing a mismatch PQ flanked by base pairs XW and YZ, is shown in Figure 2; hairpins are identified by nomenclature (PQ/XY) as described above. In total, 104 oligonucleotides were made which represent all the possible single base contexts around the 8 mispairs when rotational symmetry is taken into account. The oligonucleotides were 5′-end-labeled with ^{32}P and purified as described above. Samples of labeled DNA were combined with unlabeled carrier and annealed as described. Aliquots of each oligonucleotide solution (10 μL, 20 μM DNA polymers, 10 mM Tris Base, pH 8.0) were combined with an equal volume of Δ-[Rh(bpy)$_2$(chrysi)]Cl$_3$ solution (10 μM in deionized H$_2$O), producing irradiation samples 10 μM in DNA and 5 μM in metal complex. Irradiations were performed at 313 nm for 10 min. DNA samples of identical concentration were incubated with metal complex in the absence of light (dark controls) or irradiated in the absence of metal complex (light controls). The cleavage intensity at the two bases to the 3′ side of the mismatch site on both DNA strands was quantitated using the Imagequant software package. The total specific cleavage was converted to a "fraction cleaved" ratio by dividing by the sum of the uncleaved parent DNA band and the mismatch specific cleavage. This cleavage fraction was corrected for nonspecific nucleic acid damage by subtracting identical measurements made on the light control samples for each irradiation lane.

Determination of Mismatch Specific Association Constants for Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$. To determine the thermodynamic association constants for Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$ with mismatch sites of interest, photocleavage titrations were performed on the oligonucleotide hairpins CC/GC, TT/GC, CT/GC, AC/GC, and AA/GC (nomenclature PQ/XY as described above). Although more extensive discussion of the methodology for measuring these binding constants is included elsewhere (*10*), a brief summary follows. Photocleavage was performed with Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$ at concentrations from 3×10^{-10} to 1×10^{-3} M DNA polymers in 1× running buffer. The Rh:DNA ratio was held constant at a value of 1:10. Samples were photolyzed at 313 nm for 5 min after preincubation for 4 min. Samples were dried under vacuum, and electrophoresis was performed as described. The fraction cleaved at the mismatch site was quantitated, expressed as a fraction of the uncleaved parent DNA band, and fit to a single site, one parameter binding model.

RESULTS AND DISCUSSION

Comparison of the Binding Specificity of Phi and Chrysi Complexes. Examination of the sites of DNA photocleavage by [Rh(bpy)$_2$(chrysi)]$^{3+}$ and [Rh(bpy)$_2$(phi)]$^{3+}$ demonstrates that the presence of the additional ring on the chrysene-quinone diimine intercalating ligand produces significant differences in the DNA cleavage properties of the two complexes. As shown in Figure 3, [Rh(bpy)$_2$(chrysi)]$^{3+}$ specifically cleaves the DNA immediately to the 3′ side of the CC base mismatch; in the absence of the mismatch site, no DNA cleavage is observed until the highest concentrations of oligonucleotide (10 μM) and metal complex (10 μM) are reached. In contrast, the parent complex, [Rh(bpy)$_2$(phi)]$^{3+}$, shows a large number of intense cleavage sites even in the absence of the base mismatch. Nonetheless, in the presence of the mismatch, the phi complex shows an intense band at the same 3′ base where cleavage by the chrysi complex is observed. Although this does indicate a probable *high-affinity* interaction of the phi complex with the mismatch site, the *specificity* of this interaction is low. In contrast, [Rh(bpy)$_2$(chrysi)]$^{3+}$ interacts with the mismatch site with a high degree of specificity; as shown in Figure 3, the mismatch specific product is, by far, the predominant cleavage product. Quantitation of these lanes yields a ratio in cleavage intensity at the mismatched site versus cleavage intensity at all other sites of ~0.3 for the phi complex and 3.5 for the chrysi complex.

