INVESTIGATION OF RNA TERTIARY STRUCTURE AND FUNCTION BY TRANSITION METAL COMPLEXES

Thesis by

Ai Ching Lim

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1998

(Submitted August 22, 1997)

ACKNOWLEDGMENTS

I would like to thank my research advisor, Prof. Jackie Barton for her constant encouragement, support, and insight. Jackie's enthusiasm for chemistry is infectious, and it has been an inspiration to me. Jackie also has amazing chemical intuition and her insight into what was going on in my project has provided a great deal of help.

The Barton group is a very special set of people, and I have been truly privileged to have spent my graduate years in such a supportive and fun environment. The Barton group is both a product of and a testament to Jackie's enthusiasm about both science and people, and there are too many people, past and present, to thank individually. However, I would especially like to thank Yonchu Jenkins, for being always ready to grab an icecream during the day or to listen to my ramblings at all hours of the night. I would also like to thank Tim Johann, for being a good friend. I am very grateful to Prof. Christy Chow and Debbie Brown for showing me the ropes when I got started. I would also like to thank Brian Hudson, Brian Jackson, Susanne Lin and Mo Renta in particular, but I have had so many good times with people in the Barton group that I owe thanks to everybody.

Also at Caltech, I would like to thank Bryan Hathorn and Gary Mines especially for good advice and discussions. The Okumura group has also made life more fun for me, and I would especially like to thank Prof. Mitchio Okumura, for comet-watching parties and other such activities.

The wonderful support staff at Caltech, including Mo Renta, Dian Buchness, Tom Dunn, Suzanna Horvath, Steve Gould, Ellie Noe, and Chris Smith, have made so many things work smoothly, and I am very grateful to them.

I would also like to thank the people outside Caltech who have been a part of my life throughout graduate school, especially Jackie Baptist and Deborah Gross, for their friendships and email advice.

I am deeply grateful to my aunt, Shirley Lim, and my uncle, Charles Bazerman, for being always ready to welcome me into their home when I needed a break from Pasadena, and for their perspective and support on the graduate school experience. I would also like to thank my brother, Kean Hao, for his emails and conversations about non-science and real-life related topics, which cheered me up considerably.

I cannot thank my parents enough for their neverending love and support. This thesis and even going to graduate school alone would not have been possible without their encouragement. Their tacit belief in me throughout my life has been more helpful than I could say.

Finally, I would like to thank Jim Spotts for being a constant loving and understanding presence. The joy and laughter and silliness that you have brought me has enriched my life beyond words, and I am a better and happier person for knowing you.

ABSTRACT

Phenanthrenequinone diimine (phi) complexes of rhodium(III) were employed to probe RNA secondary and tertiary structure. These complexes bind via intercalation in open major grooves of RNA and upon irradiation promote strand scission. By probing both synthetic and natural molecules containing a variety of tertiary motifs, a systematic and predictive understanding of the factors involved in RNA recognition by these complexes is sought. The metal complex $Rh(phen)_2phi^{3+}$ (phen = 1,10-phenanthroline) recognizes and cleaves synthetic triple helices selectively over double helices. The cleavage sites are dependent upon maximizing overlap between the phi ligand and the basepairs, and minimizing charge repulsion between the metal complex and protonated bases. These cleavage sites have proven useful in explaining rhodium complex cleavage in natural systems such as $tRNA^{Phe}$. With these complexes, we also seek to investigate the differences and similarities in RNA and DNA secondary and tertiary folding, by probing the tertiary structure of $tDNA^{Phe}$ compared to $tRNA^{Phe}$. These complexes have elucidated the B-form nature of the DNA duplex as well as the tertiary folding of the DNA molecule, thus shedding light on the feasibility of using DNA analogs of RNA for structural studies. These shape selective probes have also been applied to probe the tertiary structure of HIV and BIV (TAR (*trans*-activation response) RNAs. Δ-$Rh(phen)_2phi^{3+}$ binds with high affinity ($K_b = 6.1 \pm 1.3 \times 10^5\ M^{-1}$) and specificity to sites at and across from a bulge region which is the recognition element for the binding of the Tat (*trans*-activating) peptide. Importantly, the metal complex recognizes an RNA base-triple the formation of which is necessary for transactivation. Derivatives of $Rh(phen)_2phi^{3+}$, $Rh(MGP)_2phi^{5+}$ (MGP = 4-guanidylmethyl-1,10-phenanthroline) and $Rh(GEB)_2phi^{5+}$ (GEB = 4-(2-guanidylethyl)-4'-methyl-2,2'-bipyridine) where guanidinium moieties have been added to the ancillary ligands of the rhodium complex, show enhanced affinity and selectivity for HIV and BIV RNA sequences. This is due to the guanidinium moieties mimicking the arginine side

chains on the native Tat peptide, and making non-specific contacts with the phosphate backbone of the RNA. However, even without these functionalities, shape-selection, matching the shape of the small metal complex to its nucleic acid target, provides sufficient selective stabilization for RNA site discrimination. Indeed, these complexes compete effectively with the specific Tat peptides for their binding sites on their respective TAR RNAs. These complexes therefore employ shape selection to recognize structural variations along the RNA polymer which are important for protein recognition. Shape-selective recognition could also be appplied to the design of novel small molecules to target nucleic acid sites with high site-selectivity, in the development of molecules to inhibit protein recognition, and, potentially, in the design of new chemotherapeutics.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1.

Introduction: RNA Secondary and Tertiary Structure and the Development of Probes to Study RNA Structure

1.1. Introduction

The secondary and tertiary structure of RNA are of great interest to biochemists. RNA structure is critical to the various functions of RNA, such as the transfer of information in the process of protein synthesis. The one-dimensional sequence of RNA encodes this diversity of three-dimensional structure. RNA secondary and tertiary structure is involved in post-transcriptional regulation of gene expression either by making available specific sequences or through the formation of specific structural motifs. Determining the conformation and stability of the structures adopted by RNA is essential to a better understanding of their functions. Moreover, the discovery of catalytic RNA[1,2] enlarges the need for relating the structure and function of RNA.

1.2. Structural Features of RNA Compared to DNA

Canonical double helical RNA is A-form in nature, compared to DNA which adopts the B-form (Figure 1.1). A-form and B-form nucleic acids are both right handed and possess two distinct and well-defined grooves, the major and minor grooves.[3] The A form possesses a deep and narrow major groove, and a wide and shallow minor groove. The A-form bases are pushed outward towards the minor groove direction and are tilted with respect to the axis by an average of 20°. On the other hand, B-form DNA is characterized by a deep and wide major groove and a narrower minor groove, and the bases are stacked in the center with the average base planes aligned perpendicular to the axis (Figure 1.1). The global conformation of A-form and B-form nucleic acids also differ in aspects such as sugar puckering. The sugar pucker in the A-conformation is primarily C3'-endo as opposed to the B-form which is C2'-endo in conformation (Figure 1.2). The RNA is

Figure 1.1. Schematic illustration of A-DNA (left) and B-DNA (right). Side view (top) and top view (bottom). Adapted from Ref. 3.

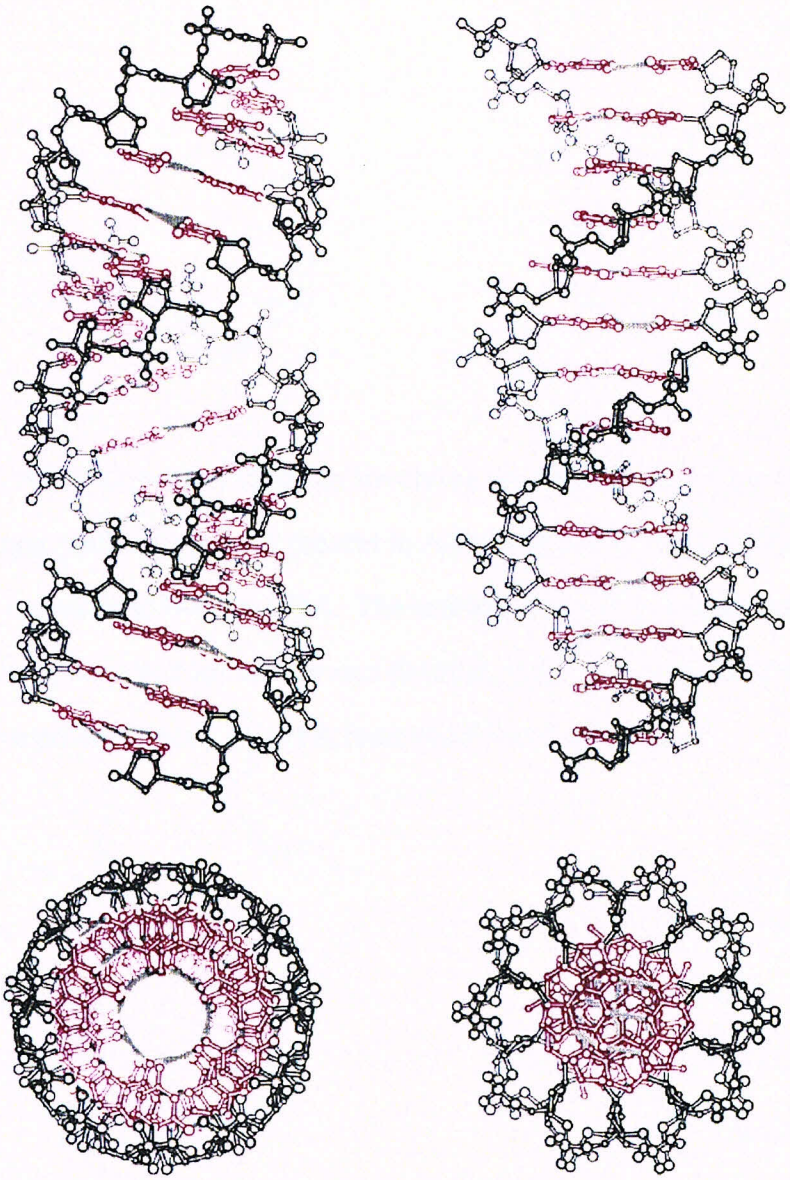
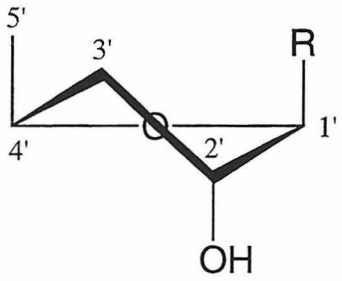
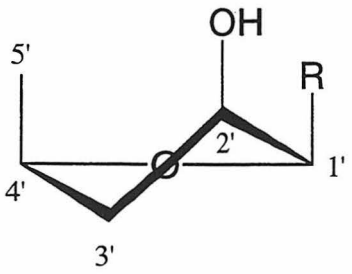


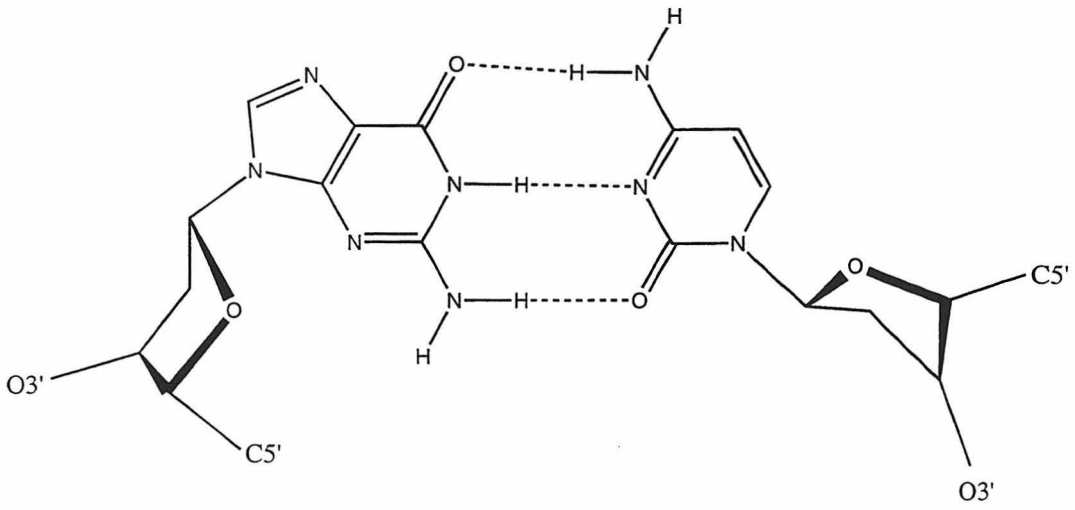
Figure 1.2. Local structural variation involving the sugar ring and bases. Schematic of C3' *endo* sugar pucker (top, left), present in A-form helices. The C2' *endo* sugar pucker (top, right) is present in B-form DNA. The rotation of the base around the glycosidic bond is categorized as *syn* (bottom, left) or *anti* (bottom, right). The *anti* conformation is exclusively used in correctly paired A-form or B-form nucleic acids.



C3' endo



C2' endo



syn *anti*

forced to adopt the A-form conformation because of the 2'-hydroxyl group. Normally paired bases in both the A and B forms adopt an anti conformation about the glycosidic bond. A summary of the characteristics of these two conformations are compared in Table 1.1.

The four bases in RNA are adenine, cytosine, guanosine and uracil, as shown in Figure 1.3. In addition, in natural RNAs, there is a wide variety of modified bases. A selection of these bases is shown in Figure 1.4. Approximately 10% of tRNA consists of these modified bases.[4] These bases can participate in novel base pairings or protein interactions, thus expanding the repertoire of RNA structure and function. Non-Watson Crick base-pairing is much more common in RNA structures than it is in DNA.

The task of recognizing RNA structures is akin to, but is certainly more complex than the recognition of double helical DNA. There is less diversity in the normal helical conformation of RNA than DNA, which can adopt A, B and Z conformations. DNA sequence within the context of a normal double helix has a large effect on conformation. For example, alternating GC tracts adopt the Z-conformation in high salt,[5] while A-tracts are bent in solution.[6] In addition, different base sequences within the helix can produce local variations in various parameters such as propeller twisting and base twists.[3] Proteins often use these variations in DNA structure for recognition, through indirect readout.[7] The structures of the complexes between TATA-box binding proteins (TBPs) and DNA solved recently with X-ray crystallography identify both direct and indirect readout interactions.[8-10] Examples of indirect readout mechanisms in these complexes are DNA bending and non-local electrostatic complementarity. These methods of protein recognition are used by RNA as well, as will be described, but are complicated by the greater diversity in RNA structure.

Table 1.1. Structural Characteristics of A-form and B-form DNA.[1]

Feature	A[2]	B[b]
Helix Diameter	23 Å	19.3 Å
Pitch	28.2 Å	33.8 Å
Base pairs per turn	11 bp	10 bp
Groove width[3] - major	2.7 Å	11.7 Å
minor	11.0 Å	5.7 Å
Groove depth[c] - major	13.5 Å	8.5 Å
minor	2.8 Å	7.5 Å
Rise	2.56 Å	3.38 Å
Inclination	10° to 20°	-5.9° to -16.4°
Base pair displacement	4.4 to 4.9 Å	-0.2 to 1.8 Å

[1] Data was originally obtained by X-ray crystallography and X-ray fiber diffraction methods and compiled from Ref. 3.
[2] Measurements for random sequence DNA.
[3] Measurements based on van der Waals radii

Figure 1.3. The numbering system for the four bases and ribose ring in RNA.

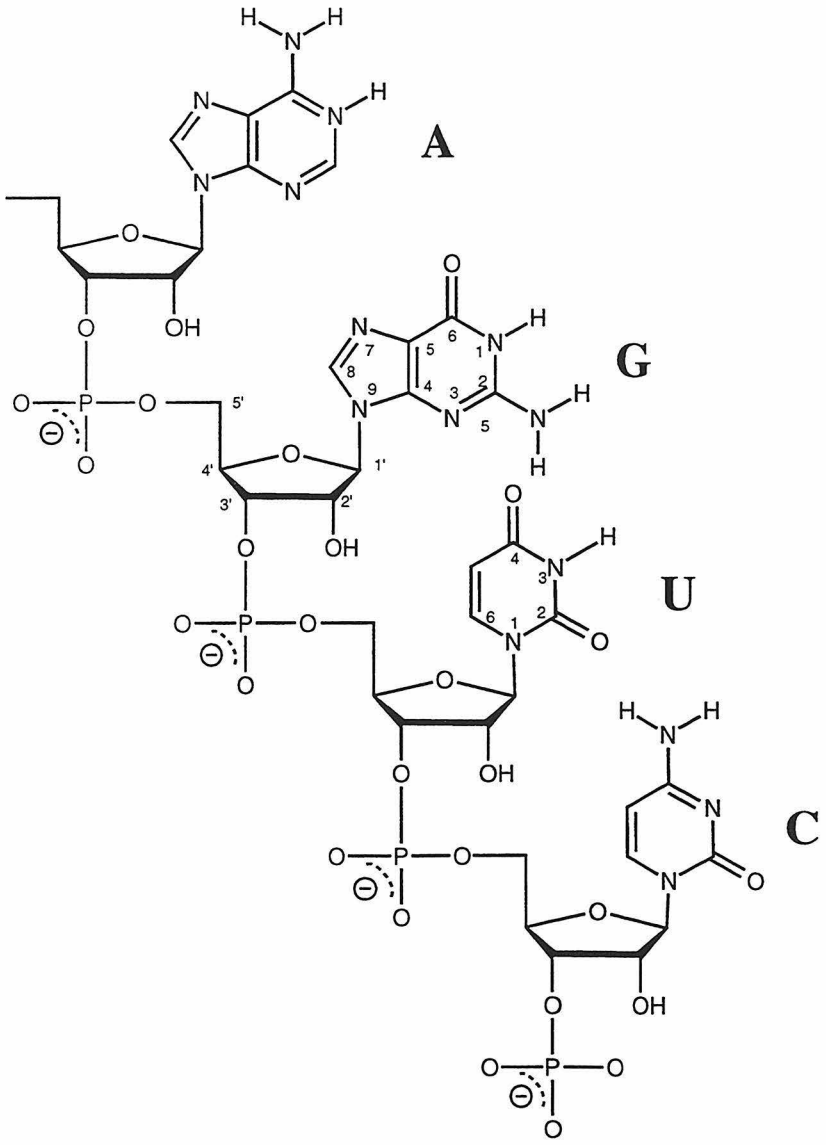
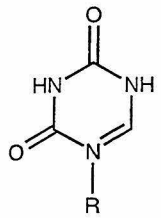
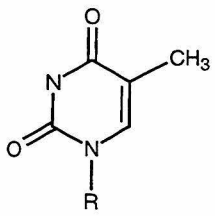

A
G
U
C

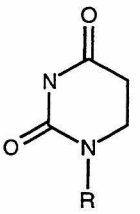
Figure 1.4. A selection of modified nucleosides from transfer RNA.



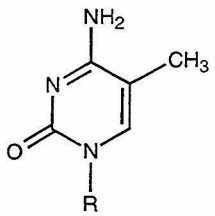
Pseudouridine (Ψ)



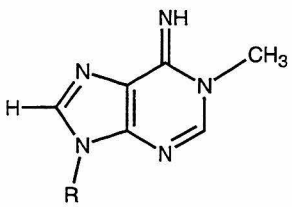
Ribothymidine (rT)



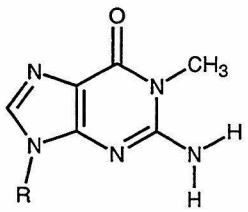
Dihydrouridine (D)



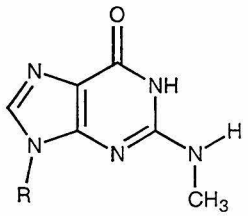
5-Methylcytidine (m^5C)



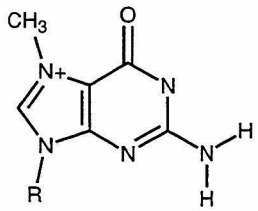
1-Methyladenosine (m^1A)



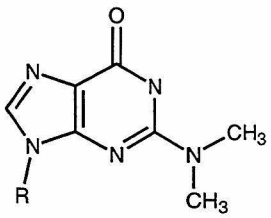
1-Methylguanosine (m^1G)



N^2-Methylguanosine (m^2G)



7-Methylguanosine (m^7G)



N^2,N^2-dimethylguanosine (m_2^2G)

1.3. RNA Secondary Structure

Among the secondary structures in RNA which can be formed are bulges, hairpin loops, base triples and pseudoknot.[11] Some of these features occur within a RNA helix, such as bulges, loops and mismatches. Some can be both intra and intermolecular RNA features, such as pseudoknots,[12] 'kissing' hairpins[13] and base triplets.[3] Examples of these motifs are shown in Figure 1.5.

Most of these structures are stabilized by base pairing, including non-Watson-Crick basepairs which often make use of non-canonical bases. Non-Watson Crick base pairing such as Hoogsteen and reverse Hoogsteen base pairing (Figure 1.6) add to the flexibility of RNA structure. The G-U wobble base is a very common mispair.[14] A single base pair (G3-U70) in the acceptor stem of $tRNA^{Ala}$ is the principle element for specific recognition by *E. coli* alanine-tRNA synthetase.[15] The critical G3 exocyclic 2-amino group makes important functional contacts with 2'-hydroxyl groups near it.[16] Other common mispairs include the G•A mismatch.[17]

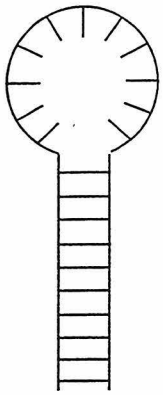
Two hairpin motifs occur in 70% of tetraloops in ribosomal RNAs- the UNCG and the GNRA motifs.[18] Both the UNCG and GNRA motifs actually form biloop structures in solution, as the two bases at the ends of the bulge actually form G•U[19] and G•A[20] mismatches respectively. Both loops are also stabilized by base stacking in the loop and by base-phosphate or base-ribose hydrogen bonding.

1.4. RNA Tertiary Structural Elements

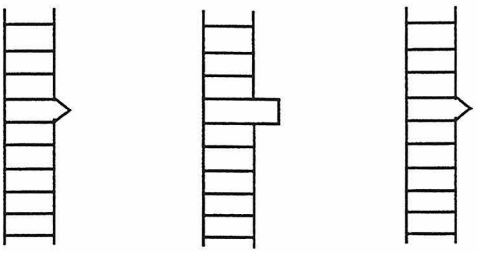
RNA tertiary structure is closely connected to secondary structural elements, as many unique secondary structural elements promote long-range tertiary interactions. Two examples of these tertiary structure elements are loops and base triples.

It has been suggested that GNRA tetraloops may be involved in specific long-range tertiary interactions, with each A in position 3 or 4 of the loop interacting with a C-G base pair in the duplex, and G in position 3 interacting with a U-A base pair.[21] This has been

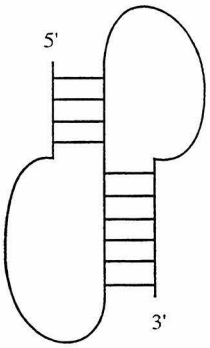
Figure 1.5. A schematic of various secondary structural motifs in RNA. Shown here are (from left to right, top) a stem-loop structure, and a one- and two-base bulge respectively. On the bottom, from left to right, are shown two different pseudoknot structures, and a triple-strand interaction. Adapted from Ref. 11.