Recognition of Mismatch Sites in Single Sequence Contexts. Photocleavage data for Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$ with oligonucleotides containing varied DNA base mismatches are included in Figure 4. In this case, the strongest cleavage is observed at the CC, CA, and CT sites. An intermediate level of cleavage occurs at the TC, AC, TC, and TT sites. A much smaller, but still detectable, level of cleavage is observed at the GA, GG, TG, and AG sites. Negligible levels of photocleavage are observed in the absence of a mismatch site; the bands at the very top of the GC Watson−Crick paired gel lane to the far right are artifactual and were observed in the light and dark control samples. Beyond serving as an additional demonstration of the recognition properties of chrysi complexes, however, these constant sequence context data also provided the opportunity to make a first examination of the proposed relationship between DNA mismatch recognition and the helix destabilization caused by the mispaired bases.

In these first experiments, where the base sequence around the mismatch is held constant, we make the assumption that, to first order, the efficiencies in photocleavage for the different sites are similar enough to allow direct comparison of photocleavage intensities from mismatch site to mismatch site. On that basis, the DNA photocleavage fraction for Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$ at each mismatch in Figure 4 and, in addition, that of a previously reported experiment in a different sequence context (*7*) were quantitated by integration; nonspecific light damage was controlled by subtracting

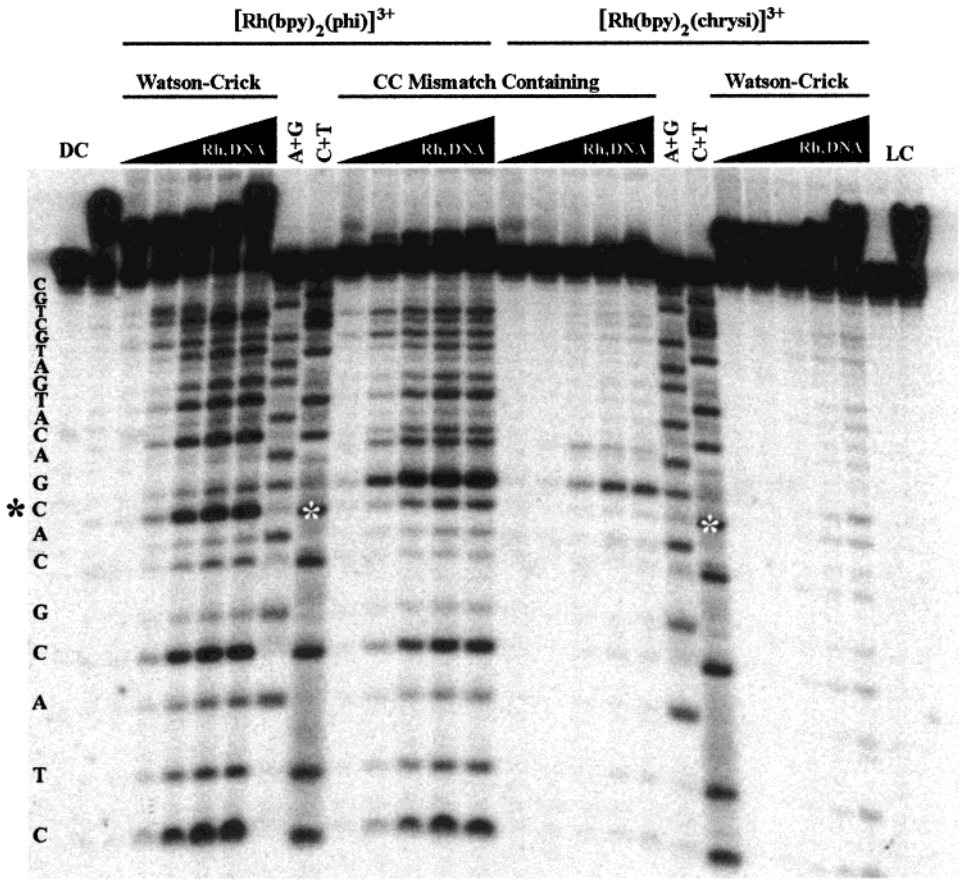

FIGURE 3: DNA photocleavage by $[Rh(bpy)_2(chrysi)]^{3+}$ and $[Rh(bpy)_2(phi)]^{3+}$ on DNA samples with and without a CC base mismatch. Samples of the 26mer 5′-end-labeled oligonucleotide: 5′-GCCCTACGCA***C***GACATGATGCTGCGG-3′ combined with one of two complementary strands (one producing a properly paired "Watson−Crick" duplex and the other containing a CC mismatch at the site indicated by boldface italic in the sequence and with asterisks on the gel) were irradiated for 15 min with equal concentrations of either $[Rh(bpy)_2(chrysi)]^{3+}$ or $[Rh(bpy)_2(phi)]^{3+}$. Sample concentrations increased from 1×10^{-9} to 1×10^{-5} M in 10-fold increments as indicated by the gradient bars above the lanes. Lanes marked DC are DNA samples incubated with metal in the absence of light; those labeled LC are DNA samples irradiated in the absence of metal. Note the high level of specificity shown by the chrysi complex for the mismatch site; in addition, in the absence of a mismatch, cleavage of the DNA by the complex is only observed at the highest concentrations. In contrast, although the phi complex does show affinity for the mismatch site, its level of specificity is very low as numerous other high-affinity sites for the complex exist both in the presence and in the absence of the mispaired bases.