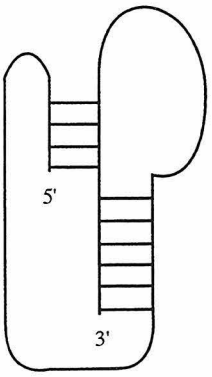


Hairpin

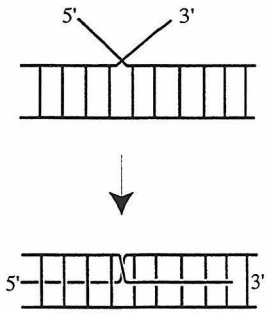


Bulges



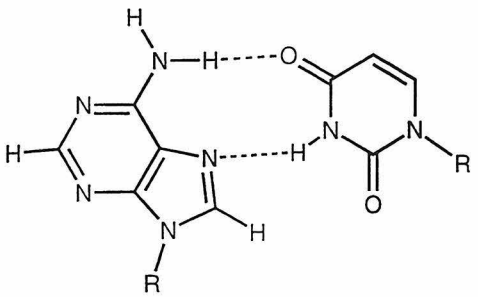


Pseudoknots

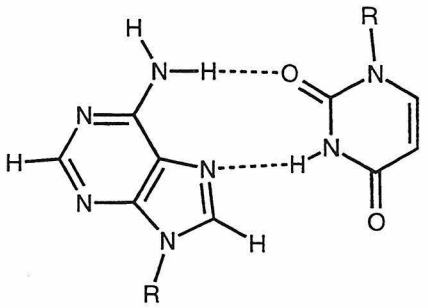


Triplex

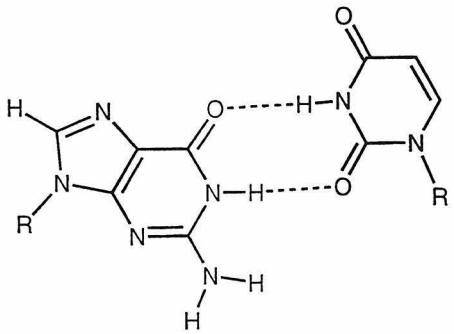
Figure 1.6. Non-Watson-Crick basepairs in RNA. A•U Hoogsteen basepair (top, left) and a A•U reverse Hoogsteen basepair (top, right). On the bottom is shown a G•U wobble basepair (left) and a G•U reverse wobble basepair (right).



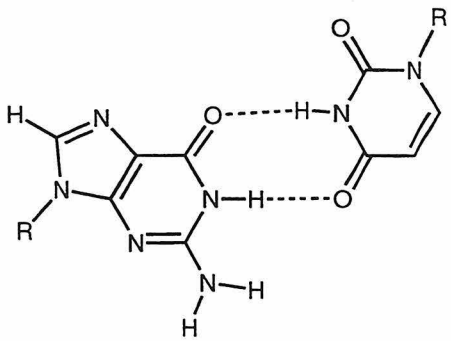
A•U Hoogsteen



A•U Reverse Hoogsteen



G•U Wobble



G•U Reverse Wobble

supported by a recent crystal structure showing this long range interaction,[22] where a GAAA tetraloop interacts with the minor groove of an RNA duplex helix. Many of the interactions observed involve the 2'-hydroxyls of the riboses.

Base triples in RNA can occur when a third strand interacts with the major groove of a duplex.[3] Some base triples are important for the stabilization of the three-dimensional structure of RNA. The base triples m^7G46-[G22-C13], G45-[m^2G10-C25] and A9-[A23-U12] help to generate the structural core of $tRNA^{Phe}$.[23,24] The crystal structure of the P4-P6 domain of the Tetrahymena intron shows various base triples, such as the reverse Hoogsteen basepair U224-A248 interacting with A151.[25]

1.4.1. Role of the 2'-Hydroxyl Group in Stabilizing Tertiary Interactions

The 2'-hydroxyl group is the distinguishing feature on RNA compared to DNA. As a primary structural element, it has effects on both the secondary and tertiary structure of RNA. In addition to forcing the RNA double helix into the A-conformation, the hydroxyl group participates in a range of tertiary interactions. The 2'-OH can form contacts with bases, and RNA tertiary structure can also be stabilized by 2'-OH- phosphate interactions. A good example of the versatility of 2'-OH is on the Tetrahymena intron P4-P6 domain, where it is a ubiquitous donor and acceptor of hydrogen bonds to phosphates and bases, and to other 2'-OH groups.[25]

The use of deoxyribonucleotide substitution in RNA polymers by chemical synthesis permits an evaluation of the role of 2'-hydroxyl groups in the stability of small RNAs. It has been estimated that the interaction of a 2'-OH group contributes on the order of 1 kcal/mol to helix stability on part of the Tetrahymena ribozyme.[26] This method has also been used to study the role of 2'-hydroxyl groups in the hammerhead ribozyme catalysis[27] and other small RNAs. Unfortunately, this technique can not be used for naturally occurring RNAs.

1.4.2. Role of Metal ions in RNA Structure

1.4.2.1. Structural Role of Metal Ions

It has been found that metal ions play vital roles in stabilizing RNA structure.[28] Divalent metal cations, such as Mg^{2+} and Mn^{2+}, in particular, have been found to be crucial for RNA folding and activity.[29] There are four specifically bound Mg^{2+} ions to $tRNA^{Phe}$ in the orthorhombic crystal form[30] and three in the monoclinic form.[31] Several of these magnesium ions are directly coordinated to oxygens of the phosphates. Upon the addition of sub-millimolar concentrations of magnesium ions, the $tRNA^{Phe}$ core undergoes a substantial rearrangement in tertiary structure, passing from an open form with an apparent interstem angle of approximately 150 degrees to a conformation with an angle around 70 degrees.[32] Thus, the magnesium ions are crucial in the correct folding and activity of tRNA molecules.

It has been shown that the Tetrahymena ribozyme is stabilized by Mg^{2+} ions as well.[25] These ions can coordinate the 2'-hydroxyl groups on the RNA and help stabilize the tertiary structure. The tertiary structure of the P4/P6 domain nucleates around a magnesium ion core, with 5 magnesium ions binding at the center of the molecule.[33] The Mg^{2+} ions tie phosphate groups and guanine bases together to provide a platform for specific stacking and hydrogen bonding of the bases.

1.4.2.2. Catalytic Role of Metal Ions

As well as providing a structural role, metal ions also serve a catalytic role in RNA. These roles are often interconnected, with the same metal ion being necessary both for the correct folding and then activity of an RNA molecule.

For example, metal ions such as Mg^{2+} and Mn^{2+} play a crucial role in the mechanism of self-splicing of the Tetrahymena ribozyme.[34] The metal ion contributes directly to catalysis by coordination to the 3' oxygen atom in the transition state.[35] An *in vitro* selection process has been used to engineer variants of the Tetrahymena ribozyme that

are capable of cleaving an RNA substrate in reaction mixtures containing Ca^{2+} as the divalent cation.[36] Similarly, small RNA sequences that bind specifically to zinc, probably relying on one or two direct ion coordinations, have been selected.[37]

The hammerhead ribozyme also shows a dependence on divalent metal ions for self-cleavage.[28] In the absence of added metal ions, the global structure of the hammerhead is extended. Two separate ion-dependent stages are involved in the folding of the hammerhead ribozyme into the active conformation.[38]

Ribozymes with different metal dependencies for catalytic activities have also been engineered, for example, a 'leadzyme', consisting of two helical regions separated by an internal asymmetric loop and dependent on the presence of a lead ion for cleavage at a specific phosphodiester bond.[39]

1.4.5. Role of Water in Stabilizing RNA Structure

Water molecules also play a part in RNA structure by locking the sugar pucker and mediating 3'-->5' intrastrand 02'- 04' hydrogen bonds.[3] The crystal structure of an 8mer RNA duplex shows that the first shell water molecules in the major groove involve half the phosphate oxygens, and join the two strands.[40] In the minor groove, tandem water molecules link the 2'-hydroxyl groups of adjacent nucleotides in base-pair steps. The structure provides evidence for the role of the 2'-hydroxyl groups in the thermodynamic stabilization of RNA. The crystal structure of the domain A of *Thermus flavus* ribosomal 5S RNA shows two U-G and G-U base-pairs which are stabilized by H-bonds supplied via three water molecules to compensate for the lack of base-pair hydrogen bonds, thus showing the importance of highly ordered internal water in stabilizing an RNA structure.[41]

1.5. Specific RNA Molecules and Their Structural Elements

1.5.1. Transfer RNAs

Even small RNA molecules can exhibit a wide variety of motifs. The classical cloverleaf tRNA has been well studied by crystallographical means.[24,42] A crystal structure of $tRNA^{Phe}$ is shown in Figure 1.7. Among the motifs found have been base triplets, G•U wobble base pairs, and phosphate-basepair contacts. These tertiary interactions all contribute to the tight folding of the RNA to produce an active tRNA in aminoacylation and subsequent translation.

1.5.2. Tetrahymena Ribozyme

The Tetrahymena ribozyme is a widely studied RNA molecule. A wide range of interesting motifs such as G•U, A•A and G•A mismatches can be found. A recent crystallographic structure (Figure 1.8) of the P4-P6 domain shows, among other interactions, metal phosphate coordination, base triples, a base quadruple, and the involvement of the 2'-OH in a wide variety of interactions.[25]

1.5.3. Hammerhead Ribozyme

The hammerhead ribozyme is a small catalytic RNA motif made up of three base-paired stems and a core of highly conserved, non-complementary nucleotides essential for catalysis (reviewed in Refs. 43 and 44). A crystal structure of the RNA (Figure 1.9) shows that the base-paired stems are A-form helices and that the core has two structural domains.[45,46] The first domain is formed by the sequence 5'-CUGA following stem I and is a sharp turn identical to the uridine turn of transfer RNA, whereas the second is a non-Watson-Crick three-base-pair duplex with a divalent-ion binding site.

Figure 1.7. Tertiary structure of $tRNA^{Phe}$. Based on the X-ray crystallographical structure from Refs. 23 & 24.

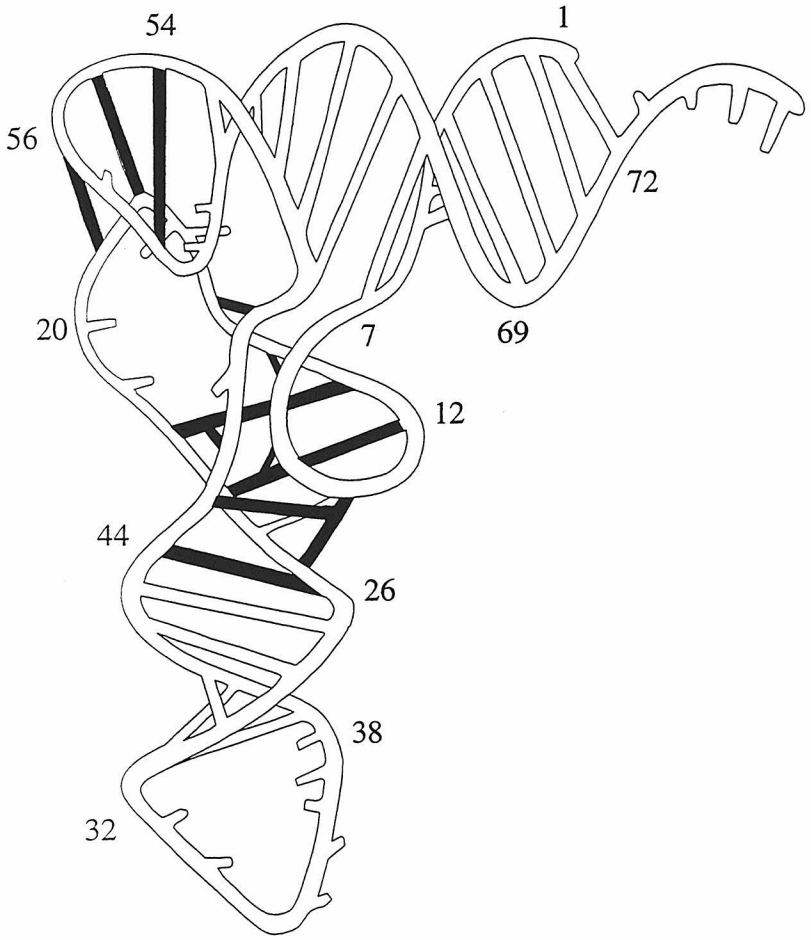
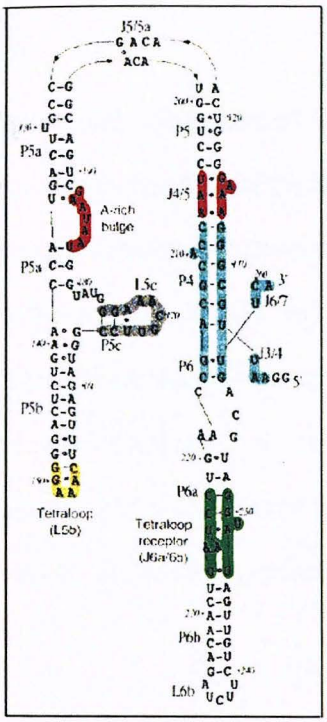

54
1
56
72
20
7
69
12
44
26
38
32

Figure 1.8. Structure of the P4/P6 elements of the Tetrahymena intron. Shown on the left is the secondary sequence and structure, while shown on the right is the corresponding tertiary structure based on X-ray crystallographical coordinates from Ref. 25. Specific contacts occur between the P4 helix (light blue) and the A-rich bulge (orange), and between the tetraloop receptor (green) and the GAAA tetraloop (gold).


J5/5a
P5a
A-rich bulge
P5c
L5c
P4
J6/7
J3/4
P6
P5b
Tetraloop (L5b)
Tetraloop receptor (J6a/5a)
P6a
P6b
L6b

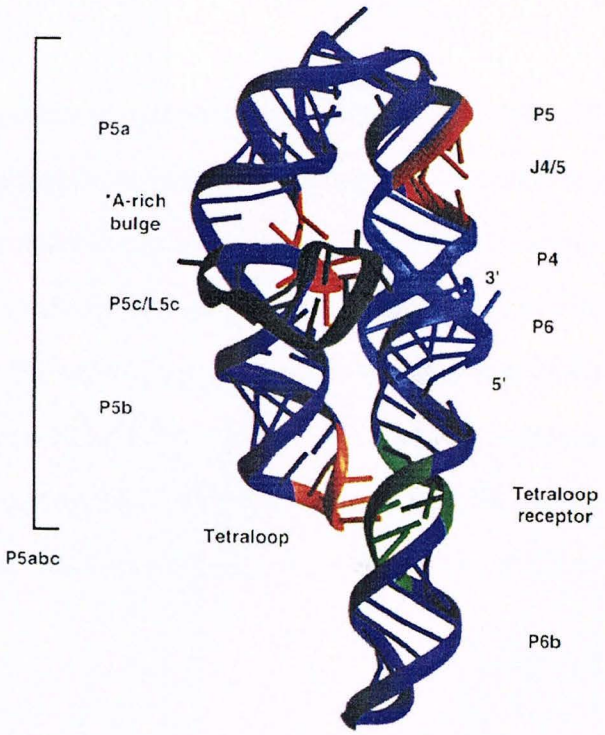
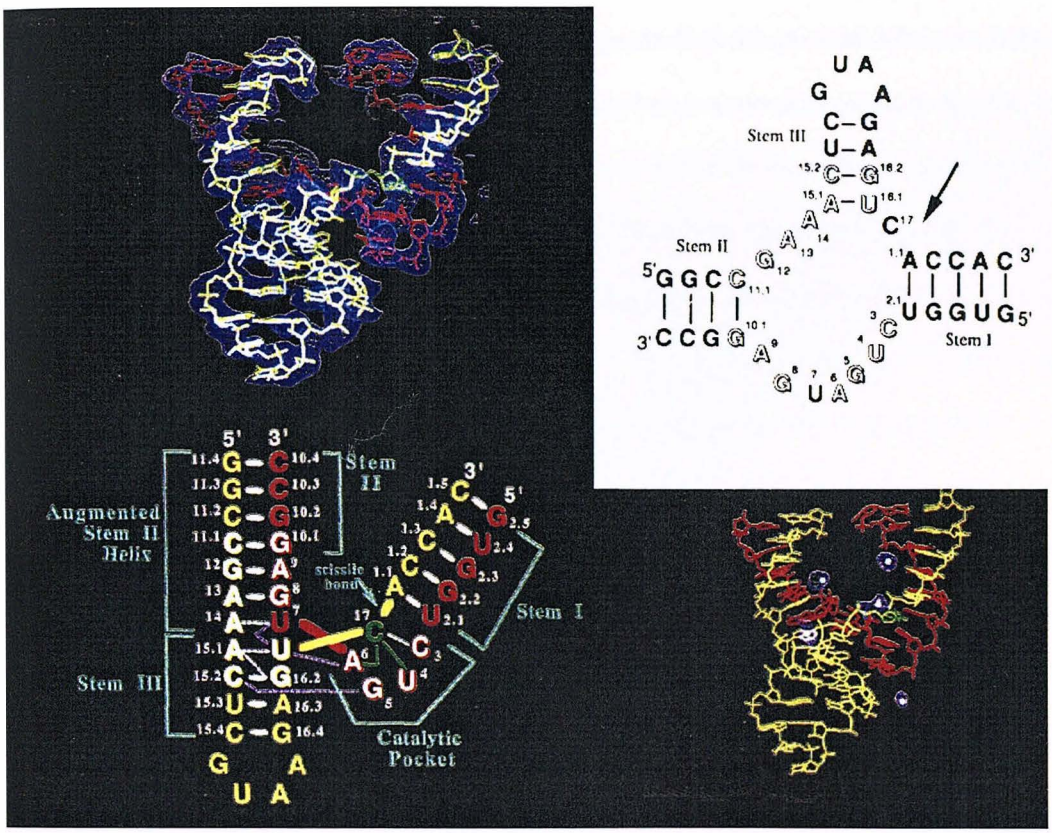

P5a
*A-rich bulge
P5c/L5c
P5b
P5abc
P5
J4/5
P4
3'
P6
5'
Tetraloop receptor
Tetraloop
P6b

Figure 1.9. Structure of the hammerhead ribozyme (adapted from Ref. 46). The secondary structure of the hammerhead (top right) and the noncrystallographic symmetry-averaged Fourier electron density map is shown superimposed on the refined structure of the hammerhead RNA (top left). A color-coded schematic diagram that shows the structural architecture is shown on the bottom left, where the enzyme strand is shown in red and the substrate strand in yellow. The active site cytosine is shown in green. On the bottom right is shown the refined hammerhead ribozyme structure with the same color coding; the refined positions of the five Mg(II) cations are indicated by light blue spheres.


Augmented Stem II Helix
Stem II
Stem III
Stem I
scissile bond
Catalytic Pocket
Stem III
Stem II
Stem I

1.6. RNA-Protein Interactions

RNA structure is also very important in determining RNA-protein interactions. The modes of protein-RNA recognition are complementary to that of protein-DNA recognition, with direct readout, indirect readout, and sequence-dependent deformability being important factors.[47] The various aminoacyl tRNA synthetases (aaRS) complexed with their cognate RNAs provide examples of various modes of protein-RNA recognition. Interactions between the proteins and RNA in both the major and minor groove are known. The crystal structure of yeast AspRS-$tRNA^{Asp}$ shows that the protein interacts with the tRNA from the major groove side.[48] On the other hand, several sequence-specific interactions between GlnRS in the minor groove of $tRNA^{Gln}$ have been observed.[49] A single base pair (G3-U70) in the acceptor stem of $tRNA^{Ala}$ is the principle element for specific recognition by *E. coli* AlaRS,[15] where the protein interacts with the 2-amino group of the guanine in the minor groove of RNA. Thus, synthetases interact with their cognate tRNAs using a range of different strategies, including specific and non-specific interactions.

The NMR solution structures of both the HIV-1 and HIV-2 TAR-argininamide complexes[50,51] have pointed to the importance of both base-triple interactions and amino-acid sidechain-base interactions in the stabilization of RNA structure. In both cases, there are contacts made between the guanidinium side-chain on the argininamide, reminiscent of the targeting of guanine by the guanidinium side chain of arginine in the recognition of DNA by zinc finger proteins.[52] This leads to a change in the RNA structure, involving the formation of an intramolecular U•A-U base triple. The opening of the major groove of the RNA by a 3-base bulge widens the groove enough to allow the RNA-protein recognition to occur.

The recognition of a widened major groove in RNA by proteins appears to be a common feature in RNA-protein recognition. The solution structure of a HIV-1 Rev peptide bound to stem-loop IIB of the Rev response element (RRE) RNA was solved by

NMR spectroscopy.[53] It was found that the α-helical Rev peptide binds in the major groove of the RNA near a purine-rich internal loop. Several arginine side chains make base-specific contacts, and an asparagine residue contacts a G• A basepair. The phosphate backbone adjacent to a G•G base pair adopts an unusual structure that allows the peptide to access a widened major groove. This is necessary in order to provide hydrogen bonding and van der Waals interaction between protein side chains and the exposed edges of the base pairs. It is noteworthy that RNA recognition by proteins has components of both direct and indirect readout. Specific interactions between amino acids and the bases occur, while the RNA molecules also contain elements which accommodate the peptides better.

1.7. Therapeutic Applications of the Understanding of RNA Structure

There is also a need to understand the structure of RNA in the context of the design of new therapeutics. The antisense strategy of gene therapy depends on the interaction of oligonucleotides or oligonucleotide analogs with single stranded messenger RNA to block transcription.[54] Thus, it is necessary to understand both the secondary and tertiary structure of RNA in order to design oligonucleotides which bind specifically to their targets, and are not hindered by any local structures which may form. In addition, the use of ribozymes to target and cleave specific sequences of mRNA depend upon an intimate knowledge of structure and availability of the target sequence.[55]

A knowledge of the structure of RNA can also lead to the design of small molecules which target RNA. Many potent antibiotics target specific RNAs, especially ribosomal RNAs, and halt the process of translation.[56] Aminoglycoside antibiotics interfere with ribosomal protein synthesis and with intron splicing. Various lines of evidence suggest that RNA is the molecular target for various aminoglycosides, but little is known about the recognition process.[57] It has been proposed that some antibiotics target specific sites based on shape recognition, fitting into specific RNA structures. An example of this is the interaction between the aminoglycoside paromomycin and the 16S rRNA.[58]

1.8. Physical Methods for Studying RNA Structure

1.8.1. X-ray Crystallography

X-ray crystallography can produce the highest resolution information on RNA structure.[59] Transfer RNAs proved to be easy to crystallize, and structures of $tRNA^{Phe}$ [24] and $tRNA^{Asp}$ [42] were obtained early on. Other RNA molecules proved more recalcitrant, until fairly recently when a large number of new RNA structures have been solved. One drawback of X-ray crystallography is that the packing and conformation of the RNA in the solid state may differ from that in solution.

1.8.2. NMR Spectroscopy

NMR spectroscopy is the other widely used physical technique for the determination of RNA structure. It has been useful in the determination of the structure of both small RNA molecules and RNA-protein complexes. If a large number of distance and torsion angle constraints are determined, detailed information on the groove width, displacement or base tilt can be obtained.[60] In addition, tertiary interactions, such as base-2'-hydroxyl interactions can be elucidated. Some dynamic structural information such as the rates of base pair openings can be obtained by this method. However, while capable of providing a great deal of information and detail on the structure and dynamics of RNA, NMR spectroscopy is limited to smaller molecules of RNA, below 100 oligonucleotides in size. In addition, the preparation of isotopically enriched ^{13}C and ^{15}N-labeled oligonucleotides is often necessary in order to elucidate the disposition of the bases.[61]

1.8.3. Other Physical Methods

Fluorescence energy resonance transfer (FRET) has been used to obtain a 3-dimensional model for the hammerhead ribozyme[62] which is in agreement with the crystal structure. It has also been used to demonstrate the bending of RNA helices for a series of

double-stranded molecules containing bulge loops of unopposed adenosine.[63] However, this method is unable to provide fine details on the base-pairing and local conformation of the RNA.

Another physical method which has been used to study RNA conformation and interactions with other molecules is IR spectroscopy. FTIR spectroscopy has successfully been used to define the local conformation of RNA,[64] as well as RNA-drug interactions.[65] Here again, however, this is not a very high-resolution method of studying RNA structure.

1.9. Biological Probes of RNA Structure

Nucleases, while not able to provide the detail on RNA structure that NMR and X-ray crystallography do, have been used extensively to probe the structure of large RNA molecules (reviewed in Ref. 66). A variety of nucleases exist, some of which are specific for single stranded RNA such as RNase S1[67], or double stranded regions, such as RNase V1.[68] In addition, there exist base specific RNases such as RNase *Bc* (C- and U-specific).[69] These base-specific RNases are useful in probing differential accessibility of specific bases in folded RNAs, targeting bases which are not involved in double helical regions.

RNase H is a divalent cation-dependent nuclease which is selective for the RNA strand of DNA-RNA hybrids.[70] Hybridization of various DNA probes to an RNA molecule followed by cleavage of the RNA strand with RNase H provides information on which portions of the RNA are single stranded and hence are available for DNA-RNA duplex formation.

1.10. Chemical Probes of RNA Structure

A wide variety of chemical probes have been used for the study of RNA structure. In many cases, these probes carry out the same chemistry on DNA as on RNA. However,

the reactivity of the RNA phosphodiester backbone is different than the DNA backbone owing to the presence of the 2'-OH group. Thus, RNA is much more susceptible to hydrolysis than DNA. In addition, the presence of the uracil instead of the thymine group on RNA leads to slightly different chemistry on the bases.

In addition, the structure of RNA is much more diverse, presenting many more single stranded and accessible regions than duplex DNA. Thus, the process of probing RNA structure is complementary but not identical to the task at hand for DNA.

1.10.1. Phosphodiester Backbone

Fe(II)-ethylenediaminetetraacetic acid ($Fe(EDTA)^{2-}$) has been widely used to probe the structure of RNA . Upon the addition of a reducing agent, the ferrous ions bind EDTA and generate short-lived hydroxyl radicals which promote strand scission of the sugar-phosphate backbone. The hydroxyl radicals cleave at sites which are more solvent-accessible, thus providing information on which regions of a folded RNA molecule are protected. Various derivatives of this complex have also proven to be effective secondary and tertiary structure mapping agents. Methidiumpropyl-EDTA•Fe(II)(MPE•Fe(II) is an intercalator moiety tethered to EDTA.[71] This reagent binds to and cleaves double-helical regions preferentially. MPE•Fe(II) has been used to investigate the structure of Drosophila melanogaster ribosomes.[72] $Fe(EDTA)^{2-}$ has also proven to be a versatile means of site-specifically probing large RNAs, by covalently tethering it to a portion of the RNA. Hydroxyl radicals generated by a tethered Fe(II)-ethylenediaminetetraacetic acid linker to RNA have been used to probe the structure of rRNA.[73] A run-off transcript of yeast $tRNA^{Phe}$ has been synthesized with uridine-EDTA at a single nucleotide position and the resulting set of cleavage fragments upon treatment with ferrous ions and a reducing agent were in general agreement with the three-dimensional structure derived from X-ray analysis.[74]

The active form of bis(1,10-phenanthroline)copper ($Cu(phen)^{2+}$) also depends upon the presence of H_2O_2 to cleave RNA.[75,76] With diffusion, the hydroxyl radicals can then cleave neighboring nucleotides at a distance of up to 1.5 nm to the $Cu(phen)^{2+}$ binding site. $Cu(phen)^{2+}$ cleaves at predominantly single-stranded and loop structures which are not involved in base-pairing interactions on RNA, providing a complement to $Fe(EDTA)^{2-}$ cleavage.

Fe bleomycin has also been used to facilitate oxidative cleavage of tRNAs.[77-80] Damage appears to occur at junctions between single- and double-stranded regions, and many RNA nicks occur at or in proximity to guanosine residues.[78] However, a consensus sequence has not been identified. Cleavage is very specific and occurs at the same site as at its tDNA analog,[81] suggesting that Fe bleomycin is recognizing a specific tertiary structure. It is believed that the bithiazole and C-terminal substituent of bleomycin bind to minor groove structures on the RNA.[80] However, the high site-specificity and lack of cleavage similarity between different related tRNAs means that this reagent can not be used for chemical mapping in the same fashion as more classical reagents. Fe(II)•bleomycin has also been shown to preferentially cleave at RNA duplex/triplex junctions.[82]

The enediyne family of anticancer antibiotics also target RNA, but with no apparent sequence selectivity.[83] Neocarzinostatin, esperamicin and calicheamicin were all observed to cleave tRNA substrates near the 5'-end, and all three compounds exhibited cleavage in single-stranded loop regions. Thus, these compounds are structure but not sequence specific. The high specificity and lack of predictability of these compounds make them less useful for RNA structure probing.

It is possible to carry out RNA cleavage by chemical hydrolysis of the phosphodiester bond. The greater lability of this bond due to the 2'OH of RNA means that this is more easily achieved than the corresponding cleavage on DNA. Derivatives of a 1,4,7,10-tetraazacyclododecane macrocycle which encapsulates trivalent lanthanides have been used to carry out RNA hydrolysis.[84] They cleave RNA by transesterification of the

phosphate diester linkages, but do not appear to be specific.[84] A dinuclear La^{3+} complex hydrolyzes the phosphodiester bond in diribonucleotides efficiently under mild conditions.[85]

Artificial nucleases derived from biological molecules have also proven effective for carrying out hydrolysis. For example, derivatives of guanidinium compounds imitating the bis(arginine) structural motif of staphylococcal nuclease which contain RNA-binding substructures bind to and cleave efficiently the truncated TAR sequence of HIV1.[86]

1.10.2. Base-specific Chemical Agents

There are a large number of guanosine-specific agents which attack the guanine N7 and cleave the RNA polymer upon base treatment. β-ethoxy-α-ketobutyraldehyde (kethoxal) and dimethylsulfate (DMS) both cleave RNA in this fashion and are widely used to probe the accessibility of various regions of folded RNAs. A nickel complex, NiCR, has also been shown to promote the conformation-specific oxidation of guanosine in polynucleotide RNA in the presence of an oxidant, and the reaction was shown to be strictly dependent on the solvent exposure and surface properties of the guanine N7.[87]

Diethylpyrocarbonate (DEPC) carbethoxylates purines, primarily adenines, at the N7 position, while hydrazine and 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate (CMCT) are U-specific.[88] These have proven useful in probing regions such as bulges where the major groove is open and the bases are accessible.[89,90]

1.10.3. Phosphate-specific Probes

Ethylnitrosourea (EtNU) is a reagent which ethylates the phosphates in nucleic acids. This labilizes the phosphodiester bonds of the RNA and thus induces strand scission. The reactivity of phosphates in folded or unfolded RNAs are different, and thus phosphates exposed on the surface of the RNA can be identified. For example, treatment

with EtNU has provided information on the conformational transitions in tRNA,[91] and from these data, the solution structure obtained correlates with the crystal structure.

Pb^{2+} cleaves the phosphodiester bond between U17 and G18 of yeast $tRNA^{Phe}$.[92] It appears that this highly specific cleavage of this RNA molecule is the result of Pb^{2+} hydroxide binding to RNA in such a way that the hydroxide can perform alkaline hydrolysis.[28] Lead cleavage has also been used to probe the core structure of the group I intron RNA for divalent metal ion binding sites.[93]

Psoralen has been used for crosslinking pyrimidines which are adjacent to each other in RNA molecules.[94] It is especially useful for mapping sites which are proximal to each other in the tertiary structure, but do not appear to be so in the secondary structure.

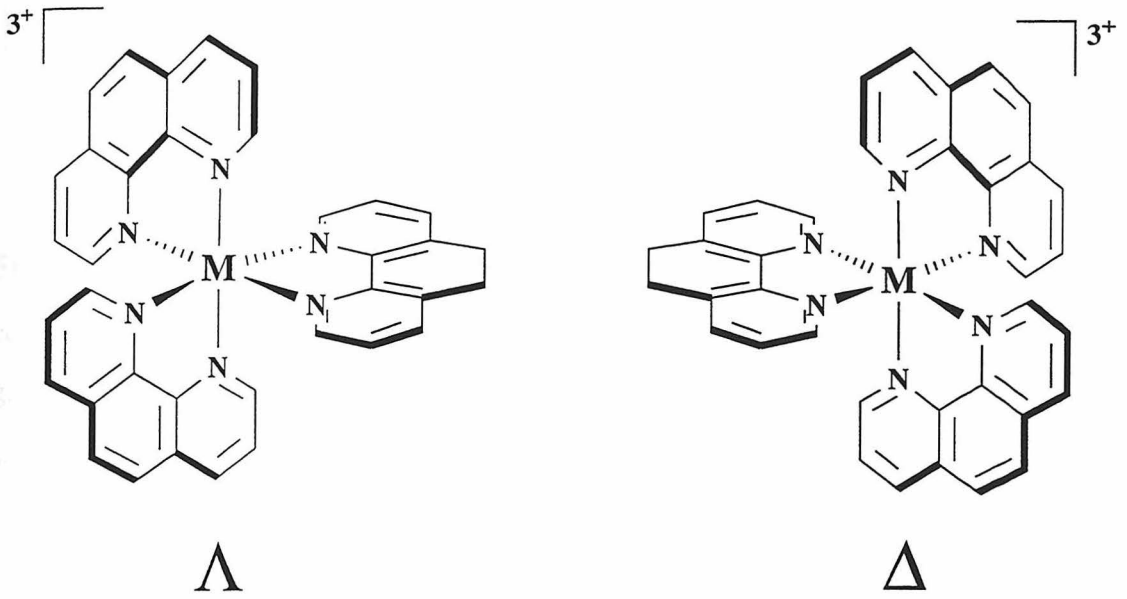
1.11. Transition Metal Complexes as Probes of DNA and RNA Structure

1.11.1. Transition Metal Complexes as DNA Probes

Transition metal complexes have been used to recognize DNA shape-selectively in this laboratory. These metal complexes are derived from the parent tris(phenanthroline) compound (Figure 1.10) and are rigid, octahedral, and allow the appendage of ligands with a wide variety of functional groups on them. The metal centers can be varied to produce an array of complexes with different spectroscopic and reactive characteristics. These complexes may interact with DNA in 3 modes: through electrostatic interaction with the negatively charged phosphate-sugar backbone, through surface binding with the minor groove, and through intercalation in the major groove (Figure 1.11). A selection of molecules is shown in Figure 1.12.

$Rh(phen)_2phi^{3+}$ (phen = phenanthroline, phi = 9,10-phenanthrenequinone diimine) has been shown through NMR studies[95] and biophysical methods to intercalate into duplex DNA from the major groove with high affinity.[96,97] A high resolution NMR structure of a related complex, Δ-α–$Rh[(R,R)\text{-}(Me_2Trien)_2)]phi^{3+}$ (Me_2Trien = 2,9-diamino-4,7-diazadecane), shows this major groove intercalation.[98] Upon photoactivation, the rhodium

Figure 1.10. Schematic of a generic tris(phenanthroline) metal complex.



Tris(phenanthroline) complexes

Figure 1.11. CPK model illustrating the different binding modes of transition metal complexes to B-form DNA (base-pairs in blue, sugar-phosphate backbone in purple). The green $[Ru(bpy)_3]^{2+}$ binds electrostatically to the helix; the Δ-$[Ru(phen)_3]^{2+}$ (yellow) is shown intercalated into the major groove, while its enantiomer Λ-$[Ru(phen)_3]^{2+}$ (red) is surface-bound against the minor groove. Figure taken from Ref. 100.

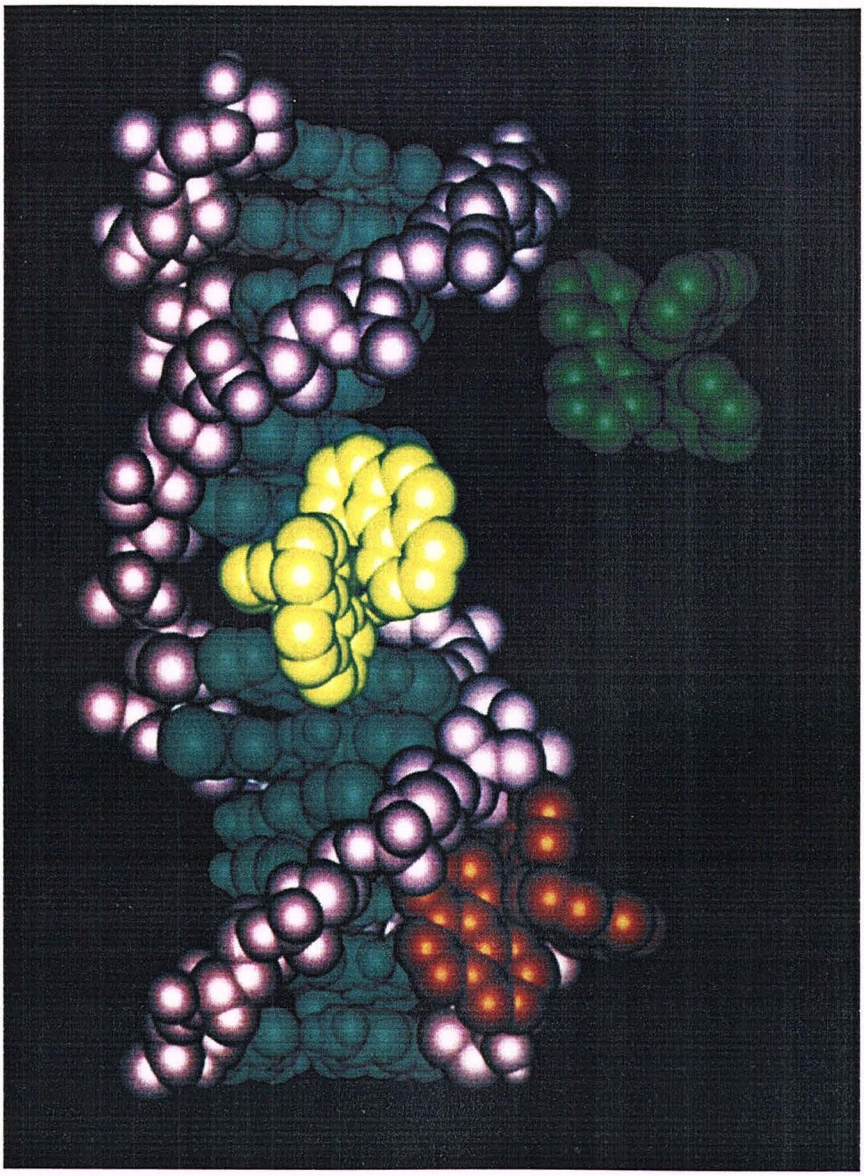
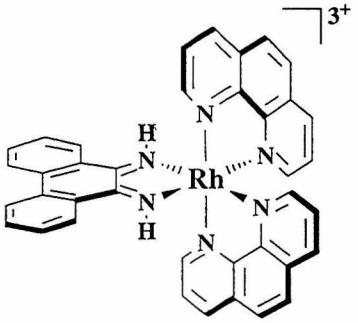
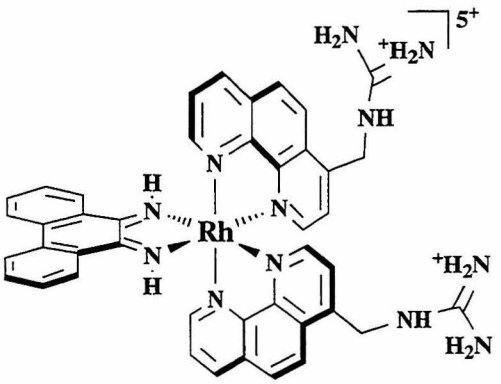


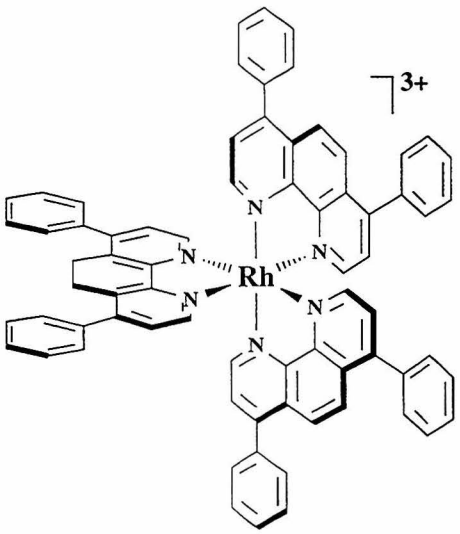
Figure 1.12. Schematic of various rhodium complexes used as structural probes on nucleic acids. Shown starting from the top left, clockwise, is $Rh(phen)_2phi^{3+}$, $Rh(MGP)_2phi^{5+}$, Δ-α-$[Rh(R, R\text{-}MeTrien)_2phi]^{3+}$ and $Rh(DIP)_3^{3+}$.



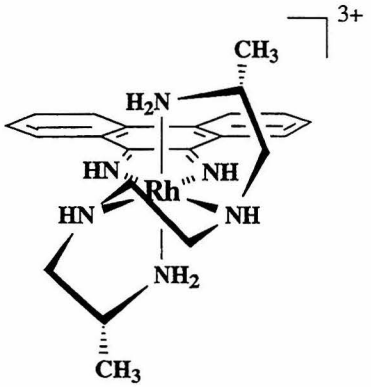
$Rh(phen)_2phi^{3+}$



$3\Delta\text{-}Rh(MGP)_2phi^{5+}$



$[Rh(DIP)]^{3+}$



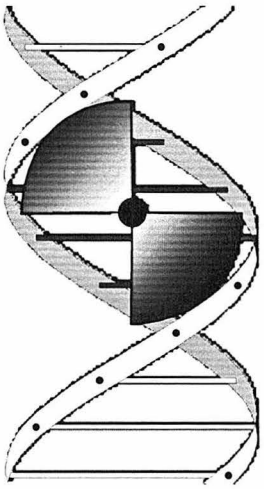
$\Delta\text{-}\alpha\text{-}Rh[(R,R)\text{-}Me_2Trien)_2)]phi^{3+}$

complex promotes direct strand cleavage. Product analysis is consistent with a photoreaction involving direct abstraction of the C3'-hydrogen atom by the photoexcited, intercalated phi ligand.[97] The DNA sites selected for targeting by $Rh(phen)_2phi^{3+}$ correlate closely with sites identified crystallographically as being more open in the major groove, in particular those 5'-pyrimidine-pyrimidine-purine-3' sites which show a large major groove opening due to a change in propeller twist.[99] This preference for open major groove sites has been attributed to a high intercalative binding affinity of the phi ligand being modulated by significant steric clashes of the overhanging phenanthroline ligands with the base paired helix; only at sites which are somewhat open in the major groove is stacking by the octahedral complex made facile.

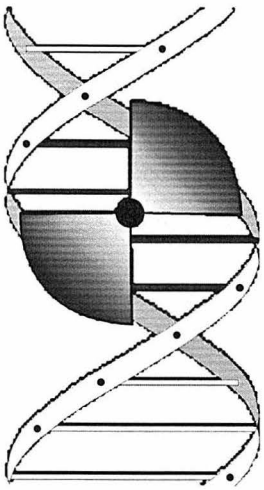
A powerful feature of these metal complexes is that they are chiral and can be resolved into Δ- and Λ-enantiomers. As B-form double helical DNA is a right-handed polymer, the Δ-enantiomer interacts preferentially with the right-handed major groove in DNA,[100] as shown in Figure 1.13 The Λ-enantiomer of $Rh(phen)_2phi^{3+}$ clashes with the backbone of the right-handed DNA, and thus binding is disfavored compared with the Δ-enantiomer.[99] Nevertheless, the Λ-enantiomer of some complexes can bind specifically to certain DNA structures. 1-Λ-$Rh(MGP)_2phi^{5+}$ (MGP=4-guanidylmethyl-1,10-phenanthroline), a derivative of $Rh(phen)_2phi^{3+}$ which has a guanidinium group appended to each phen ligand, recognizes sites with both a favored phi intercalation site and where the DNA is unwound, enhancing contacts between a guanine on the DNA and the guanidinium group on the metal complex.[101]

These metal complexes have been shown to recognize unusual DNA structures. Intron/exon junctions[102] and cruciform regions[103] in DNA are recognized and cleaved photochemically by $Rh(DIP)_3^{3+}$ (DIP = 4,7-diphenyl-1,10-phenanthroline). The exact mode of recognition is not known; however, it appears that this metal complex targets structures which are non-canonical B-form DNA in a very specific manner. Mismatches on

Figure 1.13. Schematic illustrating the interaction of the Δ- and Λ- enantiomers of a metal complex with a right-handed nucleic acid polymer. The Δ-enantiomer fits easily within the right-handed major groove whereas steric interactions are observed between the sugar-phosphate backbone and the ancillary ligands of the Λ-enantiomer, illustrating how shape complementarity can be used as the basis for stereoselectivity in binding.



Λ



Δ

DNA can also be targeted by rhodium complexes. $Rh(TMP)_2phi^{3+}$ (TMP = 3,4,7,8-tetramethyl-9,10-phenanthroline) recognizes GA mismatches preferentially over GT mismatches and the normal GC basepair,[104] and $Rh(bpy)_2chrysi^{3+}$ (bpy = bipyridine, chrysi = 5,6 chrysene quinone diimine) targets certain mismatches which are thermodynamically destabilized.[105]

1.11.2. Transition Metal Complexes as RNA Probes

$Rh(phen)_2phi^{3+}$ may be particularly useful in probing RNA tertiary structure because, despite binding through shape-selection to B-DNA sites, the complex neither binds nor cleaves A-form duplexes.[106] This poor reactivity with RNA duplexes is consistent with the preferential intercalation of the complex in the major groove; since an A-form helix contains a narrow and deep major groove, intercalation by the octahedral complex is precluded.[106-108] Analogous studies on $tRNA^{Phe}$ [108] as the photoactivation experiments on DNA[97] indicate equivalent products and efficiencies, suggesting that to cleave RNA, the phi ligand must be intercalated in a mode which is similar to that on DNA.

With photoactivation, $Ru(phen)_3^{2+}$ targets primarily guanosines in a pathway consistent with singlet oxygen mediation.[107] $Ru(TMP)_3^{2+}$ cleaves at a subset of solvent accessible sites cleaved by $Ru(phen)_3^{2+}$, making it a useful A-form probe, when its cleavage sites are compared to those of $Ru(phen)_3^{2+}$.

$Rh(DIP)_3^{3+}$ has been used to recognize target G-U mismatches very specifically on $tRNA^{Phe}$, $tRNA^{Asp}$ and microhelices.[109] Photoactivated cleavage occurs on the 3' side of the uracil. The difference in recognition of various G-U mismatches appears to reflect the differences in base stacking with the flanking Watson Crick base pairs.

1.11.3. Transition Metal Complexes as Inhibitors of Nucleic Acid-Protein Interactions

Rhodium probes have also been used as DNA intercalators in the presence of protein factors. The complex $Rh(DPB)_2phi^{3+}$ (DPB = 4, 7-diphenylbpy) binds strongly in an enantioselective fashion to its recognition sequence, 5'-d(CTCTAGAG)-3', and inhibits DNA cleavage by the restriction enzyme *Xba*I at that same recognition sequence.[110] It has been shown that $Rh(MGP)_2phi^{5+}$ and Δ-α–$Rh[(R,R)\text{-}(Me_2Trien)_2)]phi^{3+}$ can specifically inhibit transcription by various RNA polymerases.[111] Thus, these metal complexes can specifically inhibit nucleic-acid protein interactions by binding tightly to their recognition sequences. In addition to the binding affinity provided by the intercalating ligand, these complexes can be engineered with functionalities which mimic those on amino acids, thus providing additional specificity and affinity. For example, the guanidinium groups on $Rh(MGP)_2phi^{5+}$ contact guanine groups at its DNA recognition site.[101]

$Rh(phen)_2phi^{3+}$ recognizes sites in RNA which are open in the major groove, thus making use of one of the strategies of protein-RNA recognition. In addition, transition metal complexes can also recognize such features as G-U mismatches and triple base sites, which have been shown to be RNA-protein recognition features. The appendage of amino-acid side-chain functionalities on the ancillary ligands of these complexes could broaden their functionality. Thus, these metal complexes could be used as protein mimics/inhibitors.

This thesis explores the application of transition metal probes to RNA higher-order structure recognition. By probing both synthetic and natural molecules containing a variety of tertiary motifs, a systematic and predictive understanding of the factors involved in RNA recognition by these complexes is sought. With these complexes, we also seek to elucidate the differences and similarities in RNA and DNA secondary and tertiary folding. Finally, we will explore whether site specific targeting by the complexes can be used to inhibit

RNA-binding proteins and peptides, thus leading to insights into the design of better pharmaceuticals.