an identical integrated volume from the control samples for each irradiation lane. The resulting fractions were plotted versus the thermodynamic destabilization caused by the incorporation of each mismatch into the DNA helix (Figure 5). These values were taken from thermal denaturation experiments and allow calculation of the thermodynamic consequences of mismatch incorporation in any sequence context (*11*−*15*). It is clear from these plots that the extent to which Δ-$[Rh(bpy)_2(chrysi)]^{3+}$ promotes DNA photocleavage increases with increasing thermodynamic destabilization at the mismatch sites.

Recognition of Mismatch Sites in Multiple Sequence Contexts. Although the data for a constant sequence context discussed above do suggest a relationship between increased thermodynamic destabilization caused by a DNA mismatch and more facile recognition by $[Rh(bpy)_2(chrysi)]^{3+}$, a broader look at its mismatch-specific cleavage shows that this correlation is not absolute. To make this examination, DNA photocleavage experiments were performed with $[Rh(bpy)_2(chrysi)]^{3+}$ at each of the eight base mismatches in all possible sequence contexts. The intensities of all mismatch specific photocleavage bands were expressed as a cleavage fraction as described earlier. Examples of these metal complex-specific cleavage fractions, plotted against the thermodynamic destabilization associated with the CC and GG mismatch sites, are shown as Figure 6. Complete plots for all the mismatches, in addition to the gel data from which they are derived, are included in the Supporting Information.

Unlike the previous sets of constant sequence context data, the degree to which mismatch-induced helix destabilization appears to be correlated with the observed photocleavage efficiency varies greatly from mismatch to mismatch. In the two cases presented, the CC mismatch (Figure 6A) shows no correlation between the variables while the GG mismatch (Figure 6B) displays an increase in photocleavage intensity with increasing destabilization at the mismatch site. The other mismatches, whose helix destabilizing effects lie between the very disruptive CC and the, on average, stabilizing GG mismatch, display intermediate degrees of correlation between the two variables. In general, experiments with less destabilizing mismatches yield a stronger correlation between photocleavage intensity and helix disruption.