References

1. Cech, T. R. *Annu. Rev. Biochem.* **1988**, *59*, 543-568.
2. Altman, S. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1989**, *62*, 1-36.
3. Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag:, 1984.
4. McClosky, J. A.; Nishimura, S. *Acc. Chem. Res.* **1977**, *10*, 403-410.
5. Rich, A.; Nordheim, A.; Wang, A. H.-J. *Annu. Rev. Biochem.* **1984**, *53*, 791-846.
6. Klug, A.; Jack, A.; Viswamitra, M. A.; Kennard, O.; Shakked, Z.; Steitz, T. *J. Mol. Biol.* **1979**, *131*, 669-680.
7. Travers, A. *Curr. Biology* **1997**, *7*, R252-R254.
8. Geiger, J. H.; Hahn, S.; Lee, S.; Sigler, P. B. *Science* **1996**, *272*, 830-836.
9. Nikolov, D. B.; Chen, H.; Halay, E. D.; Usheva, A. A.; Hisatake, K.; Lee, D. K.; Roeder, R. G.; Burley, S. K. *Nature* **1995**, *377*, 119-128.
10. Tan, S.; Hunziker, Y.; Sargent, D. F.; Richmond, T. J. *Nature* **1996**, *381*, 127-134.
11. Wyatt, J. R.; Tinoco, I. in *The RNA World* ; Gesteland, R. F. and Atkins, J. F. Eds.; Cold Spring Harbor Laboratory Press: Plainview, N. Y., 1993, pp 465-496.
12. Puglisi, J. D.; Wyatt, J. R.; Tinoco, I. *Acc. Chem. Res.* **1991**, *24*, 152-158.
13. Chang, K. Y.; Tinoco, I. *J. Mol. Biol.* **1997**, *269*, 52-66.
14. Gautheret, D.; Konings, D.; Gutell, R. R. *RNA* **1995**, *1*, 807-814.
15. Hou, Y. M.; Schimmel, P. *Nature* **1988**, *333*, 140-145.
16. Musier-Forsyth, K.; Schimmel, P. *Nature* **1992**, *357*, 513-515.
17. Gautheret, D. F.; Konings, D.; Gutell, R. R. *J. Mol. Biol.* **1994**, *242*, 1-8.

18. Woese, C. R.; Gutell, R.; Gupta, R.; Noller, H. F. *Microbiol. Rev.* **1983**, *47*, 621-669.

19. Cheong, C.; Varani, G.; Tinoco, I. *Nature* **1990**, *346*, 680-682.

20. Heus, H. A.; Pardi, A. *Science* **1991**, *253*, 191-194.

21. Michel, F.; Westhof, E. *J. Mol. Biol.* **1990**, *216*, 585-610.

22. Pley, H. W.; Flaherty, K. M.; Mckay, D. B. *Nature* **1994**, *372*, 111-113.

23. Quigley, G. J.; Rich, A. *Science* **1976**, *194*, 796-806.

24. Kim, S.-H.; Sussman, J. L.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Wang, A. H.; Seeman, N. C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4970-4974.

25. Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K. H.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678-1685.

26. Pyle, A. M.; Cech, T. R. *Nature* **1991**, *350*, 628-631.

27. Perreault, J. P.; Labuda, D.; Usman, N.; Yang, J. H.; Cedergren, R. *Biochemistry* **1991**, *30*, 4020-4025.

28. Pan, T.; Long, D. M.; Ohlenbeck, O. C. in *The RNA World* ; Gesteland, R. F. and Atkins, J. F. Eds.; Cold Spring Harbor Laboratory Press: Plainview, N. Y., 1993, pp 271-302.

29. Pyle, A. M. *Metal Ions Biol. Systems* **1996**, *32*, 479-520.

30. Quigley, G. A.; Tetter, M. M.; A., R. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 64-68.

31. Jack, A.; Ladner, J. E.; Rhodes, D.; Brown, R. S.; Klug, A. *J. Mol. Biol.* **1977**, *111*, 315-318.

32. Friederich, M. W.; Hagerman, P. J. *Biochemistry* **1997**, *36*, 6090-6099.

33. Cate, J. H.; Hanna, R. L.; Doudna, J. A. *Nature Struc. Biol.* **1997**, *4*, 553-.

34. Grosshans, C. A.; Cech, T. R. *Biochemistry* **1989**, *28*, 6888-6894.

35. Piccilli, J. A.; Vyle, J. S.; Caruthers, M. H.; Cech, T. R. *Nature* **1993**, *361*, 85-88.

36. Lehman, N.; Joyce, G. F. *Nature* **1993**, *361*, 182-185.

37. Ciesiolka, J.; Yarus, M. *RNA* **1996**, *2*, 785-793.

38. Bassi, G. S.; Murchie, A. I. H.; Lilley, D. M. J. *RNA* **1996**, *2*, 756-768.

39. Pan, T.; Uhlenbeck, O. C. *Nature* **1992**, *358*, 560-563.

40. Egli, M.; Portmann, S.; Usman, N. *Biochemistry* **1996**, *35*, 8489-8494.

41. Betzel, C.; Lorenz, S.; Furste, J. P.; Bald, R.; Zhang, M.; Schneider, T. R.; Wilson, K. S.; Erdmann, V. A. *FEBS Letts.* **1995**, *351*, 159-164.

42. Westhof, E.; Dumas, P.; Moras, D. *J. Mol. Biol.* **1985**, *184*, 119-145.

43. Birikh, K. R.; Heaton, P. A.; Eckstein, F. *Eur. J. Biochemistry* **1997**, *245*, 1-16.

44. Mckay, D. B. *RNA* **1996**, *2*, 395-403.

45. Pley, H. W.; Flaherty, K. M.; Mckay, D. B. *Nature* **1994**, *372*, 68-74.

46. Scott, W. G.; Finch, J. T.; Klug, A. *Cell* **1995**, *81*, 991-1002.

47 Steitz, T. A. in *The RNA World* ; Gesteland, R. F. and Atkins, J. F.Eds.; Cold Spring Harbor Laboratory Press: Plainview, N. Y., 1993, pp 219-237.

48. Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A.; Podjarny, A.; Rees, B.; Thierry, J. C.; Moras, D. *Science* **1991**, *252*, 1682-1689.

49. Rould, M. A.; Perona, J. J.; Söll, D.; Steitz, T. A. *Science* **1989**, *246*, 1135-1142.

50. Brodsky, A. S.; Williamson, J. R. *J. Mol. Biol.* **1997**, *267*, 624-639.

51. Puglisi, J. D.; Tan, R.; J., C. B.; Frankel, A. D.; Williamson, J. R. *Science* **1992**, *257*, 76-80.

52. Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1992**, *61*, 1053-.

53. Battiste, J. L.; Mao, H. Y.; Rao, N. S.; Tan, R. Y.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R. *Science* **1996**, 1547-1551.

54. Stein, C. A.; Cheng, Y. C. *Science* **1993**, *261*, 1004-1012.

55. Kijima, H.; Ishida, H.; Ohkawa, T.; Kashanisabet, M.; Scanlon, K. J. *Pharmacology & Therapeutics* **1995**, *68*, 247-267.

56. Gale, E. F.; Cundliffe, E.; Reynolds, P. E.; Richmond, M. H.; J., W. M. *The Molecular Basis of Antibiotic Action*; John Wiley & Sons: London, 1981.

57. Wang, Y.; Rando, R. R. *Chemistry & Biology* **1995**, *2*, 281-290.

58. Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. *Science* **1996**, *274*, 1367-1371.

59. Holbrook, S. R.; Kim, S. H. *Biopolymers* **1997**, *44*, 3-21.

60. Schmitz, U.; James, T. L. *Meth. Enzym.* **1995**, *261*, 3-44.

61. Batey, R. T.; Battiste, J. L.; Williamson, J. R. *Meth. Enzym.* **1995**, *261*, 300-322.

62. Tuschl, T.; Gohlke, C.; Jovin, T. M.; Westhof, E.; Eckstein, F. *Science* **1994**, 785-789.

63. Gohlke, C.; Murchie, A. I. H.; Lilley, D. M. J.; Clegg, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11660-11664.

64. Dagneaux, C.; Liquier, J.; Taillandier, E. *Biochemistry* **1995**, *34*, 16618-16623.

65. Neault, J. F.; Tajmirriahi, H. A. *J. Phys. Chem. B* **1997**, *101*, 114-116.

66. Schein, C. H. *Nature Biotech.* **1997**, *15*, 529-536.

67. Wrede, P.; Wurst, R.; Vournakis, J.; Rich, A. *J. Biol. Chem.* **1979**, *254*, 9608-9616.

68. Lowman, H. B.; Draper, D. E. *J. Biol. Chem.* **1986**, *261*, 5396-5403.

69. Kuchino, Y.; Nishimura, S. *Meth. Enzym.* **1989**, *180*, 154-163.

70. Hostomsky, Z.; Hostomska, Z.; Matthews, D. A. *Ribonucleases H*; Cold Spring Harbor Laboratory Press: Plainview, N.Y., 1993.

71. Hertzberg, R. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 313-315.

72. Han, H.; Schepartz, A.; Pellegrini, M.; Dervan, P. B. *Biochemistry* **1994**, *33*, 9831-9844.

73. Heilek, G. M.; Noller, H. F. *Science* **1996**, *272*, 1659-1662.

74. Han, H. Y.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4955-4959.

75. Hermann, T.; Heumann, H. *RNA* **1995**, *1*, 1009-1017.

76. Sigman, D. S.; Landgraf, R.; Perrin, D. M.; Pearson, L. *Metal Ions in Biol. Systems* **1996**, *33*, 485-513.

77. Hecht, S. M. *Bioconjugate Chem.* **1994**, *5*, 513-526.

78. Carter, B. J.; Devroom, E.; Long, E. C.; van der Marel, G. A.; van Boom, J. H.; Hecht, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9373-9377.

79. Battigello, J. M.; Cui, M.; Carter, B. J. *Metal Ions Biol. Systems* **1996**, *33*, 593-617.

80. Holmes, C. E.; Abraham, A. T.; Hecht, S. M.; Florentz, C.; Giege, R. *Nucleic Acids Res.* **1996**, *24*, 3399-3406.

81. Holmes, C. E.; Hecht, S. M. *J. Biol. Chem.* **1993**, *268*, 25909-25913.

82. Kane, S. A.; Hecht, S. M.; Sun, J. S.; Garestier, T.; Helene, C. *Biochemistry* **1995**, *34*, 16715-16724.

83. Battigello, J. M. A.; Cui, M.; Roshong, S.; Carter, B. J. *Bioorg. & Med. Chemistry* **1995**, *3*, 839-849.

84. Morrow, J. R. *Metal Ions Biol. Systems* **1996**, *33*, 561-592.

85. Yashiro, M.; Ishikubo, A.; Komiyama, M. *J. Biochem.* **1996**, *120*, 1067-1069.

86. Kurz, K.; Gobel, M. W. *Helv. Chim. Acta* **1996**, *79*, 1967-1979.

87. Chen, X. Y.; Woodson, S. A.; Burrows, C. J.; Rokita, S. E. *Biochemistry* **1993**, *32*, 7610-7616.

88. Peattie, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1760.

89. Peattie, D. A.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4679-4682.

90. Weeks, K. M.; Crothers, D. M. *Science* **1993**, *261*, 1574-1577.

91. Vlassov, V. V.; Giegé, R.; Ebel, J. P. *FEBS Lett.* **1980**, *120*, 12-16.

92. Dirheimer, G.; Ebel, J. P.; Bonnet, J.; Gangloff, J.; Keith, G.; Krebs, B.; Kuntzel, B.; Roy, A.; Weissenbach, J.; Werner, C. *Biochimie* **1972**, *54*, 127-144.

93. Streicher, B.; Vonahsen, U.; Schroeder, R. *Nucleic Acids Res.* **1993**, *21*, 311-317.

94. Lipson, S. E.; Hearst, J. E. *Meth. Enzymology* **1988**, *164*, 330-341.

95. David, S. S.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 2984-2985.

96. Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520-4522.

97. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303-2312.

98. Hudson, B. P.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1995**, *117*, 9379-9380.

99. Campisi, D.; Morii, T.; Barton, J. K. *Biochemistry* **1994**, *33*, 4130-4139.

100. Barton, J. K. *Science* **1986**, *233*, 727-734.

101. Terbrueggen, R. H.; Barton, J. K. *Biochemistry* **1995**, *34*, 8227-8234.

102. Lee, I.; Barton, J. K. *Biochemistry* **1993**, *32*, 6121-6127.

103. Kirshenbaum, M. R.; Tribolet, R.; Barton, J. K. *Abstracts Papers Of the American Chemical Society* **1988**, *196*, 92-INOR.

104. Campisi, D., Doctoral Thesis, Dept. of Chemistry, California Institute of Technology: Pasadena, 1995.

105. Jackson, B. A.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, submitted.

106. Chow, C. S.; Hartmann, K. M.; Rawlings, S. L.; Huber, P. W.; Barton, J. K. *Biochemistry* **1992**, *31*, 3534-3542.

107. Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839-2841.

108. Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972-982.

109. Chow, C. S.; Barton, J. K. *Biochemistry* **1992**, *31*, 5423-5429.

110. Sitlani, A.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12589-12590.

111. Johann, T., Doctoral Thesis, Dept. of Chemistry, California Institute of Technology: Pasadena, 1997.

Chapter 2:

Chemical Probing of $tDNA^{Phe}$ with Transition Metal Complexes: A Structural Comparison of RNA and DNA‡

2.1. Introduction

Substantial interest is focused on understanding the three-dimensional folding of RNAs. This understanding is essential in delineating RNA recognition and reactivity. Despite recent advances in synthetic methodology,[1-5] RNA synthesis, either chemically or enzymatically, is not easily accomplished in pure form. In addition, RNA is prone to degradation by nucleases, and few structural probes are available to characterize the products once obtained. In contrast, stable DNA polymers may be prepared in high yield, and both enzymatic and chemical probes are available for characterization. Given these considerations, many laboratories have begun studies of DNA analogs of RNA polymers,[6,7] or of deoxyribonucleotide polymers containing ribonucleotides at selected positions.[8,9]

But how similar is DNA to RNA, structurally and functionally? Studies on the hammerhead ribozyme have shown that the deoxy analog of a substrate oligonucleotide bearing only a ribonucleotide at the cleavage site is active.[10] The all deoxy hammerhead domain is inactive, but an oligodeoxynucleotide containing as few as four ribonucleotides in the catalytic domain shows activity.[9] In addition, an all-DNA external guide sequence can make a potential substrate susceptible to cleavage by RNase P.[8] Thus, where the 2'-OH is not directly involved in the chemistry of the cleavage site, the secondary and tertiary structure must be conserved to some degree. It has also been demonstrated that tDNAs[1] may be aminoacylated by their cognate aminoacyl synthetases with the correct amino acids in the three cases (phenylalanine, lysine and methionine) tried thus far, as

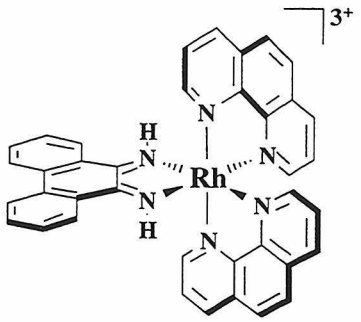
---

‡ Adapted from A. Lim; J. K. Barton *Biochemistry* **1993**, 11029-11034.

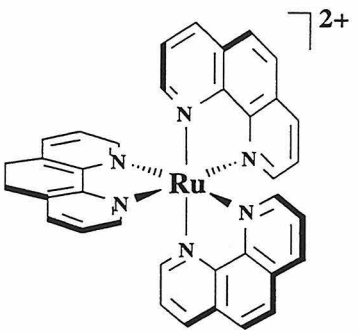
[1]. tDNAs are synthetic DNA oligomers, analogues of particular full-length tRNA, with dT substituted in the positions occupied by ribouridine or its derivatives. In addition, unmodified bases are used instead of modified ones in the case of the tDNA studied in our laboratory.

long as the 3' -riboadenosine is retained.[6,7] Thus the ribose backbone is not required for aminoacylation. Kinetic analysis of aminoacylation of mutant yeast $tRNA^{Phe}$ transcripts[11] has shown that the G20, G34, A35, A36 and G73 nucleotides are the major recognition sites; these five nucleotides must, therefore, still be positioned in essentially the correct positions for aminoacylation in the deoxynucleotide analog. It has also been established through studies utilizing mung bean nuclease and the restriction endonucleases *Hha*I and *Cfo*I[12] that the classic cloverleaf structure is retained in these tRNA analogs.

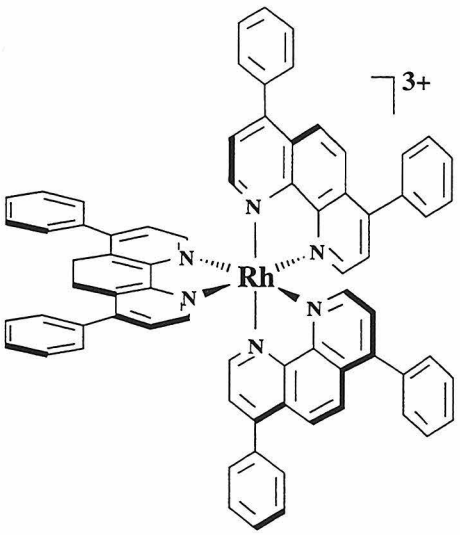
In view of the wealth of knowledge about $tRNA^{Phe}$ structure[13,14] and its interactions with its cognate synthetase, we decided to examine in greater structural detail the parallels between $tRNA^{Phe}$ and a chemically synthesized $tDNA^{Phe}$ using a series of chemical probes for local nucleic acid structure. In our laboratory, transition metal complexes have been designed based upon *shape selection* as *local* structural probes for both DNA and RNA.[15,16] Furthermore, the coupling of photoactivated strand cleavage to this recognition provides a sensitive means to mark structurally distinct regions with single nucleotide resolution. Figure 2.1 illustrates the shape-selective probes bis(phenanthroline)(9,10-phenanthrenequinone diimine)rhodium(III) [$Rh(phen)_2phi^{3+}$], tris(4,7-diphenyl-1,10-phenanthroline)rhodium(III) [$Rh(DIP)_3^{3+}$], and tris(3,4,7,8-tetramethylphenanthroline)ruthenium(II) [$Ru(TMP)_3^{2+}$]. $Rh(phen)_2phi^{3+}$ has been shown to intercalate in the major groove of double helical DNA at 5'-pyr-pyr-pur-3' steps[17-20] and consistent with this recognition, to target sites of tertiary interaction in folded RNAs.[21-23] $Rh(DIP)_3^{3+}$ targets unusual DNA tertiary structures, such as cruciforms, Z-DNA, and Holliday junctions,[24,25] while targeting primarily GU mismatches in RNAs.[26] $Ru(TMP)_3^{2+}$ binds against the shallow minor groove of RNA and DNA double helices which are A-like in character.[21,27,28]



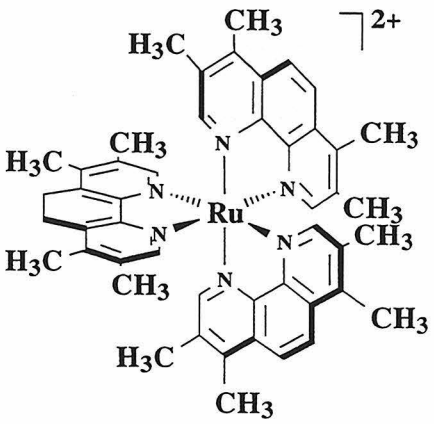
$Rh(phen)_2phi^{3+}$



$[Ru(phen)]^{2+}$



$[Rh(DIP)]^{3+}$



$[Ru(TMP)]^{2+}$

Figure 2.1. Schematic illustration of the shape-selective probes $Rh(phen)_2phi^{3+}$, $Rh(DIP)_3^{3+}$ and $Ru(TMP)_3^{2+}$.

Here we apply these different shape-selective probes to compare and contrast the structure of $tDNA^{Phe}$ and $tRNA^{Phe}$. In general we find the overall folding of the tDNA to resemble that of tRNA, but local variations and potential recognition elements differ. In particular, double helical regions of the molecule differ in that these regions adopt primarily the B-conformation.

2.2. Experimental

Materials. Native yeast $tRNA^{Phe}$ from brewer's yeast (Boehringer Mannheim) was 3'-end-labeled with cytidine 3',5'-[5'-^{32}P]-bisphosphate using T4 RNA ligase. It was then gel purified on a 10% denaturing polyacrylamide gel, located by autoradiography, excised, and eluted from the gel slice in 45 mM Tris, 45 mM boric acid, and 1.25 mM EDTA, pH 8.0 The eluted $tRNA^{Phe}$ was ethanol-precipitated twice and stored frozen in 10 mM Tris-HCl, pH 7.5. $tDNA^{Phe}$ was chemically synthesized on a 1 μM scale on an Applied Biosystems 392 DNA/RNA synthesizer using the phosphoramidite method. Unusual bases were not incorporated into the polymer. The oligonucleotide was then purified twice by HPLC, first with the dimethoxytrityl (DMT) group on and subsequently with the DMT group off using a C_{18} column (Dynamax). The oligonucleotide was then purified on a 10% denaturing polyacrylamide gel, located by UV shadowing, excised, and eluted from the gel in 45 mM Tris, 45 mM boric acid and 1.25 mM EDTA, pH 8.0. The eluted tDNA was concentrated by a Centricon 10 (Amicon) device, desalted by washing twice with water, and stored in 10 mM Tris-HCl, pH 7.5. The purified oligonucleotide has a UV maximum at 257 nm and an extinction coefficient at 260 nm of 8500 $M^{-1}cm^{-1}$ /nucleotide. The tDNA was then 3'-end-labeled with [α-^{32}P]-ddATP using terminal deoxytransferase or 5'-end-labeled with [γ-^{32}P]-ATP using T4 polynucleotide kinase. The labeled material was gel-purified by the same method as for $tRNA^{Phe}$ and stored frozen in 10 mM Tris-HCl, pH 7.5. $[Rh(phen)_2phi]Cl_3$, $[Rh(DIP)_3]Cl_3$, $[Ru(TMP)_3]Cl_2$ and $[Ru(phen)_3]Cl_2$ were prepared in our laboratory by

published procedures.[15] All metal stock solutions were freshly prepared in either ethanol or 10 mM Tris-HCl, pH 7.5.

Photocleavage Reactions. The end-labeled $tRNA^{Phe}$ and $tDNA^{Phe}$ were renatured by heating to 70°C for 1 min in 10 mM Tris-HCl, pH 7.5 (10 mM $MgCl_2$ optional), and slowly cooling to room temperature prior to use. A typical 20 μL cleavage mixture contained labeled RNA or DNA, 2.5-10 μM metal complex (freshly diluted in H_2O), the appropriate buffer (50 mM Tris, 20 mM NaOAc, 18 mM NaCl, pH 7.0, or 5 mM Tris, 50 mM NaCl, pH 7.0; 10 mM $MgCl_2$ optional to both buffers) and was brought to a final concentration of 100 μM in nucleotides with either carrier $tRNA^{Phe}$ or $tDNA^{Phe}$. The mixture was incubated for a maximum of 5 minutes at room temperature, and was then irradiated at 365 nm in the case of the $Rh(phen)_2phi^{3+}$ mixtures and 313 nm for the $Rh(DIP)_3^{3+}$ mixtures at ambient temperature using a 1000-W Hg/Xe lamp and monochrometer (Oriel model 77250). The $Ru(TMP)_3^{2+}$ and $Ru(phen)_3^{2+}$ mixtures were irradiated at 442 nm using a He-Cd laser (Liconix model 4200 NB) at 19 mW for 10 to 30 minutes at ambient temperature. The reaction mixtures were ethanol precipitated and washed at least three times with ethanol to remove buffer salts, and then dried on a SpeedVac Concentrator (Savant). In addition, the dry ruthenium complex mixtures were treated with either 1M piperidine in H_2O (for DNA) at 90°C or 1M aniline (for RNA) at 60°C for 30 minutes to reveal sites of modification.

Sequencing Gels. The cleavage products were analyzed on 20% polyacrylamide-8M urea gels and viewed by autoradiography. The full-length $tRNA^{Phe}$ and cleavage products were identified by coelectrophoresing with $Ru(phen)_3^{2+}$ (G-specific) reactions,[21] diethyl pyrocarbonate (DEPC) (A-specific) and hydrazine (U-specific) reactions,[29] and the full-length $tDNA^{Phe}$ and cleavage products by piperidine formate (A+ G-specific) and hydrazine hydrate (C+T-specific) Maxam Gilbert reactions.[30] The fragments produced by the metal complex cleavage possess 3' and 5' phosphate termini, and thus could be directly compared with the chemical sequencing lanes.

2.3. Results

2.3.1. Cleavage of $tDNA^{Phe}$ by $Rh(phen)_2phi^{3+}$

Cleavage sites induced by $Rh(phen)_2phi^{3+}$ on the folded tDNA were determined using gel electrophoresis after reaction with tDNA labeled either at the 3'- or 5'-terminus. Specific sites of cleavage are observed on $tDNA^{Phe}$ which are similar but not identical to those found upon irradiation of $tRNA^{Phe}$ with $Rh(phen)_2phi^{3+}$.[23] Figure 2.2 shows cleavage of both 3'-labeled $tDNA^{Phe}$ and $tRNA^{Phe}$. At a $Rh(phen)_2phi^{3+}$ concentration of 5 μM after 5 minutes of irradiation, several strong sites are visible on $tDNA^{Phe}$. Strongest cleavage occurs at residues C28, C48 and C49. Rhodium-induced cleavage is also seen at A29, C40, T50 and C60, and weak cleavage sites may also be discerned at C32, G45, G46, T55, C56, T59, C61, T69 and C70. Identical sites of cleavage are apparent in experiments conducted with 5'-labeled tDNA, as shown in Figure 2.3, where, in addition, another strong site is seen at residue A9. Very weak cleavage is seen at C13 and T17. When cleavage experiments were conducted on denatured tDNA (heated to 90°C for 5 minutes prior to irradiation, and irradiated at that temperature), no cleavage is seen. This observation is consistent with the requirement of the folded structure for recognition and cleavage by the metal complex. No difference in cleavage patterns was observed on tDNA upon the addition of 10 mM $MgCl_2$ to the buffer as seen in Figure 2.4, in contrast to the loss of cleavage observed at tertiary sites in the case of $tRNA^{Phe}$.[23] It is also noteworthy that shorter irradiation times are generally required for the cleavage of $tDNA^{Phe}$ compared to $tRNA^{Phe}$, implying that the DNA is cleaved more efficiently than the RNA.

As illustrated in Table 2.1 and Figure 2.5, the specific sites obtained on $tDNA^{Phe}$ may be contrasted to those sites targeted by the rhodium complex on the native tRNA, also shown in the gel, as well as to those sites found on the $tRNA^{Phe}$ transcript, which, like the tDNA analog, lacks modified bases.[23] It is apparent that the same regions of the molecule are targeted on the tRNA and tDNA analogs. However the relative intensity of

Figure 2.2. Autoradiogram showing cleavage of 3'-^{32}P-labeled $tDNA^{Phe}$ (lanes 1-5) and $tRNA^{Phe}$ (lanes 6-11) by $Rh(phen)_2phi^{3+}$ in 50 mM Tris, 20 mM NaOAc, 18 mM NaCl, pH 7.0. Lanes 1 and 2: A+G and C+T reactions on $tDNA^{Phe}$ respectively. Lane 3: cleaved $tDNA^{Phe}$ after incubation with 5 μM $Rh(phen)_2phi^{3+}$ and irradiation for 4' at 365 nm. Lane 4: labeled DNA with 5 μM $Rh(phen)_2phi^{3+}$ but without irradiation. Lane 5: labeled DNA irradiated for 4' at 365 nm in the absence of metal complex. Lanes 6, 7 and 8: A-, G- and U-specific reactions on labeled $tRNA^{Phe}$ respectively. Lane 9: cleaved $tRNA^{Phe}$ after incubation with 5 μM $Rh(phen)_2phi^{3+}$ and irradiation for 10' at 365 nm. Lane 10: labeled RNA with 5 μM $Rh(phen)_2phi^{3+}$ but without irradiation. Lane 11: labeled RNA after irradiation for 10' at 365 nm in the absence of metal complex.

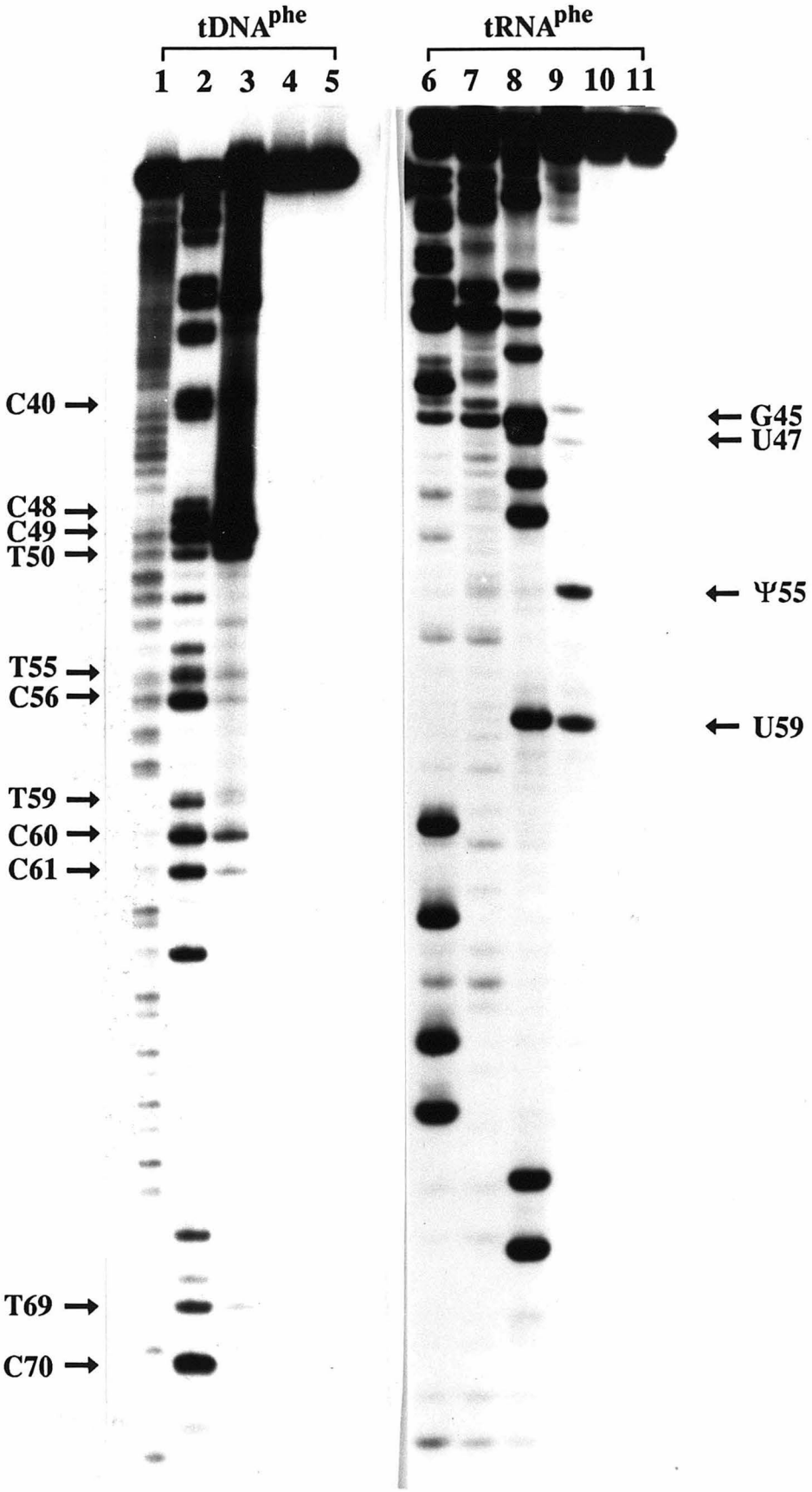

tDNA^phe
tRNA^phe
1 2 3 4 5
6 7 8 9 10 11
C40
C48
C49
T50
T55
C56
T59
C60
C61
T69
C70
G45
U47
Ψ55
U59

Figure 2.3. Greyscale representation of an autoradiogram showing cleavage of 5'-^{32}P-labeled $tDNA^{Phe}$ by $Rh(phen)_2phi^{3+}$ in 50 mM Tris, 20 mM NaOAc, 18 mM NaCl, pH 7.0. Lane 1: A+G reaction on $tDNA^{Phe}$. Lane 2: cleaved $tDNA^{Phe}$ after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 5' at 365 nm. Lane 3: labeled DNA with 10 μM $Rh(phen)_2phi^{3+}$ but without irradiation. Lane 4: labeled DNA irradiated for 5' at 365 nm in the absence of metal complex.

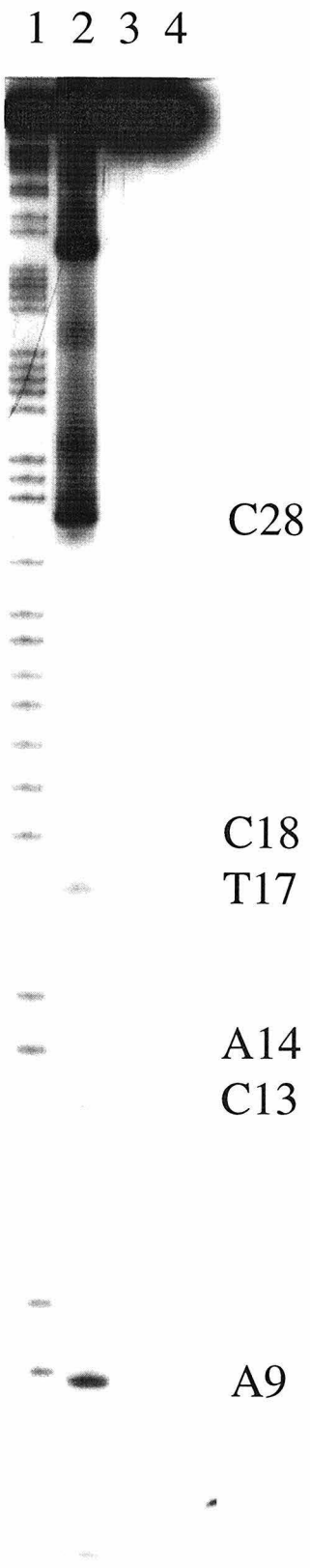

1 2 3 4
C28
C18
T17
A14
C13
A9

Figure 2.4. Greyscale representation of an autoradiogram showing the effect of magnesium ion on $Rh(phen)_2phi^{3+}$ photocleavage of 3'-^{32}P-labeled $tDNA^{Phe}$. Samples were run in 50 mM Tris HCl, 20 mM NaOAc, 18 mM NaCl, pH 7.0. Lanes 1 & 2: A+G and C+T Maxam Gilbert sequencing lanes. Lanes 3 & 4: 5 μM $Rh(phen)_2phi^{3+}$, 2' and 5' irradiation respectively at 313 nm. Lane 5: 5 μM $Rh(phen)_2phi^{3+}$. Lane 6: 5' irradiation at 313 nm. Lane 6: control. Lanes 7 - 12: identical to lanes 3-6, but with the addition of 10 mM $MgCl_2$ in the buffer.

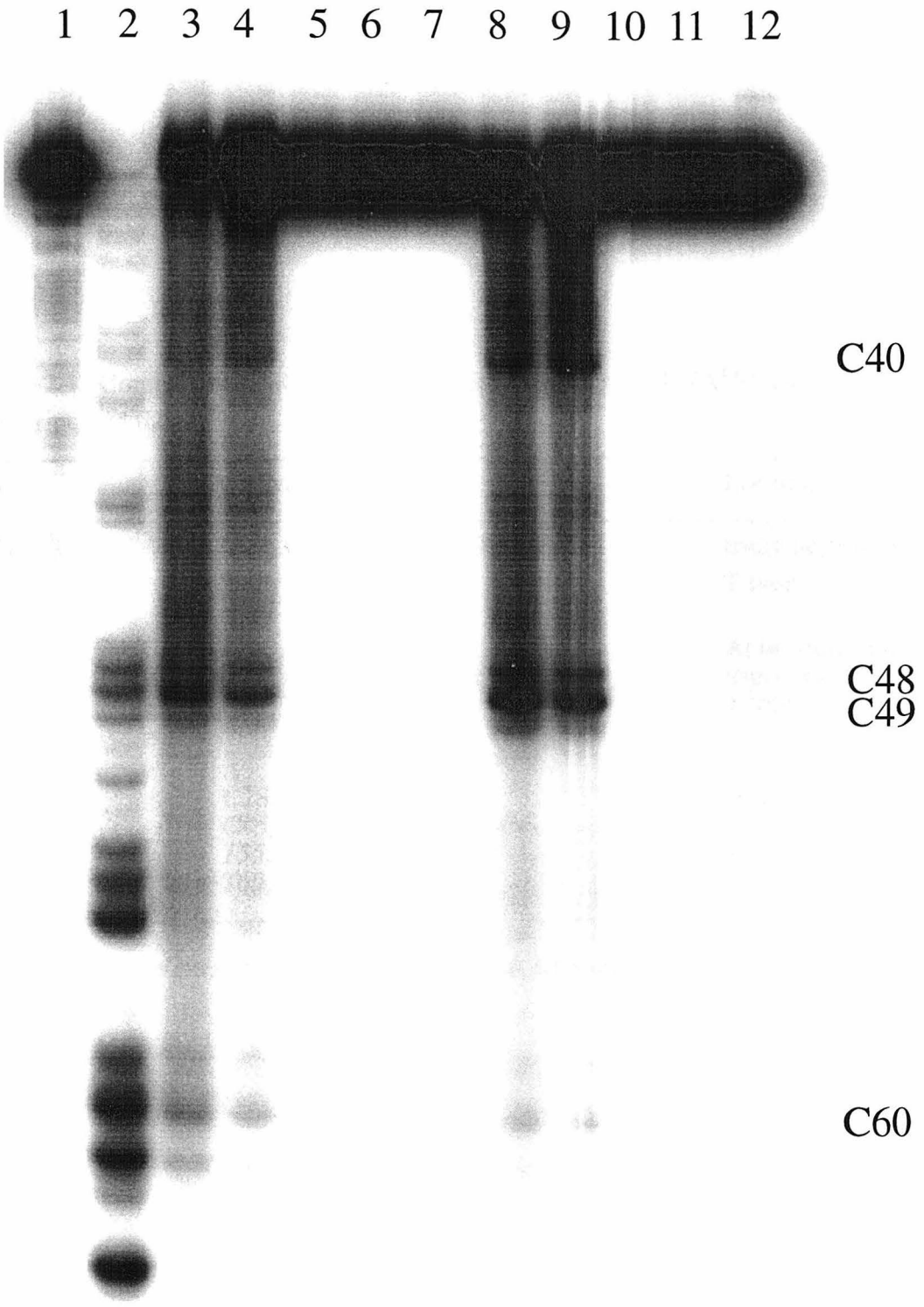
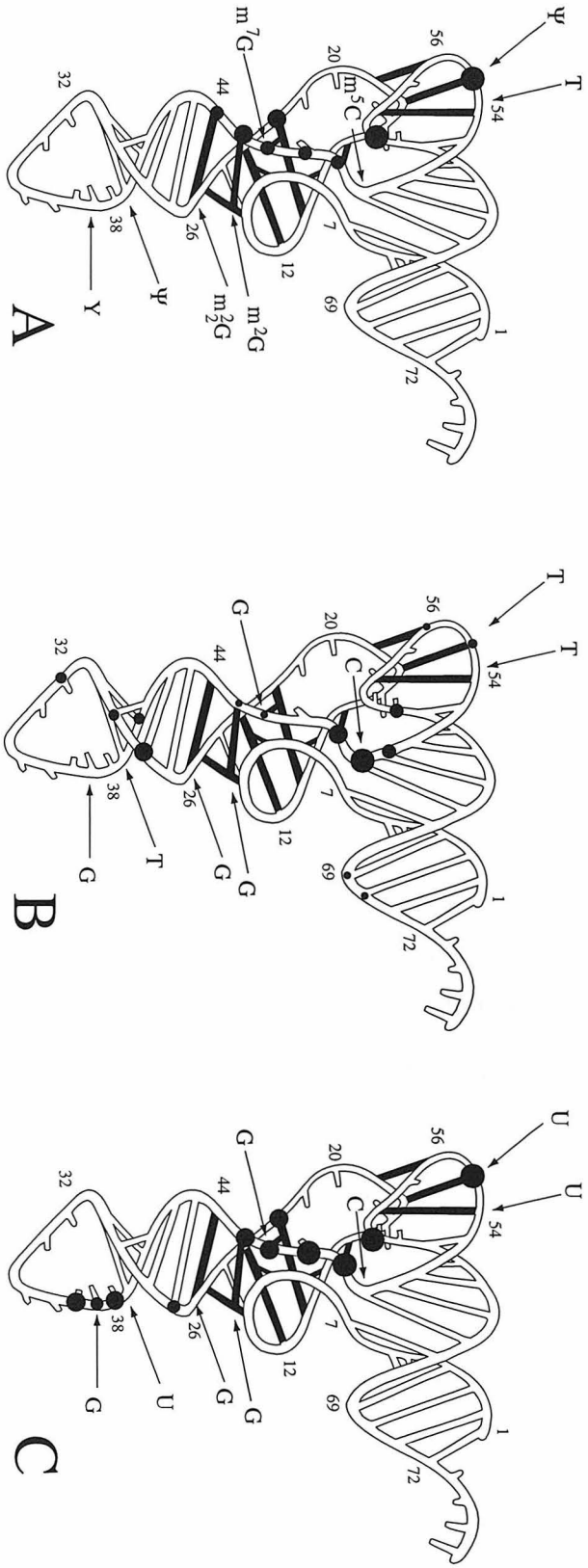

1 2 3 4 5 6 7 8 9 10 11 12
C40
C48
C49
C60

Table 2.1. Comparison of $Rh(phen)_2phi^{3+}$ cleavage sites on $tRNA^{Phe}$, the $tRNA^{Phe}$ transcript, and $tDNA^{Phe}$.

Species	Cleavage sites[a]	Location
$tRNA^{Phe}$	**G22**, **G45**, *m^7G46*, **U47**, *C48*	triply bonded core
	Ψ55, **U59**	T loop
$tRNA^{Phe}$ trans.	*C27*, *A36*, *G37*, A38	Anticodon stem/loop
	G22, *G45*, *G46*, **U47**, **C48**	triply bonded core
	U55, **U59**	T loop
$tDNA^{Phe}$	A9	D loop
	C28, A29, *C32*, C40	Anticodon stem/loop
	G45, G46, **C48**, **C49**, T50	triply bonded core
	T55, *T56, T59*, C60, *C61*	T loop
	T69, *C70*	Acceptor stem

[a] Strong cleavage sites are in boldface; very weak sites are italicized.

Figure 2.5. Ribbon diagrams showing cleavage sites on $tRNA^{Phe}$ (**A**), $tDNA^{Phe}$ (**B**), and $tRNA^{Phe}$ transcript (**C**) by $Rh(phen)_2phi^{3+}$. The solid circles indicate the sites of cleavage, with the size reflecting the relative cleavage intensity. Arrows indicate the bases which are modified in the native $tRNA^{Phe}$.


A
Ψ
T
m7G
m5C
Y
Ψ
m2G
m2G
B
T
T
G
C
T
G
G
G
C
U
U
G
C
U
G
G
G

cleavage among the regions differ for the tRNAs versus tDNA, and some regions on tDNA are targeted which are not evident on the tRNAs. For example, on $tRNA^{Phe}$ cleavage is evident primarily within two regions of the molecule, most strongly within the folded corner of the molecule where the D and TΨC loops interact at Ψ55 and U59, and also, but more weakly, along the central region created through the extensive tertiary interactions centered around m^7G46. On tDNA, instead, it is this central tertiary structure which is most strongly targeted, and the primary site of cleavage is shifted to C48. Cleavage is also evident on tDNA at T55 and T59, but again shifted in that greater cleavage is evident at C56 and C60. The tRNA transcript represents an intermediate case; here cleavage is also found in both tertiary folds of the molecule, despite the replacement of Ψ55 by U55, and within the central portion of the molecule, where cleavage is centered on U47.

The sites not present on the tRNAs but which are targeted by the rhodium complex on $tDNA^{Phe}$ occur along the stem regions. On the $tDNA^{Phe}$ acceptor and anticodon stems, there is a strong correspondence between 5'-pyr-pyr-pur-3' steps and cleavage sites, as shown in Table 2.2. On B-form DNA these sites are characteristically recognized by $Rh(phen)_2phi^{3+}$.[19] For example, $Rh(phen)_2phi^{3+}$ targets 5'-CCA-3' on the anticodon stem, a sequence which is not cleaved on $tRNA^{Phe}$ but which represents the primary recognition site for the complex on duplex DNA [19]. This targeting is therefore consistent with the notion that the tDNA differs in structure from the tRNA in adopting a more B-like conformation within the double helical acceptor and anticodon stems. This observation is also consistent with an NMR study which indicated that an RNA pentameric helix with the sequence (5'rCUGUG/5'rCACAG), corresponding to the TΨC stem in yeast $tRNA^{Phe}$, adopts an A-form conformation, whereas the analogous DNA pentamer is found to adopt the B-form conformation.[31] It is noteworthy, however, that these sites are generally weaker in intensity than the sites of tertiary interaction.

Table 2.2. Correspondence of selected cleavage sites on $tDNA^{Phe}$ to 5'-pyr-pyr-pur-3' sites in double helical regions.

Cleavage site(s)	5'-pyr-pyr-pur-3' steps	Location
A9	T7 T8 A9	Acceptor stem/D loop
C28 A29	C27 C28 A29	Anticodon stem
C32	C32 T33 G34	Anticodon loop
C40	C40 T41 G42	Anticodon stem
T69 C70	T69 C70 G71	Acceptor stem

Furthermore, the cleavage is weaker in intensity than is commonly observed in DNA oligonucleotides.