It is also noteworthy that when individual data points out of these data set are examined, there is a lack of a simple correlation between the thermodynamic destabilization caused by the incorporation of a base mismatch and $[Rh(bpy)_2(chrysi)]^{3+}$-sensitized photocleavage intensity. In some cases, sites with negative destabilization values (i.e., the presence

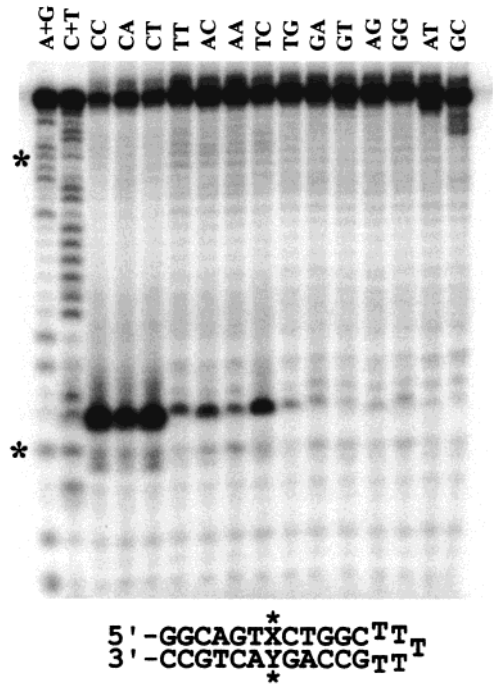

FIGURE 4: Mismatch-specific DNA photocleavage by Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$ on a family of hairpin oligonucleotides with 365 nm light. Mismatch-specific photocleavage was examined on a set of hairpin oligonucleotides shown in the inset where ***XY*** was substituted with all the possible combinations of DNA base mismatches (CC, AC, CA, ..., etc.) and the two Watson−Crick base pairs. Photocleavage samples [6.7 μM DNA, 1.7 μM Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$, in buffer (50 mM Tris Base, 20 mM sodium acetate, 18 mM NaCl, pH 7.0)] were prepared 4 min before irradiation at 313 nm for 5 min. Samples were dried and electrophoresed as described, and the resultant autoradiograph is shown.

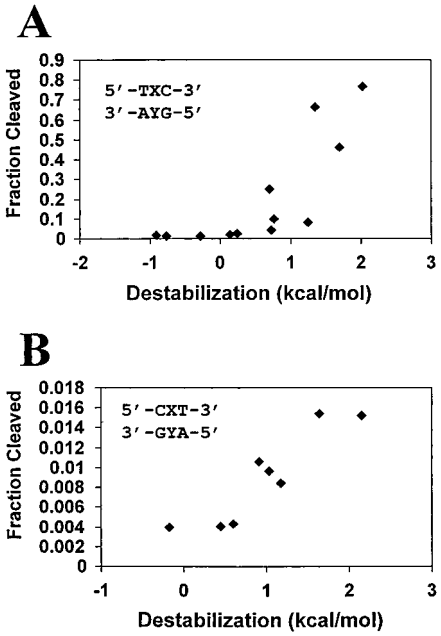

FIGURE 5: Plot of photocleavage at the mismatch site versus mismatch-dependent thermodynamic destabilization in the (A) 5′-T***X***C-3′ and the (B) 5′-C***X***T-3′ sequence context. Data from the photocleavage gel shown in Figure 4 and in ref *7* were extracted and quantitated using the Image Quant software package. Cleavage at the mismatch site was expressed as a fraction of the parent DNA band and plotted versus the thermodynamic destabilization caused by the mismatch as determined in ref *11*.

of the mismatch stabilizes the duplex compared to the Watson−Crick DNA) show a higher cleavage fraction than more destabilized sites, while sites with greater destabilization display less photocleavage.

Mismatch Specific Association Constants of [Rh(bpy)$_2$(chrysi)]$^{3+}$. The assumption that the photocleavage intensity observed at a mismatch is directly proportional to the binding affinity of [Rh(bpy)$_2$(chrysi)]$^{3+}$ at the site may not be wholly valid. In an attempt to control for the effects that differences in photocleavage efficiency might have on DNA photocleavage intensities, the thermodynamic association constants of Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$ at a variety of mismatch sites were therefore determined. Equilibrium binding isotherms for the complex at mismatch sites were constructed by photocleavage titration. By increasing the concentrations of both the rhodium complex and a short DNA oligonucleotide hairpin containing the mismatch site, it is possible to generate binding curves which are unaffected by the particular value of the DNA photocleavage efficiency at the site (*10*). Parallel experiments were conducted on various mismatch-containing oligonucleotide hairpins. The resulting binding constants are summarized in Table 1.