2.3.2. Cleavage of $tDNA^{Phe}$ by $Rh(DIP)_3^{3+}$

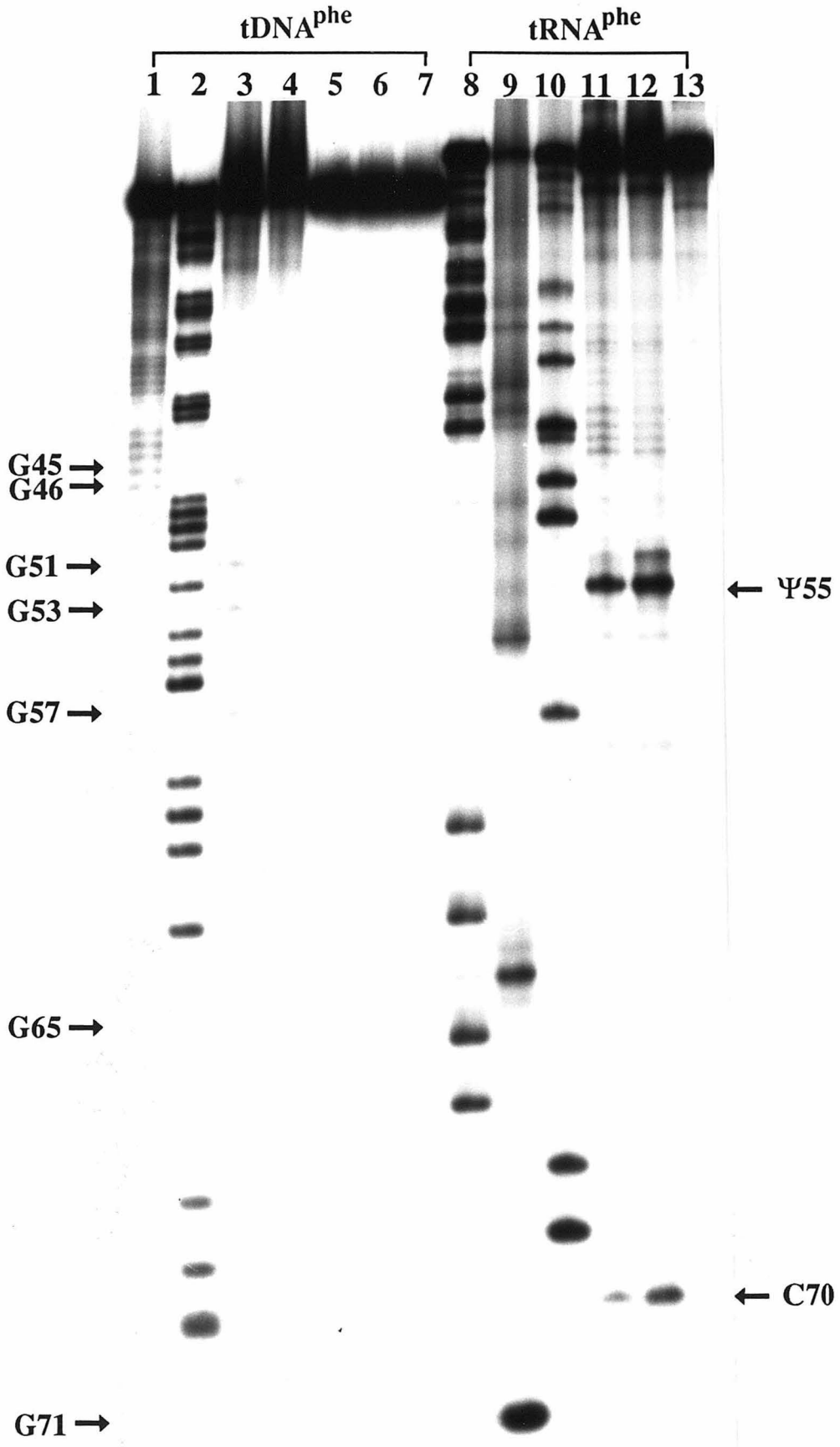
$Rh(DIP)_3^{3+}$ has been shown to target selectively two sites on $tRNA^{Phe}$, Ψ55 and C70. Targeting of Ψ55 appears to depend upon the complex tertiary folding of the tRNA where the D and TΨC loops interact; in the tRNA transcript, which lacks the pseudouracil, $Rh(DIP)_3^{3+}$ targets U55.[26] Importantly, targeting of C70 appears to depend upon the recognition of the neighboring GU mismatch. In several RNAs, $Rh(DIP)_3^{3+}$ has been shown to target the base to the 3'-side of a GU mismatch within double helical regions.[26] In the absence of a strong recognition site, on DNAs a faint reaction at guanines has generally been found.[25]

The comparison in cleavage by $Rh(DIP)_3^{3+}$ on $tRNA^{Phe}$ and $tDNA^{Phe}$ is quite striking, as shown in Figure 2.6. Although strong cleavage is evident by the rhodium complex at Ψ55 and C70 on $tRNA^{Phe}$, only a faint reaction at guanines is evident on $tDNA^{Phe}$. The same absence of specific cleavage on $tDNA^{Phe}$ is evident with higher concentrations of rhodium and longer irradiation times, or the addition of 10 mM $MgCl_2$.

2.3.3. Cleavage of $tDNA^{Phe}$ by $Ru(TMP)_3^{2+}$ and by $Ru(phen)_3^{2+}$

$Ru(TMP)_3^{2+}$ has been shown to selectively bind A-form helical regions of RNAs and DNAs.[27,28] Since $Ru(TMP)_3^{2+}$ promotes nucleic acid strand cleavage in a nucleic acid base-dependent reaction mediated by singlet oxygen, comparative cleavage experiments are always carried out using $Ru(phen)_3^{2+}$, which also promotes nucleic acid cleavage with irradiation in a reaction mediated by singlet oxygen, but which shows no preference for A-form regions because of its smaller size and hydrophobicity compared to $Ru(TMP)_3^{2+}$. Earlier studies have shown that some differences in cleavage are evident between $Ru(phen)_3^{2+}$ and $Ru(TMP)_3^{2+}$ on $tRNA^{Phe}$, in that some protection from

Figure 2.6. Autoradiogram showing cleavage of 3'-^{32}P-labeled tDNA^{Phe} and tRNA^{Phe} by $Rh(DIP)_3^{3+}$ in 50 mM Tris, 20 mM NaOAc, 18 mM NaCl, pH 7.0. Lanes 1 and 2: A+G and C+T reactions, respectively, on tDNA^{Phe}. Lane 3: cleaved tDNA^{Phe} after incubation with 2.5 μM $Rh(DIP)_3^{3+}$ and irradiation for 6' at 313 nm. Lane 4: cleaved tDNA^{Phe} after incubation with 10 μM $Rh(DIP)_3^{3+}$ and irradiation for 10' at 365 nm. Lane 5: labeled DNA with 10 μM $Rh(DIP)_3^{3+}$ but without irradiation. Lane 6: labeled DNA upon irradiation for 6' at 313 nm. Lane 7: labeled DNA without metal or irradiation. Lanes 8, 9 and 10: A-, G- and U- specific reactions respectively on tRNA^{Phe}. Lane 11: cleaved tRNA^{Phe} after incubation with 2.5 μM $Rh(DIP)_3^{3+}$ and irradiation for 6' at 313 nm. Lane 12: cleaved tRNA^{Phe} after incubation with 10 μM $Rh(DIP)_3^{3+}$ and irradiation for 10' at 365 nm. Lane 13: cleaved tRNA^{Phe} after incubation with 10 μM $Rh(DIP)_3^{3+}$ but without irradiation.


tDNA^phe
tRNA^phe
1 2 3 4 5 6 7 8 9 10 11 12 13
G45
G46
G51
G53
G57
G65
G71
Ψ55
C70

cleavage is observed neighboring the helical regions of the molecule by $Ru(TMP)_3^{2+}$.[21] No strong selectivity was apparent, however. On $tDNA^{Phe}$, no differences in cleavage between $Ru(phen)_3^{2+}$ and $Ru(TMP)_3^{2+}$ are observed, as shown in Figure 2.7. Instead, the same amount of cleavage by both complexes is seen at guanine residues and the adjoining residues. Since some differences are evident with tRNA but not with tDNA, this result also supports a difference in conformation between helical regions of $tRNA^{Phe}$ and $tDNA^{Phe}$.

2.4. Discussion

2.4.1. Comparisons in Structures of $tRNA^{Phe}$ and $tDNA^{Phe}$

Chemical probing of the structure of $tDNA^{Phe}$ with the shape-selective transition metal complexes indicates that globally the tertiary folding of the DNA analog resembles that of $tRNA^{Phe}$ but that some differences in local conformation are evident. The metal complexes provide site-selective probes of different regions of the molecule: the acceptor and anticodon stems, the central region where triply bonded bases interact, and the region where the D and TΨC loops associate. A computer graphic representation of the crystal structure of $tRNA^{Phe}$ which highlights the similarities and differences between $tDNA^{Phe}$ and $tRNA^{Phe}$ as probed by the various rhodium complexes is shown in Figure 2.8.

Both $Rh(phen)_2phi^{3+}$ and $Rh(DIP)_3^{3+}$ provide probes of the tertiary folding in the D and TΨC loops. In $tRNA^{Phe}$, intense cleavage by $Rh(phen)_2phi^{3+}$ is observed at Ψ55 and U59. On tDNA, cleavage is maintained by $Rh(phen)_2phi^{3+}$ in the vicinity of T55 and T59, but shifted by one base to T56 and C60, and of lower intensity. With $Rh(DIP)_3^{3+}$, despite strong cleavage at Ψ55 in $tRNA^{Phe}$, and consistent cleavage in a series of tRNA mutants containing substitutions within that region of the molecule, no specific cleavage in the vicinity of T55 is detectable on tDNA. These results may indicate that the basic folding within this region is maintained on the DNA analog, but perhaps is less flexible in structure and therefore less accessible to association with the bulky $Rh(DIP)_3^{3+}$.

Figure 2.7. Autoradiogram showing the comparison of $Ru(TMP)_3^{2+}$ and $Ru(phen)_3^{2+}$ photocleavage on tDNA (lanes 1-8) and $tRNA^{Phe}$ (lanes 9-17). Lanes 1 & 2: A+G and C+T Maxam Gilbert sequencing lanes. Lanes 3 & 5: cleaved tDNA after incubation with 10 μM $Ru(phen)_3^{2+}$ and $Ru(TMP)_3^{2+}$ respectively, irradiated at 442 nm for 30'. Lanes 4 & 6: labeled DNA with 10 μM $Ru(phen)_3^{2+}$ and $Ru(TMP)_3^{2+}$ respectively, no irradiation. Lanes 7 & 8: labeled DNA after 30' irradiation at 442 nm respectively. Lanes 9, 10 & 11: A-, G- and U- specific reactions respectively on $tRNA^{Phe}$. Lanes 12 & 14: cleaved $tRNA^{Phe}$ after incubation with 10 μM $Ru(phen)_3^{2+}$ and $Ru(TMP)_3^{2+}$ respectively, irradiated at 442 nm for 30'. Lanes 13 & 15: labeled $tRNA^{Phe}$ with 10 μM $Ru(phen)_3^{2+}$ and $Ru(TMP)_3^{2+}$ respectively, no irradiation. Lanes 16 & 17: labeled $tRNA^{Phe}$ after 30' irradiation at 442 nm respectively. All samples except lanes 8 and 17 were treated with aniline.

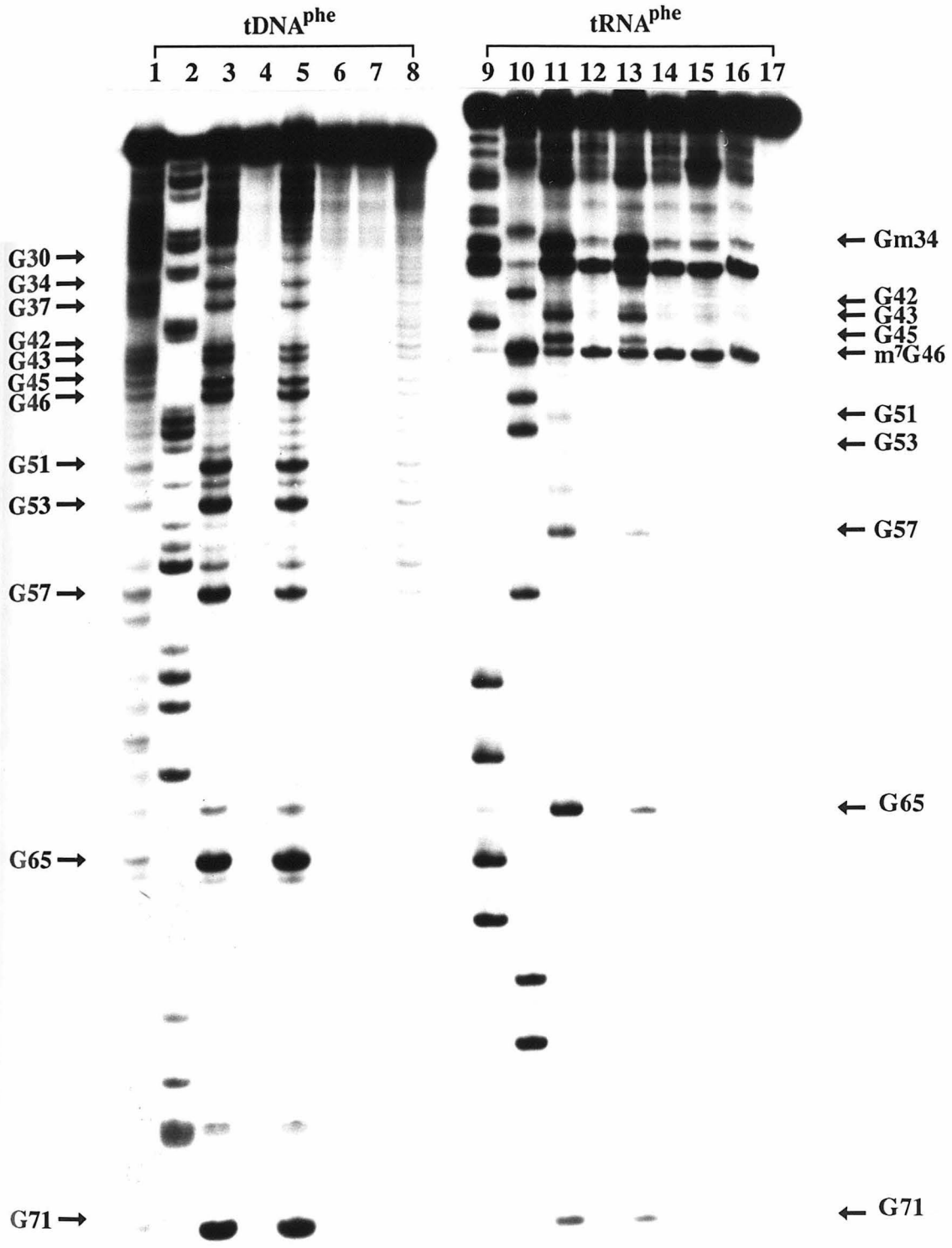
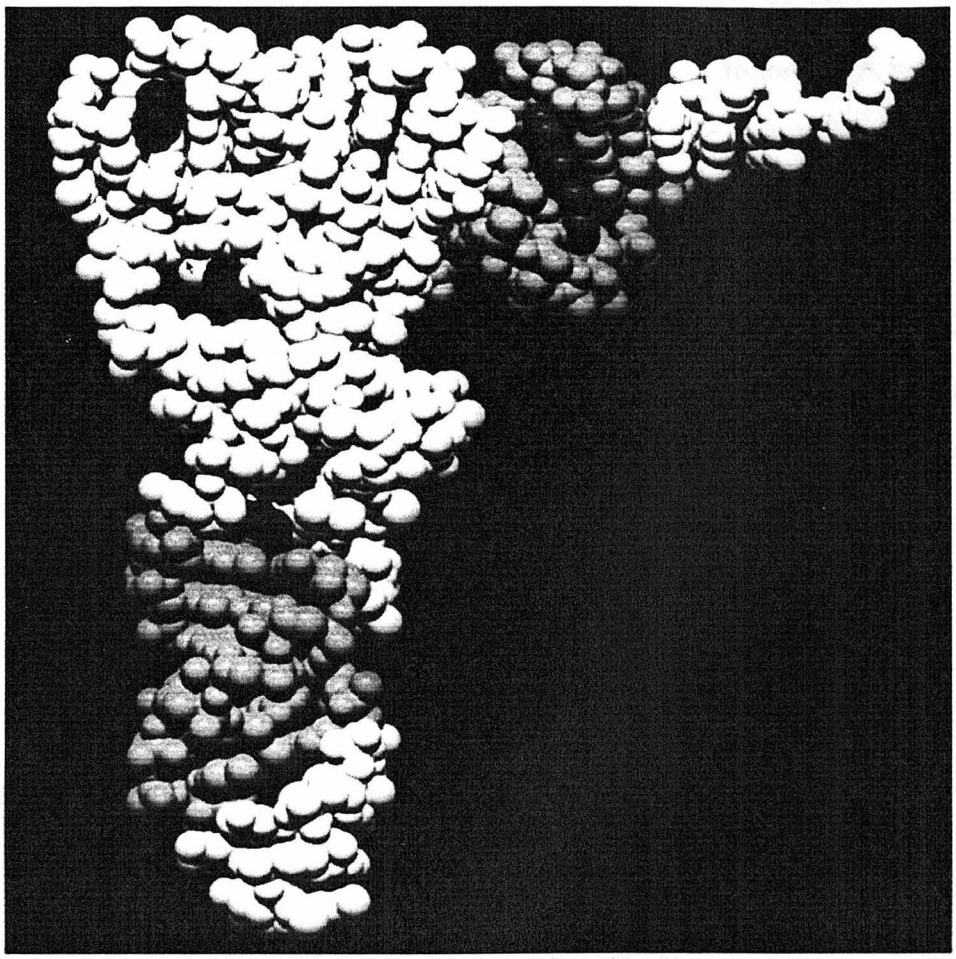

tDNA^phe
tRNA^phe
1 2 3 4 5 6 7 8
9 10 11 12 13 14 15 16 17
G30
G34
G37
G42
G43
G45
G46
G51
G53
G57
G65
G71
Gm34
G42
G43
G45
m7G46
G51
G53
G57
G65
G71

Figure 2.8. Computer graphic representation of the crystal structure of $tRNA^{Phe}$ highlighting the similarities and differences between $tDNA^{Phe}$ and $tRNA^{Phe}$ as probed by rhodium complexes. Double helical regions which are A-form in $tRNA^{Phe}$ and B-form in $tDNA^{Phe}$ are shown in light grey. The GU mismatch recognized by $Rh(DIP)_3^{3+}$ on $tRNA^{Phe}$ is shown in grey; the corresponding GT mismatch on DNA is not recognized. Bases involved in tertiary and/or triple interactions which are recognized by $Rh(phen)_2phi^{3+}$ on $tRNA^{Phe}$, $tRNA^{Phe}$ transcript and $tDNA^{Phe}$ are shown in dark gray.



The overall structure and folding of tRNA is also defined by the triple base interactions in the central core of the molecule, and on $tRNA^{Phe}$ these sites are specifically targeted by $Rh(phen)_2phi^{3+}$. On tDNA this region is also specifically targeted, and indeed with the DNA analog this region represents the primary area for rhodium-induced cleavage. That this region in $tDNA^{Phe}$ is specifically recognized by $Rh(phen)_2phi^{3+}$ supports the notion that the DNA analog adopts an overall folded structure like that of tRNA. Nonetheless the details of this folding differ. Cleavage on tDNA occurs strongly at positions C48 and C49, compared to the tRNA transcript where cleavage is centered on U47, and the native $tRNA^{Phe}$ where cleavage is found most strongly around m^7G46. It is noteworthy that C48 and C49 are recognized, even though they do not lie within a double helical 5'-pyr-pyr-pur-3' region like most of the other targeted sites. Therefore, this points to this region being structured in some fashion.

The clear difference in targeting of the DNA and RNA analogs is apparent within the double helical stem regions. $Rh(phen)_2phi^{3+}$ targets 5'-pyr-pyr-pur-3' sites on B-form DNA since these sites are open and accessible in the major groove for intercalative binding.[19,20] On $tRNA^{Phe}$ and a series of tRNA mutant transcripts, no specific cleavage within the acceptor or anticodon stems by $Rh(phen)_2phi^{3+}$ is observed.[23] In sharp contrast, on tDNA specific cleavage is observed at 5'-pyr-pyr-pur-3' sites within the putatively double helical regions. This targeting therefore indicates that, like the tRNA analogs, tDNA contains double helical regions, but, unlike the tRNA analogs, tDNA contains B-form helical regions. For tDNA the major groove is accessible to intercalation, whereas in tRNAs in these double helical regions the major groove is not accessible, likely because the helix is more A-like in character. Experiments with $Ru(TMP)_3^{2+}$ also support this conclusion.

It is noteworthy that the anticodon loop in tDNA neither resembles that of the native $tRNA^{Phe}$ nor of the $tRNA^{Phe}$ transcript, based upon cleavage by $Rh(phen)_2phi^{3+}$. As may be evident in Figure 2.3, on tDNA, cleavage occurs within 5'-C32-T33-G34-3', a

5'-pyr-pyr-pur-3' step, suggesting that this loop in tDNA may maintain the B-like helicity of the anticodon stem. No specific cleavage is apparent in the anticodon loop of the native tRNA, and some cleavage though different is evident in the more flexible loop of the tRNA transcript.

The observation that there is no difference in the cleavage pattern on $tDNA^{Phe}$ by $Rh(phen)_2phi^{3+}$ in the presence of $MgCl_2$ suggests that the metal ion has no significant effect on the folding of that molecule. This is in contrast to the loss of cleavage on tertiary sites on $tRNA^{Phe}$ which was attributed to either a tightening or a loosening of the tertiary structure.[23] Thus, $tDNA^{Phe}$ is less flexible than $tRNA^{Phe}$ in terms of a Mg^{2+}-induced conformational change; this may be true of the tDNA molecule in general.

Also of particular note is the absence of recognition of the GT mismatch within the anticodon stem of tDNA by $Rh(DIP)_3^{3+}$. On $tRNA^{Phe}$ and microhelices derived from it, $Rh(DIP)_3^{2+}$ has been shown to target GU mismatches, specifically cleaving to the 3' side of the U, within double helical regions.[26] A GU mismatch does not stack well with the flanking base pair to the 3' side of the U, and the wobble-paired U residue is pushed away from the helix interior into the major groove of RNA. Yet no specific recognition of the GT mismatch in tDNA by $Rh(DIP)_3^{3+}$ is evident. The absence of recognition may be a consequence of the change in conformation of the region to the B-form, where perhaps the GT wobble does not present the structural perturbation[32] that is apparent within the A-form helix. We cannot, however, eliminate the effects on recognition presented by the methyl group on T. This difference in recognition between RNA and DNA is important, since it is known that the G3-U70 mismatch is critical for identity of *E. coli* $tRNA^{Ala}$ [33,34] and $tRNA^{Asp}$.[35] The fact that the $tDNA^{Phe}$ is still recognized by its cognate synthetase[6] implies that the GU mismatch is not crucial for the identity of yeast $tRNA^{Phe}$. It would be interesting to compare aminoacylation experiments on *E. coli* $tDNA^{Ala}$ and $tRNA^{Ala}$. GU mismatches appear generally to serve as recognition elements in the interactions of proteins and RNA. The substitution of deoxy analogs

within RNA molecules must therefore take this effect on the structure of the GU mismatch into account. It is not unreasonable to consider that, like $Rh(DIP)_3^{3+}$, RNA-binding proteins may also lose an ability to distinguish the GU mismatch within a DNA polymer.

2.4.2 Implications for Studies of RNA and DNA Structure

Because of relative ease in synthesis and greater stability, DNA analogs have increasingly been exploited in studies to explore RNA structure, reaction, and function.[8-10] Yet few structural probes are available to examine, locally along the polymer, the structural consequences in converting from the ribose to deoxyribose form. The present study illustrates the utility of transition metal complexes in comparing with high sensitivity the differences and similarities in the local structure of DNA and RNA analogs. The results indicate that globally the structure of tRNA and tDNA analogs resemble one another. Yet distinctions are apparent, in particular within double helical regions. Moreover detailed structural perturbations, such as the GU mismatch, may vary considerably and thus offer different distinguishing elements for protein recognition. The results described here therefore provide support for the application of DNA analogs in studies of RNA structure and function. However, the present work also serves to underscore the importance and powerful application of transition metal complexes to probe local variations in RNA and DNA structure, and how, if chemical probing is conducted in concert with studies of RNA function, the results of such studies may be far more clearly understood.

References

1. Usman, N.; Ogilvie, K. K.; Jiang, M. Y.; Cedergren, R. L. *J. Am. Chem. Soc.* **1987**, *109*, 7845-7854.

2. Usman, N.; Nicoghosian, K.; Cedergren, R. L.; Ogilvie, K. K. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5764-5768.

3. Scaringe, S. A.; Francklyn, C.; Usman, N. *Nucleic Acids Res.* **1990**, *18*, 5433-5441.

4. Gasparutto, D.; Livache, T.; Duplaa, A.-M.; Bazin, H.; Favario, S.; Guy, A.; Molko, D.; Roget, A.; Teoule, R. *C. R. Acad. Sci. Paris* **1992**, *315*, 1-6.