These data demonstrate that Δ-[Rh(bpy)$_2$(chrysi)]$^{3+}$ can bind very strongly to mismatch sites with association constants as high as $2(1) \times 10^7$ M^{-1}. Such high binding constants compare very favorably with the previously reported value of $4(2) \times 10^4$ M^{-1} for binding of the complex to properly matched B-DNA (*7*). Like the photocleavage data discussed above, these association constants exhibit some correlation between mismatch-induced helix destabilization and metal complex binding, but the correlation is not strict. For the first series of constant-context association constants (Table 1, top section), the more destabilized sites generally have higher binding constants. However, in comparing the association constants at the TT and CT mismatches, for example, it is apparent that even though the TT has a higher destabilization value by approximately 38%, the binding constant at CT exceeds TT by 42%.

These particular data on the CC mismatch sites also highlight the risk associated with equating photocleavage intensity with the binding affinity of the metal complex at a mismatch site. The tightest binding site, CC/AT, has an observed cleavage fraction of 0.14. The weakest binding site, CC/AG (CC/CT reversed), has a comparable observed cleavage fraction of 0.13. However, it is the site with intermediate binding affinity, CC/GC, which shows the strongest cleavage at 0.40. It is not surprising that the binding of the rhodium complex to the highly disordered CC mismatch sites might result in different photocleavage efficiencies depending on the orientation adopted by the intercalator in the DNA.

A Model for Mismatch Recognition. Perhaps it is not surprising that the single characteristic of thermodynamic destabilization does not show a tight correlation with these DNA photocleavage data. The chrysenequinone diimine ligand, because of its steric bulk, was designed to be selective for open or disordered sites in DNA. It is reasonable to assume that mismatch sites which have a destabilizing effect on the thermodynamics of the DNA duplex will be more likely to have such open and disordered sites at or near the mismatch. Nonetheless, it is the DNA structure, not the thermodynamic destabilization itself, which is recognized by

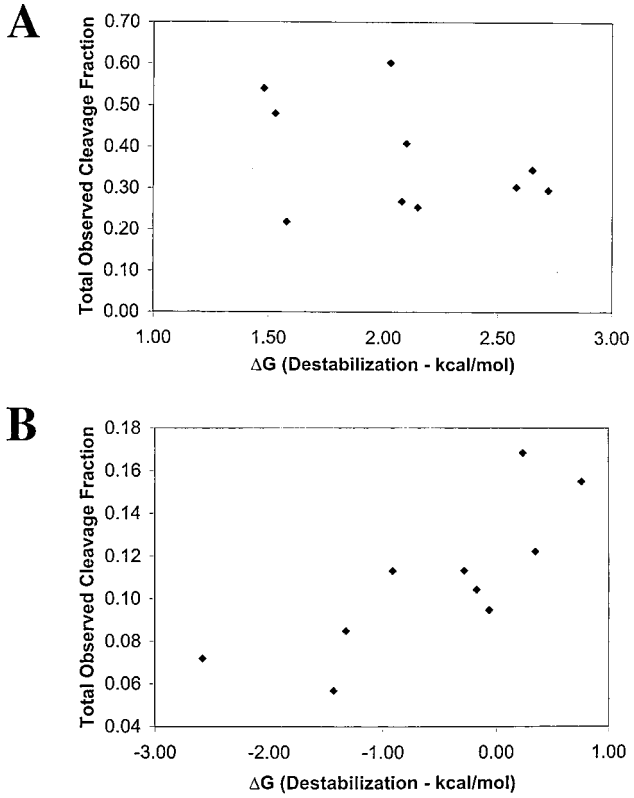

FIGURE 6: Quantitated 313 nm photocleavage by Δ-$[Rh(bpy)_2(chrysi)]^{3+}$ on DNA hairpins containing (A) CC and (B) GG base mismatches plotted versus thermodynamic destabilization as determined in ref *11*. Photocleavage intensity, adjusted for damage observed in irradiations of the DNA in the absence of metal complex, was quantitated at the two bases to the 3′ side of each mismatch site on both strands of the DNA. This value, serving as an approximation of the total cleavage at each mismatch in each sequence context, was expressed as a cleavage fraction by dividing by the sum of the quantitated photocleavage bands and the uncleaved full-length DNA band. No correlation between photocleavage and the thermodynamic destabilization is evident for the CC mismatch, but for the GG mismatch some correlation is apparent.

the complex. As a result, in these analyses, the thermodynamic parameter is actually standing proxy for an unknown quantity, the extent that each of these mismatches, in all their different sequence contexts, generate open and accessible structures in the DNA which facilitate binding of the metal complex.

It is also important to note that such open and disordered structures are not the *only* parameters which dictate the binding affinity of $[Rh(bpy)_2(chrysi)]^{3+}$ for a given mismatch site. The binding affinity of this metal complex for the DNA is primarily provided by the stacking of the intercalating ligand into the DNA helix. This interaction will clearly be influenced by the identity of the bases on either side of the mismatch and how those bases are positioned to participate in stacking. In a related manner, this binding energy will be affected by the nature of the mismatch itself. If the mismatched base is forced into an extrahelical conformation by the binding of the complex, any stacking interactions with the mismatch will not occur; if the mismatch is still within the helix, however, the particular geometry and stacking surface presented to the ligand will have a large effect on the binding affinity.

Table 1: Thermodynamic Binding Constants of Δ-$[Rh(bpy)_2(chrysi)]^{3+}$ at Varied Mismatch Sites[a]

mismatch site[b]	thermodynamic destabilization[c] (kcal/mol)	K_B[d] ($\times 10^6$ M^{-1})
CC/GC	1.48 (0.20)	10.4 (7.3)
TT/GC	1.10 (0.20)	2.31 (0.36)
CT/GC	0.80 (0.72)	3.28 (0.65)
CA/GC	0.61 (0.51)	2.71 (0.50)
AA/GC	0.18 (0.30)	0.29 (0.26)
CC/AT	2.72 (0.16)	21.6 (11.4)
CC/AG	2.15 (0.18)	0.84 (0.10)

[a] Mismatch thermodynamic values obtained from refs *11*−*15*. [b] Mismatch sites are identified with the general nomenclature PQ/XY where PQ identifies the nature of the base mismatch itself and XY the bases flanking the site to the 5′ and 3′ sides, respectively, on one strand of a DNA duplex. [c] Mismatch destabilization values from refs *11*−*15* were generated from thermodynamic parameters of duplex formation measured from thermal melting of oligonucleotides in 1 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7.0. Errors indicated in parentheses. [d] Binding constants of Δ-$[Rh(bpy)_2(chrysi)]^{3+}$ to mismatch-containing oligonucleotides were measured by photocleavage titration in 1× running buffer (50 mM Tris Base, 20 mM NaOAc, 18 mM NaCl, pH 7.0) at ambient temperature. Standard deviations indicated in parentheses.

Nonetheless, the results presented here do suggest a mechanistic hypothesis for the recognition of DNA base mismatches and other destabilizing lesions in DNA. Metallointercalators, and intercalators in general, must disrupt the structure of double-helical DNA upon binding. The energetic cost of forcing apart the stacked DNA bases, although compensated for by favorable energetic contributions upon binding, reduces the overall binding affinity of these molecules. Mismatch sites represent places in a nucleic acid polymer where these costs have already been paid and, as a result, are preferential sites for intercalation. As demonstrated in the data for $[Rh(bpy)_2(phi)]^{3+}$, however, this preferential binding is not enough to generate specificity. To make a molecule specific for these destabilized regions, the intercalator must be expanded to include an energetic cost to discourage binding at properly paired sites. The chrysi ligand, with its additional aromatic ring, imposes this energetic cost and shows increased specificity for these sites.