5. Miligan, J. F.; Groebe, D. R.; Witherall, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783-8798.

6. Khan, A. S.; Roe, B. A. *Science* **1988**, *241*, 76-79.

7. Perreault, J.-P.; Pon, R. T.; Jiang, M.; Usman, N.; Pika, J.; Ogilvie, K. K.; Cedergren, R. *Eur. J. Biochem.* **1989**, *186*, 87-93.

8. Perreault, J.-P.; Altman, S. *J. Mol. Biol.* **1992**, *226*, 399-409.

9. Yang, J. H.; Usman, N.; Chartrand, P.; Cedergren, R. *Biochemistry* **1992**, *31*, 5005-5009.

10. Dahm, S.; Uhlenbeck, O. C. *Biochimie* **1990**, *72*, 819-823.

11. Sampson, J. R.; Behlen, L. S.; DiRenzo, A. B.; Uhlenbeck, O. C. *Biochemistry* **1992**, *31*, 4161-4167.

12. Paquette, J.; Nicoghosian, K.; Qi, G.; Beauchemin, N.; Cedergren, R. *Eur. J. Biochem.* **1990**, *189*, 259-265.

13. Quigley, G. J.; Rich, A. *Science* **1976**, *194*, 796-806.

14. Kim, S.-H.; Sussman, J. L.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Wang, A., H.; Seeman, N. C.; Rich, A. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 4970-4974.

15. Chow, C. S.; Barton, J. K. *Meth. Enzym.* **1992**, *212*, 219-242.

16. Pyle, A. M.; Barton, J. K. *Prog. Inorg. Chem.* **1990**, *38*, 413-475.

17. Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520-4522.

18. Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432-9434.

19. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303-2312.

20. David, S. S.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 2984-2985.

21. Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839-2841.

22. Chow, C. S.; Hartmann, K. M.; Rawlings, S. L.; Huber, P. W.; Barton, J. K. *Biochemistry* **1992**, *31*, 3534-3542.

23. Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972-982.

24. Kirshenbaum, M. R.; Tribolet, R.; Barton, J. K. *Abs. Pap. American Chemical Society* **1988**, *196*, 92-INOR.

25. Lee, I.; Barton, J. K. *Biochemistry* **1993**, *32*, 6121-6127.

26. Chow, C. S.; Barton, J. K. *Biochemistry* **1992**, *31*, 5423-5429.

27. Mei, H. Y.; Barton, J. K. *Abs. Pap. American Chemical Society* **1988**, *195*, 476-INOR.

28. Mei, H. Y.; Barton, J. K. *J. Am. Chem. Soc.* **1986**, *108*, 7414-7416.

29. Peattie, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1760.

30. Sambrook, J.; Fritsch, E. F.; Manniatis, T. *Molecular Cloning: a Laboratory Manual*; 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.

31. Hall, K. B.; McLaughlin, L. W. *Biochemistry* **1991**, *30*, 10606-10613.

32. Quignard, E.; Fazakerley, G. V.; van der Marel, G.; van Boom, J. H.; Guschlbauer, W. *Nucleic Acids Res.* **1987**, *15*, 3397-3409.

33. Hou, Y.-M.; Schimmel, P. *Nature* **1988**, *333*, 140-145.

34. Hou, Y.-M.; Schimmel, P. *Biochemistry* **1989**, *28*, 6800-6804.

35. Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A.; Podjarny, A.; Rees, B.; Thierry, J. C.; Moras, D. *Science* **1991**, *252*, 1682-1689.

Chapter 3:

$Rh(phen)_2phi^{3+}$ as a Shape-selective Probe of Triple Helices

3.1. Introduction

$Rh(phen)_2phi^{3+}$ (phen = 1,10-phenanthroline; phi = 9,10-phenanthrenequinone diimine) (Figure 3.1) is a unique site-selective probe for RNA structures that preferentially targets sites on folded RNA which are open or accessible for intercalative stacking from the major groove.[1] These sites include major groove recognition sites for proteins[2] as well as sites of RNA tertiary interaction where triply bonded bases are arranged to permit stacking from the major groove. In studies which probed rhodium recognition sites on tRNAs and tRNA mutants,[3] $Rh(phen)_2phi^{3+}$ was shown to target the triply bonded base segments which together establish the folded structure of the RNA. In order to characterize in more detail this recognition of triple bases by $Rh(phen)_2phi^{3+}$, and in so doing to explore the scope of the rhodium complex as a chemical probe for RNA tertiary structure, we have undertaken a study of rhodium reaction with synthetic RNA triple helices.

Triple helical structures were first described based upon fiber diffraction studies of the RNA polymers poly(U)•poly(A)-poly(U).[4] More recently substantial interest has been focused on DNA and RNA triplexes in developing strategies for site-specific targeting for both genomic mapping[5] and therapeutic applications.[6] Two triplex motifs have been elucidated, a purine•purine-pyrimidine (R•R-Y) motif[7] and a pyrimidine•purine-pyrimidine (Y•R-Y) motif.[4,8] In the DNA R•R-Y motif, the purine third strand was shown to bind antiparallel in the major groove of the Watson-Crick purine strand through Hoogsteen hydrogen bonds[7] while in the Y•R-Y motif, the pyrimidine third strand is parallel to the purine Watson Crick strand.[6] Examples of both motifs have been identified also with RNA.[9-11] Intramolecular DNA triple helices (termed H-DNA) have been observed in supercoiled plasmids[12] or may be formed from a

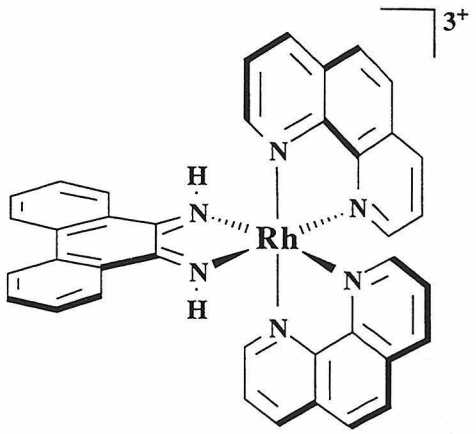


Figure 3.1. Schematic illustration of $Rh(phen)_2phi^{3+}$.

single strand of chemically synthesized oligonucleotide.[13] A recent crystal structure of parallel and antiparallel $(G•GC)_2$ triple helix fragments formed from the overlap of overhanging DNA duplexes has been described.[14] In both the parallel and antiparallel triplex fragments, the normal B-form DNA Watson-Crick components of the triplets are relatively unperturbed by the third strand.

A structural understanding of RNA triple helices is of interest in part because of the variety of triple base interactions which serve to direct the folding of complex RNA molecules. The base triples m^7G46-[G22-C13], G45-[m^2G10-C25] and A9-[A23-U12] help to generate the structural core of $tRNA^{Phe}$.[13,15,16] Crystallographical data have established the presence of A•U-A and A•C-G triple bases in the P4-P6 domain of the Tetrahymena intron.[17] These triples function in contributing to the tertiary folding of the intron. Chemical probes which preferentially target triply bonded bases in RNA molecules, therefore, could be powerfully applied as an aid in elucidating RNA tertiary structure and folding.

$Rh(phen)_2phi^{3+}$ recognizes RNA in a manner akin to its shape-selective recognition of DNA sites. $Rh(phen)_2phi^{3+}$ has been shown through NMR studies[18] and biophysical methods[19,20] to intercalate into duplex DNA from the major groove with high affinity. A high resolution NMR structure of a related complex, Δ-α–Rh[(R,R)-$(Me_2Trien)_2$)]phi^{3+} (Me_2Trien = 2,9-diamino-4,7-diazadecane), shows this major groove intercalation.[21] Upon photoactivation, the rhodium complex promotes direct strand cleavage; product analysis is consistent with a photoreaction involving abstraction of the C3'-hydrogen atom directly by the photoexcited, intercalated phi ligand,[20] and analogous studies on $tRNA^{Phe}$ [1] indicate equivalent products and efficiencies. The DNA sites selected for targeting by $Rh(phen)_2phi^{3+}$ correlate closely with sites identified crystallographically as being more open in the major groove, in particular those 5'-pyrimidine-pyrimidine-purine-3' sites which show a large major groove opening due to a change in propeller twist.[22] This preference for "open" major groove sites has been

attributed to a high intercalative binding affinity of the phi ligand being modulated by significant steric clashes of the overhanging phenanthroline ligands with the base paired helix; only at sites which are somewhat open in the major groove is stacking by the octahedral complex made facile.

There is precedence for intercalation into triple helices and triplex junctions both by synthetic metal complexes and by naturally occurring organic heterocycles. A dipyridophenazine complex of ruthenium luminesces intensely on binding to triple helices, and the luminescence observed is consistent with binding within the triplex stack rather than at a junction.[23] Ethidium bromide interacts with both duplex and triplex DNA, as does propidium bromide.[24-26] Several benzo(e)pyridoindole derivatives show greater affinity for triplex DNA than for duplex DNA[27,28] and in fact, can stabilize the triplex. Both an acridine intercalator[29] and an oxazolopyridocarbazole chromophore[30] covalently linked to an oligonucleotide have been shown to intercalate preferentially at duplex/triplex junctions rather than within the triplex column. Fe(II)•bleomycin has also been shown to preferentially cleave at duplex/triplex junctions.[31]

$Rh(phen)_2phi^{3+}$ may be particularly useful in probing RNA tertiary structure because, despite binding through shape-selection to B-DNA sites, the complex neither binds nor cleaves A-form duplexes.[2] This poor reactivity with RNA duplexes is consistent with the preferential intercalation of the complex in the major groove; since an A-form helix contains a narrow and deep major groove, intercalation by the octahedral complex is precluded.[1-3] This is in contrast to the DNA analog of $tRNA^{Phe}$, where cleavage is observed along the A-form stems.[32] In the case of RNA triplexes, however, the third strand itself binds within the duplex major groove and this stacked array may provide an accessible platform for intercalation by the complex. Other tertiary interactions, bulges, and mismatches, which widen the major groove of RNA helices,[33] may also permit access by the metal complex.[2]

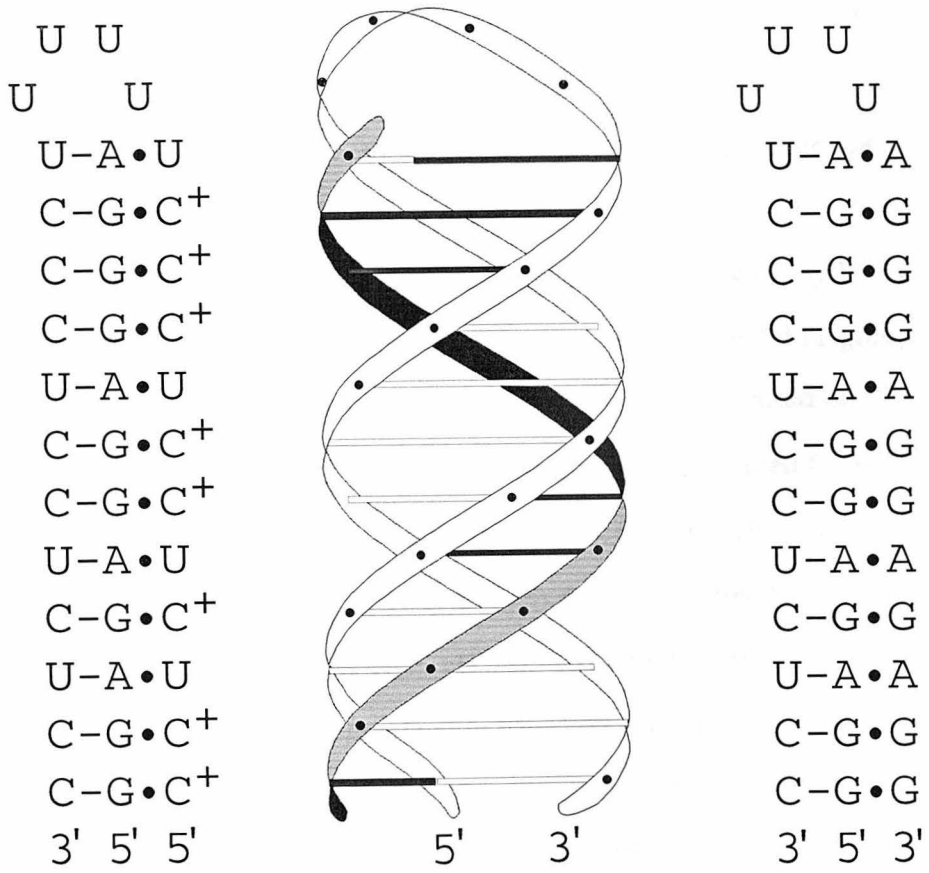
Here we focus specifically on this preferential recognition of RNA triplex structures by $Rh(phen)_2phi^{3+}$. We sought to characterize this interaction on a well-described triplex structure. Figure 3.2 illustrates the RNA triplex oligomers used as substrates in these studies. RNA or DNA pur•pur-pyr and pyr•pur-pyr (pur = purine, pyr = pyrimidine) triple helices consisting of a Watson-Crick base-paired 28mer duplex and a Hoogsteen base-paired purine or pyrimidine 12mer were formed by annealing DNA or RNA 28mer hairpin duplexes with RNA purine or pyrimidine 12mers of defined sequence. The RNA and DNA Y•R-Y synthetic triple helices have already been characterized thermodynamically.[34,35] $Rh(phen)_2phi^{3+}$ cleavage may then be compared and contrasted on both the RNA and DNA hairpin duplexes and their corresponding triplexes.

Here we find that the metal complex interacts with these triple helices in a structure-specific manner. Different cleavage patterns are seen with the pyr•pur-pyr and pur•pur-pyr motifs. Cleavage is seen on both the Watson-Crick strands of the former motif and primarily on the purine Watson-Crick strand of the latter motif. Little cleavage is seen on the Hoogsteen strand for either motif. The cleavage patterns are consistent with an intercalated model for the metal complex in the triple helix. Importantly, the metal complex shows no detectable cleavage on the A-form RNA duplex in the absence of a third strand. Similar cleavage is seen on DNA triple helices, but over a background of duplex cleavage. Targeting of synthetic RNA triple helices but not duplex regions by $Rh(phen)_2phi^{3+}$ therefore provides a basis for the chemical probing of triply bonded sites in folded RNA molecules.

3.2. Experimental Methods

RNA oligomers. RNA oligonucleotides were chemically synthesized on a 1 μmole scale on an Applied Biosystems 392 DNA/RNA synthesizer using phosphoramidite chemistry. Oligomers were desalted and failure sequences removed

Figure 3.2. Schematic of the RNA triple helices used in this study. Triple helices were formed by annealing together a 28-mer RNA hairpin duplex with either a 12-mer purine strand (5'-AGGGAGGAGAGG-3') or a 12-mer pyrimidine (5'-CCUCUCCUCCCU-3') strand.


U U
U U
U-A•U
C-G•C+
C-G•C+
C-G•C+
U-A•U
C-G•C+
C-G•C+
U-A•U
C-G•C+
U-A•U
C-G•C+
C-G•C+
3' 5' 5'
5' 3'
U U
U U
U-A•A
C-G•G
C-G•G
C-G•G
U-A•A
C-G•G
C-G•G
U-A•A
C-G•G
U-A•A
C-G•G
C-G•G
3' 5' 3'

using an oligonucleotide purification column from Applied Biosystems. Extinction coefficients at 260 nm were 10200, 7300 and 8800 $M^{-1}cm^{-1}$ /nucleotide based on nucleotide composition for the purine 12mer, the pyrimidine 12mer, and the 28mer respectively. Oligomers were 3'-end-labeled with cytidine 3',5'-[5'-^{32}P]-bisphosphate using T4 RNA ligase [36] or 5'-end-labeled with [γ-^{32}P]-ATP using T4 polynucleotide kinase. Samples were then gel purified on a 10% denaturing polyacrylamide gel, located by autoradiography, excised, and eluted from the gel slice in 45 mM Tris-HCl, 45 mM boric acid, 1.25 mM EDTA, pH 8.0. The eluted RNA oligomers were ethanol-precipitated twice and stored frozen in 10 mM Tris-HCl, pH 7.5.

DNA oligomers. DNA analogs of the RNA strands were chemically synthesized on a 1 μmole scale on an Applied Biosystems 392 DNA/RNA synthesizer using phosphoramidite chemistry. Oligonucleotides were purified twice by HPLC, first with the dimethoxytrityl (DMT) group on and subsequently with the DMT group off using a C_{18} column (Dynamax). The oligonucleotides were further purified on a 10% denaturing polyacrylamide gel, located by UV shadowing, excised, and eluted from the gel in 45 mM Tris-HCl, 45 mM boric acid, 1.25 mM EDTA, pH 8.0. The eluted DNA oligomers were concentrated by a Centricon 10 (Amicon) device, desalted by washing twice with water, and stored in 10 mM Tris-HCl, pH 7.5. Extinction coefficients at 260 nm were 9300, 7300 and 8300 $M^{-1}cm^{-1}$ /nucleotide based on nucleotide composition for the purine 12mer, the pyrimidine 12mer, and the 28mer respectively. The oligomers were then 3'-end-labeled with [α-^{32}P]-ddATP using terminal deoxytransferase or 5'-end-labeled with [γ-^{32}P]-ATP using T4 polynucleotide kinase. The labeled materials were gel-purified by the same method as for RNA oligomers and stored frozen in 10 mM Tris-HCl, pH 7.5.

Triplex Formation. Gel retardation assays were employed to establish triplex formation. In separate experiments, either the hairpin or third strand were radioactively labeled and conditions for retardation relative to the double stranded or single strand controls, respectively, were determined. Duplex alone or duplex with either purine or

pyrimidine 12mer to a final total concentration of 100 μM in nucleotides were renatured by heating to 65°C for 1 min in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine, pH 5.5, and slowly cooling to 4°C. The samples were then allowed to stand at 4°C for 4 to 12 hr. Samples were then loaded onto a 20% polyacrylamide gel at 4°C which had been prepared with 50 mM sodium acetate, 1 mM EDTA, pH 5.5. Samples were electrophoresed at low voltage at 4°C before being visualized by autoradiography.

Cleavage Reactions. $[Rh(phen)_2phi]Cl_3$ was synthesized as described previously [37]. All metal stock solutions were freshly prepared in either ethanol or 10 mM Tris-HCl, pH 7.5. The end-labeled oligomers and either carrier oligomer or $tRNA^{Phe}$ (Boehringer Mannheim) at a final concentration of 100 μM in nucleotides were renatured by heating to 65°C for 1 min in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine, pH 5.5, and slowly cooled to 4° C. 2.5-10 μM metal complex (freshly diluted in H_2O) at this temperature was then added. The 20 μL mixture was incubated between 5 min to 12 h at 4° C and was then irradiated at 365 or 313 nm at ambient temperature or on ice using a 1000-W Hg/Xe lamp and monochrometer (Oriel model 77250). The reaction mixtures were ethanol precipitated and washed at least three times with ethanol to remove buffer salts, and then dried on a SpeedVac Concentrator (Savant).

RNase Digestions. Annealed duplexes and triplexes, with carrier oligomer to a final concentration of 28 μM in nucleotides in buffer, were prepared in the manner described for the metal cleavage reactions and incubated at 4°C. 1-5 units of either RNase *Physarum polycephalum* (*Phy*M) or RNase *Bacillus cereus* (*Bc)* (Pharmacia) was added, and the samples incubated for 15 - 70 min at 4°C before being ethanol precipitated or quenched with denaturing loading dye.

Sequencing Gels. The cleavage products were analyzed on 20% polyacrylamide-8M urea gels and viewed by autoradiography. The full-length RNA oligomers and cleavage products were identified by coelectrophoresing with $Ru(phen)_3^{2+}$ (G-specific) reactions,[1] diethyl pyrocarbonate (DEPC) (A-specific) and hydrazine (U-specific)

reactions,[38] and the full-length DNA oligomers and cleavage products by piperidine formate (A+G-specific), hydrazine hydrate (C+T-specific), dimethyl sulfate (G-specific) or hydrazine/NaCl (C-specific) Maxam Gilbert reactions.[39] The fragments produced by the metal complex cleavage possess 3' and 5' phosphate termini, and thus may be directly compared with the chemical sequencing lanes.

Factors affecting triplex formation and cleavage. The annealing temperature of the triplex was critical, as incubation at room temperature alone was not sufficient to promote formation of the triple helix. It has been reported based on gel-shift data that a very similar RNA R•R-Y motif could not be formed while the DNA analog could be;[40] however, the addition of spermidine and long incubation at cooler temperatures in the experiments reported here may have helped promote the formation of the R•R-Y motif. Best results were obtained when duplex and third strand were allowed to anneal and slow-cool from 65°C to 4°C. The incubation time of the oligomers was also relatively important. Samples were incubated from 8 h to 24 hr at 4°C, as longer incubation did not seem to produce more cleavage. Based on gel retardation assays, the R•R-Y motif was formed in greater yield than the Y•R-Y motif. Once the triplex was annealed, the metal complex could be added and photolyzed immediately (5 min) or after as much as 12 hr with no difference in cleavage efficiency.

Melting temperatures of the duplexes and triplexes. RNA and DNA duplex and triplex melting temperatures were measured on the Beckman DU 7400 UV-vis spectrophotometer equipped with a 6 cell melting temperature unit. The concentration of RNA strands was kept constant at 1.33 μM ends, or the total RNA nucleotide concentration was kept constant at 50 μM. RNA solutions were prepared in triplex buffer at pH 5.5, and incubated at 90°C for 5 min. The samples were then annealed from 90°C to 4°C at 0.5 deg/min intervals, halted at 4°C for 60 min, then heated to 90°C again at 0.5 deg/min intervals. The absorption was monitored at 260 nm.

The melting temperature of the RNA duplex was found to be 79 ± 1 °C. The melting temperature curves of the triplexes also had a single inflection point, and it was not possible to distinguish when the third strand melted. However, the addition of the third strand raised the melting temperature of the duplex. The melting temperature of the R•R-Y motif was found to be 85 ± 1°C, while that of the Y•R-Y motif was found to be 84 ± 3 °C. Thus, even though the transition where the third strand melts is not observed, the addition of third strand does appear to stabilize the duplex.

Quantitation. Quantitation was accomplished using photostimulable storage phosphorimaging Kodak screen S0230 from Molecular Dynamics. A Molecular Dynamics 400S PhosphorImager was used to scan the screens, and Imagequant version 3.3 was used to analyze the data.

3.3. Results

3.3.1. Gel Retardation Assay for Triplex Formation

The RNA R•R-Y triplex was found to be substantially retarded in its mobility compared to the labeled single RNA homopurine strand (Figure 3.3). The complementary experiment, where the labeled duplex is compared to the triplex, showed the same result. Purine triplexes formed with either labeled purine or labeled duplex strand show identical gel mobilities, as shown in Figure 3.4. Conditions for triplex formation established with this assay, were applied in photocleavage experiments. The analogous RNA Y•R-Y triple helices, although substantially formed, were formed in lesser and variable yield (≤ 20%) under identical conditions (data not shown). The yields of triplex by gel shift assays varied greatly, as can be seen by comparing Figures 3.3 and 3.4.

The analogous DNA R•R-Y and Y•R-Y helices could be formed under these same conditions. Their stability differed from the analogous RNA triplexes. Figure 3.5

Figure 3.3. Autoradiogram showing gel retardation of R•R-Y triplex compared to 3'-^{32}P-labeled purine 12mer alone. Lane 1: labeled purine 12mer in Running buffer. Lane 2-9: labeled and unlabeled (10 pmoles) purine 12mer and unlabeled duplex (12 pmoles) in various buffers; successful attempts (lanes 6 and 7) are in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine, pH 6.0 and 7.0 respectively.