The shape of the plots is also noteworthy with respect to the recognition of destabilized sites by chrysenequinone diimine intercalators. As the identity of the base mismatch is changed and thermodynamic destabilization increased, the amount of DNA photocleavage displays a roughly sigmoidal dependence. This behavior is observed most clearly in Figure 5B but also to some extent in Figure 5A. This roughly sigmoidal dependence of the photocleavage intensity on destabilization can be used to suggest a model of recognition of disordered sites in DNA by $[Rh(bpy)_2(chrysi)]^{3+}$. At the bottom of the curve, where the destabilization at the mismatch site is the least, the intercalator has difficulty gaining access to the base stack to bind. As destabilization reaches a lower threshold value and continues to grow, the incremental increase in the opening of the base pairs and breakdown of the helical structure will allow incremental increases in binding and photocleavage. This corresponds to the central steep region of the sigmoidal curve. At some point, however, the destabilization of the helix will no longer be the dominant factor controlling the binding of the intercalating ligand to the helix. Once this point is reached, other factors such as the stacking of the ligand on the adjacent

base pairs will determine the extent that the molecule can bind to a given site. This situation would correspond to the region at the top of the sigmoidal curve where photocleavage intensity levels out even with increasing destabilization.

A useful analogy for this proposed model of molecular recognition is that of an opening door: At the bottom of the curve, where the door is closed or only slightly ajar, binding is still prohibited or strongly disfavored. As the door opens, the increasing access will allow tighter and tighter binding. Once the door is sufficiently open to allow "complete" access, however, opening it even wider provides no added benefit.

The data obtained in the multi-sequence context experiments are also consistent with this proposed model for recognition. In the discussion of the data presented in Figure 6, it was noted that data on the GG mismatch showed a much stronger correlation with helix destabilization than data on the CC mismatch. In the mismatches which are either stabilizing or destabilize the helix to a small extent (Table 1), increases in thermodynamic destabilization do result in corresponding increases in DNA photocleavage. This is consistent with these sites being "access limited", where small increases in the amount of destabilization have a direct and proportional effect on the amount of metal complex which is able to bind. In case of the more destabilizing mismatches, however, there is no indication of a direct relationship between thermodynamic destabilization and the extent of DNA photocleavage (Figure 6A and Supporting Information). Especially in the CC and CT cases, there is *no* obvious correlation. This observation is consistent with these mismatches being above the threshold level of destabilization where additional disorder does not translate into additional metal complex binding. In these cases, it is likely that the other factors discussed above, including stacking of the chrysi ligand and the overall structure of the sites, dominate the binding affinity and subsequent photocleavage. As a result, increases in destabilization would not have a positive effect on binding and, depending on their geometrical consequences, might result in reduced cleavage. The association constants measured for the chrysi complex at various mismatch sites are also consistent with helix destabilization being an important, but not the sole determining, factor for the binding of the complex to nucleic acids.

This porposed model for mismatch recognition therefore well represents the data presented for $[Rh(bpy)_2(chrysi)]^{3+}$. Moreover, this model may provide a general framework for for the design of small-molecule recognition agents for mismatches and other destabilized sites in nucleic acids.

ACKNOWLEDGMENT

B.A.J. acknowledges the NSF and the Parsons Foundation for predoctoral fellowships.

SUPPORTING INFORMATION AVAILABLE

Gel images and data plots of mismatch recognition by $[Rh(bpy)_2(chrysi)]^{3+}$ in multiple sequence contexts are available (4 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. (a) Erkkila, K. E., Odom, D. T., and Barton, J. K. (1999) *Chem. Rev. 99*, 2777. (b) Lippard, S. J. (1978) *Acc. Chem. Res. 11*, 211. (c) Chow, C. S., and Barton, J. K. (1992) *Methods Enzymol. 212*, 219. (d) Armitage, B. (1998) *Chem. Rev. 98*, 1171.
2. Johann, T. W., and Barton, J. K. (1996) *Philos. Trans. R. Soc. London, A 354,* 299−324.
3. (a) Terbrueggen, R. H., and Barton, J. K. (1995) *Biochemistry 34*, 8227−8234. (b) Terbrueggen, R. H., Johann, T. W., and Barton, J. K. (1998) *Inorg. Chem. 37*, 6874−6883.
4. (a) Sitlani, A., Long, E. C., Pyle, A. M., and Barton, J. K. (1992) *J. Am. Chem. Soc. 114*, 2303−2312. (b) Sitlani, A., and Barton, J. K. (1994) *Biochemistry 33*, 12100. (c) Sitlani, A., Dupureur, C. M., and Barton, J. K. (1993) *J. Am. Chem. Soc. 115*, 12589−12590.
5. (a) Krotz, A. H., Kuo, L. Y., Shields, T. P., and Barton, J. K. (1993) *J. Am. Chem. Soc. 115*, 3877. (b) Shields, T. P., and Barton, J. K. (1995) *Biochemistry 34*, 15049−15056. (c) Krotz, A. H., Hudson, B. P., and Barton, J. K. (1993) *J. Am. Chem. Soc. 115*, 12577.
6. Hudson, B. P., and Barton, J. K. (1998) *J. Am. Chem. Soc. 120*, 6877−6888.
7. Jackson, B. A., and Barton, J. K. (1997) *J. Am. Chem. Soc. 119*, 12986−12987.
8. Jackson, B. A., Alekseyev, V. Y., and Barton, J. K. (1999) *Biochemistry 38*, 4655−4662.
9. Mürner, H., Jackson, B. A., and Barton, J. K. (1998) *Inorg. Chem. 37*, 3007−3012.
10. Jackson, B. A. (1999) Ph.D Dissertation, California Institute of Technology.
11. Allawi, H. T., and SantaLucia, J. J. (1997) *Biochemistry 36*, 10581−10594.
12. Allawi, H. T., and SantaLucia, J. J. (1998) *Biochemistry 37*, 2170−2179.
13. Allawi, H. T., and SantaLucia, J. J. (1998) *Nucleic Acids Res. 26*, 2694−2701.
14. Allawi, H. T., and SantaLucia, J. J. (1998) *Biochemistry 37*, 9435−9444.
15. Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J. J. (1999) *Biochemistry 38*, 3468−3477.
16. Patel, D. J., Kozlowski, S. A., Ikuta, S., and Itakura, K. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 2663−2670.
17. Moe, J. G., and Russu, I. M. (1992) *Biochemistry 31*, 8421−8428.
18. Borden, K. L. B., Jenkins, T. C., Skelly, J. V., Brown, T., and Lane, A. N. (1992) *Biochemistry 31*, 5411−5422.
19. Faibis, V., Cognet, J. A. H., Boulard, Y., Sowers, L. C., and Fazakerley, G. V. (1996) *Biochemistry 35*, 14452−14464.
20. Lane, A. N., Jenkins, T. C., Brown, D. J. S., and Brown, T. (1991) *Biochem. J. 279*, 269−281.
21. Lane, A. N., and Peck, B. (1995) *Eur. J. Biochem. 230*, 1073−1087.
22. Kuwata, K., Liu, H., Schleich, T., and James, T. L. (1997) *J. Magn. Reson. 128*, 70−81.
23. Gervais, V., Cognet, J. A. H., Le Bret, M., Sowers, L. C., and Fazakerley, G. V. (1995) *Eur. J. Biochem. 228*, 279−290.
24. Kielkopf, C. L., Erkkila, K. E., Hudson, B. P., Barton, J. K., and Rees, D. C., submitted for publication.
25. Surf, A. K., Mao, B., Amin, S., Geacintov, N. E., and Patel, D. J. (1999) *J. Mol. Biol. 292*, 289−307.

BI9927033