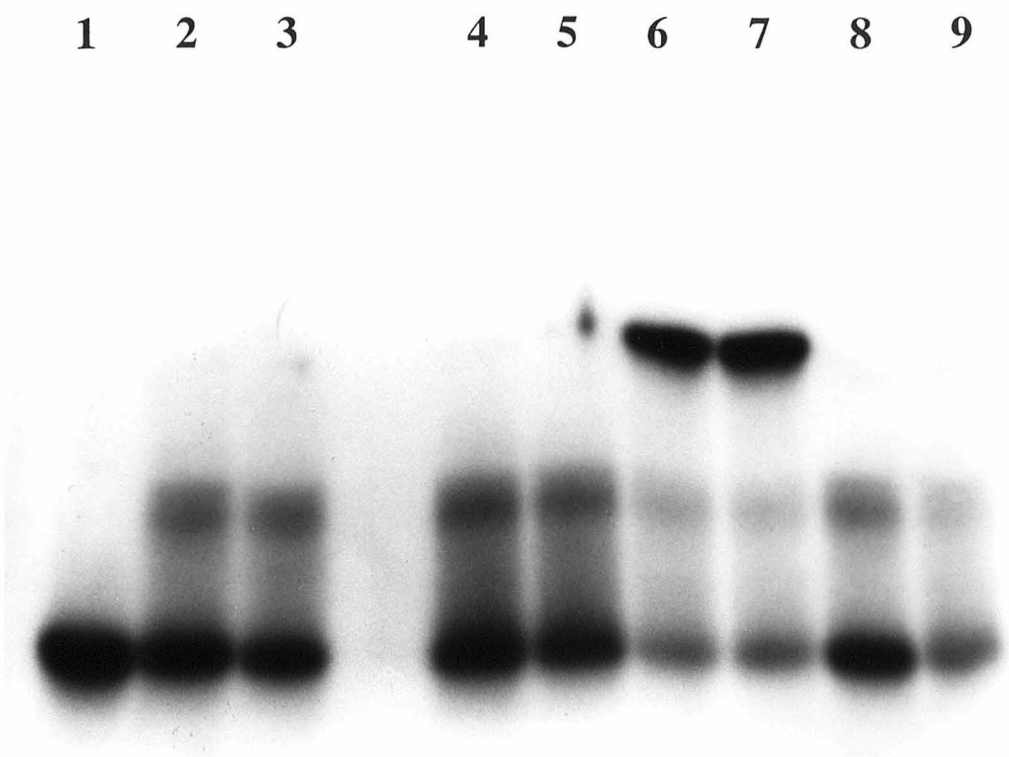

1
2
3
4
5
6
7
8
9

Figure 3.4. Greyscale representation of an autoradiogram showing the formation of RNA R•R-Y and Y•R-Y triplexes in triplex buffer, pH 5.5. Lane 1: labeled and unlabeled (10 pmoles) purine 12mer. Lane 2 & 3: labeled and unlabeled (10 pmoles) purine 12mer, with 6.1 and 9.4 pmoles 28mer duplex respectively. Lanes 4, 5, & 6: labeled and unlabeled (10 pmoles) duplex, with 20, 20 and 206 pmoles purine 12mer respectively. Lane 7: labeled and unlabeled (10 pmoles) duplex. Lane 8: labeled and unlabeled (10 pmoles) duplex, with 10 pmoles unlabeled pyrimidine third strand.

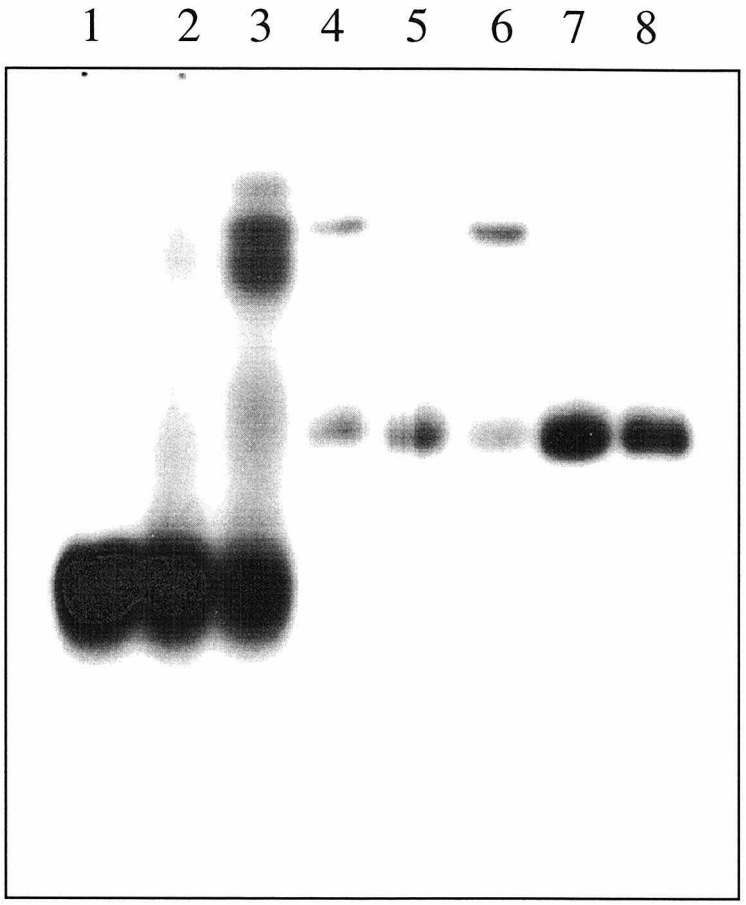
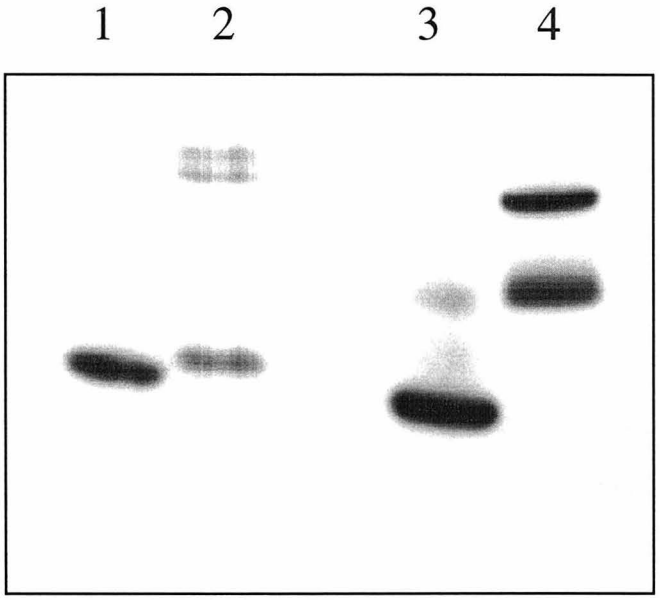

1 2 3 4 5 6 7 8

Figure 3.5. Greyscale representation of an autoradiogram showing the formation of DNA R•R-Y and Y•R-Y triplexes in triplex buffer, pH 5.5. Lane 1: labeled and unlabeled (10 pmoles) purine 12mer. Lane 2: Labeled and unlabeled (10 pmoles) purine 12mer and unlabeled duplex (12 pmoles). Lane 3: labeled and unlabeled (10 pmoles) pyrimidine 12mer. Lane 4: Labeled and unlabeled (10 pmoles) pyrimidine 12mer and unlabeled duplex (12 pmoles).


1
2
3
4

shows a gel where the pyrimidine triplex was formed in approximately 50% yield, but where the purine triplex was formed in less than 50% yield.

3.3.2. Recognition of RNA R•R-Y Triple Helices by $Rh(phen)_2phi^{3+}$

Figure 3.6 illustrates the shape-selective targeting of RNA R•R-Y triple helices by $Rh(phen)_2phi^{3+}$. In the experiment shown, the 3'-^{32}P-endlabeled hairpin RNA was incubated in the presence and absence of the R strand and then tested for photocleavage by $Rh(phen)_2phi^{3+}$. The rhodium complex shows no detectable cleavage of the RNA hairpin duplex. $Rh(phen)_2phi^{3+}$ does not appear to intercalate into an A-form helix. This observation is fully consistent with earlier studies of $tRNA^{Phe}$ and other folded RNAs.[2] However, upon the addition of the third RNA strand with the consequent formation of the R•R-Y triplex, distinct cleavage becomes evident at each purine position of the duplex strand as well as throughout the U loop of the hairpin.

Cleaved fragments are found to comigrate with 5'-phosphate products consistent with the rhodium photocleavage chemistry.[20] Also consistent with the photochemistry,[41] reduced but detectable cleavage is observed on the triplex at pH 5.5. Experiments were also conducted using the 5'-^{32}P-endlabeled hairpin strand (data not shown) and equivalent results were obtained; fragments comigrated with 3'-phosphate products. Relatively sequence-neutral cleavage was observed across the purine region and the U-loop of the triplex, but not in the hairpin duplex alone, as schematically illustrated in Figure 3.7. $Rh(phen)_2phi^{3+}$, therefore, is found to bind preferentially and, upon photoactivation, to promote cleavage throughout the triple helical region.

Importantly, the observation of cleavage across the hairpin loop of the triplex but not within the hairpin of the duplex must also indicate a conformational change in the loop which is a consequence of triplex formation. Triplex formation imparts hypersensitivity of the loop region to the rhodium complex. In earlier studies of folded RNAs,[2] $Rh(phen)_2phi^{3+}$ was found to promote cleavage within loops which were stacked

Figure 3.6. Autoradiogram showing cleavage of 3'-^{32}P-labeled RNA duplex (lanes 1-3, 7-9) and RNA R•R-Y triplex (lanes 4-6, 10-12) by $Rh(phen)_2phi^{3+}$ in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine pH 7.0 (lanes 1-6) and pH 5.5 (lanes 7-12). Lanes 1 & 7: labeled RNA duplex after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 10' at 313 nm. Lanes 2 & 8: labeled RNA duplex after incubation with 10 μM $Rh(phen)_2phi^{3+}$. Lanes 3 & 9: labeled RNA duplex after irradiation for 10' at 313 nm. Lanes 4 & 10: cleaved RNA R•R-Y triplex after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 10' at 313 nm. Lanes 5 & 11: labeled RNA R•R-Y triplex after incubation with 10 μM $Rh(phen)_2phi^{3+}$. Lanes 6 & 12: labeled RNA R•R-Y triplex after irradiation for 10' at 313 nm. Lanes 13, 14 & 15: A-, G- and U-specific reactions on the labeled RNA duplex respectively.

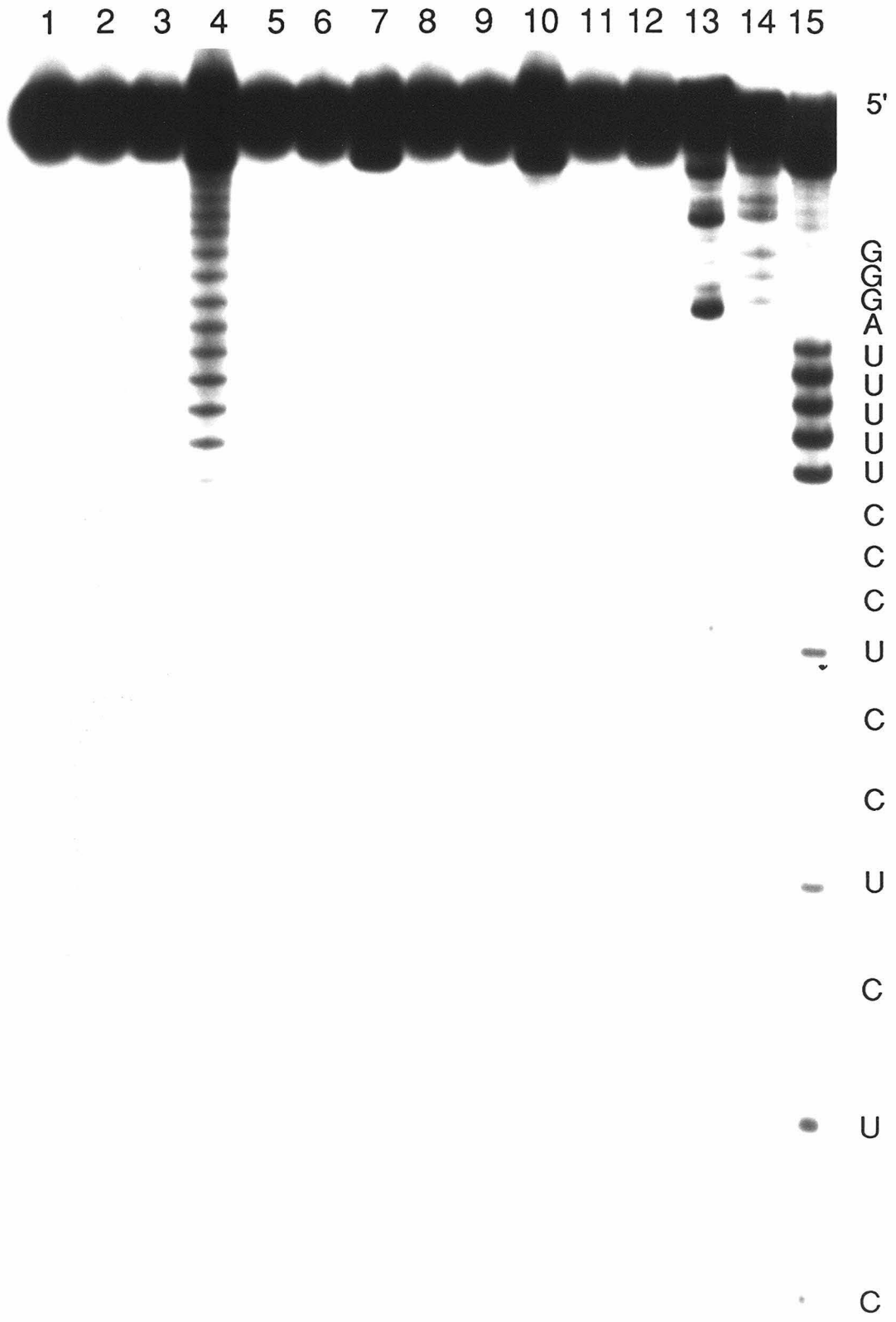
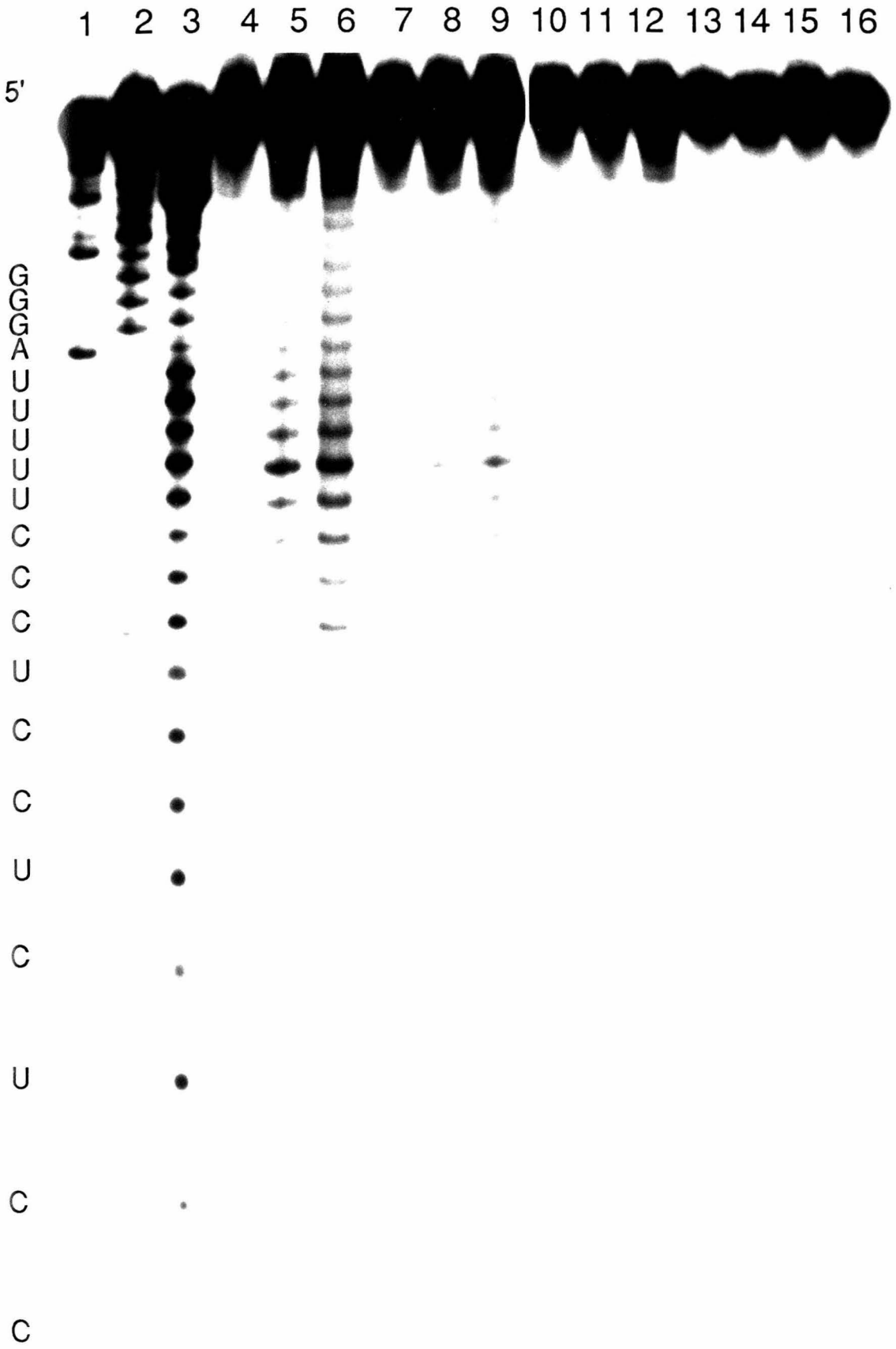

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
5'
GGGAUUUUUCCCCUCCCUUCU
C

Figure 3.7. Schematic illustration of light-induced $Rh(phen)_2phi^{3+}$ cleavage on the RNA duplex (A), RNA R•R-Y triplex (B), and RNA Y•R-Y triplex (C). Relative cleavage intensities (as determined by integration using ImageQuant) are depicted by the size of the arrows. Arrows with asterisks are approximations only, as quantitation was difficult due to the closeness of the bands. No cleavage is evident on the duplex strand alone, while relatively sequence-neutral cleavage is seen on the purine side of the Watson-Crick duplex on the R•R-Y motif. Cleavage is seen throughout the Watson-Crick duplex on the Y•R-Y motif.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
5'
GGGAUUUUUCCCUCCUCUCC

and structured. The result here suggests that formation of the triplex imparts structure to the loop, perhaps in exerting some torque onto the loop region which more tightly stacks the bases, promoting binding and cleavage by the metal complex.

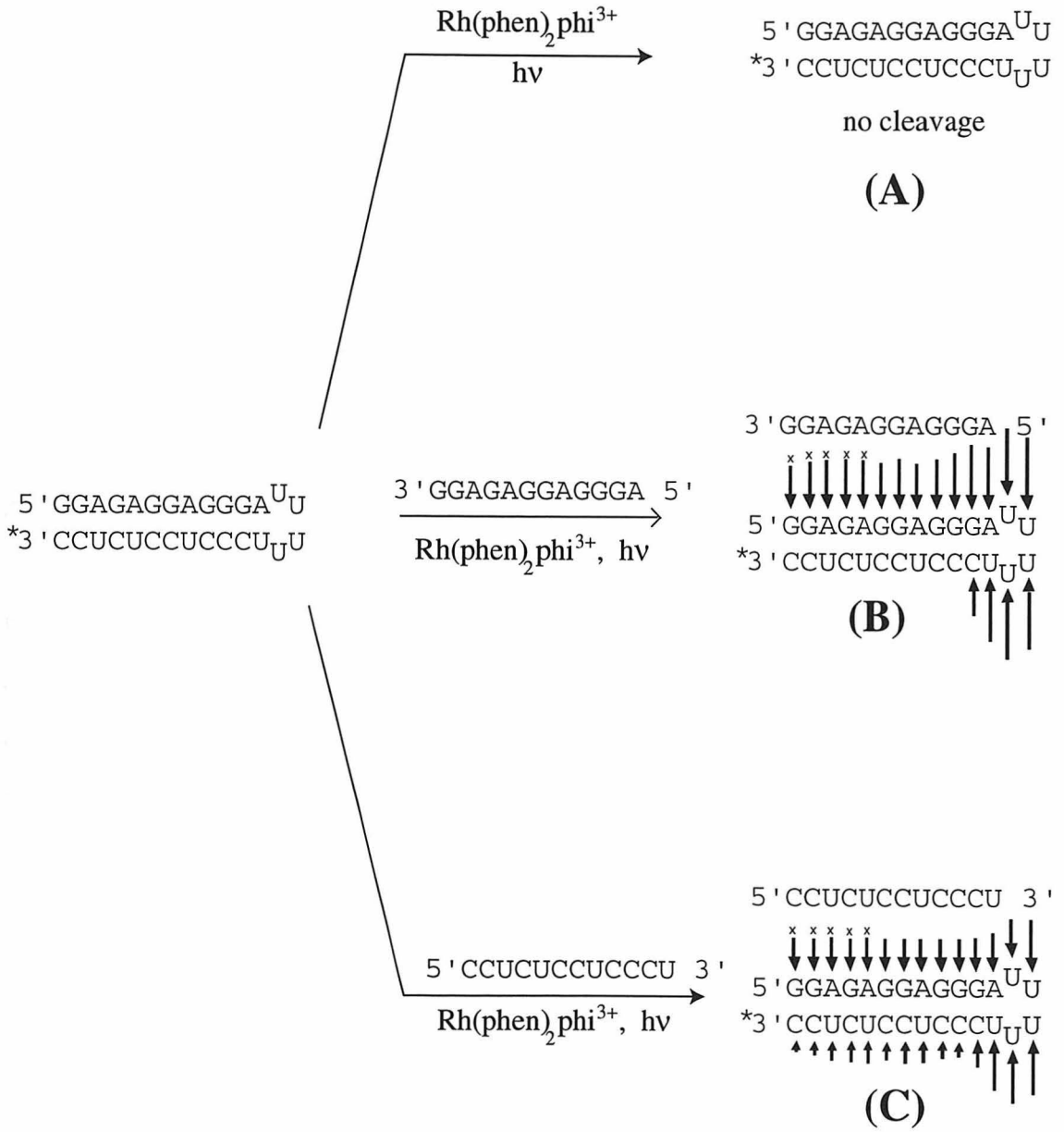
Analogous experiments were also conducted using a 3'-^{32}P-endlabeled purine strand and unlabeled hairpin duplex. Under comparable conditions, little cleavage is detected on the purine strand, both in the absence and presence of the hairpin duplex. At higher concentrations of metal complex and with longer irradiation times, some background cleavage on the labeled single strand may be detected, but no increase in cleavage intensity is observed upon addition of duplex. Consistent with earlier studies, $Rh(phen)_2phi^{3+}$, then, does not promote significant cleavage of single stranded RNA. Also, although the rhodium complex binds and cleaves the triplex along the Watson-Crick purine strand, no significant cleavage is found on the third purine strand.

3.3.3. Recognition of RNA Y•R-Y Triple Helices by $Rh(phen)_2phi^{3+}$

Figure 3.8 illustrates the shape-selective targeting of Y•R-Y triple helices by $Rh(phen)_2phi^{3+}$. Again, no significant cleavage is seen in the RNA duplex without the addition of the third strand. In the presence of the third Y strand, and in contrast to the R•R-Y motif, cleavage by $Rh(phen)_2phi^{3+}$ is observed across both the purine and pyrimidine strands of the Watson-Crick duplex. Phosphorimager quantitation with baseline correction showed a somewhat greater amount of cleavage on the purine side of the duplex than on the pyrimidine side (Figure 3.7). This observation is to be contrasted to those with the R•R-Y motif, in which no cleavage was apparent on the pyrimidine side of the hairpin duplex.

As in the R•R-Y motif, cleavage by the rhodium complex is also observed across the hairpin loop in the triplex but not the duplex. This result supports the idea that triplex formation with either motif leads to a conformational change or torque in the flanking loop region. No significant cleavage is evident on the third Y strand of the triplex in

Figure 3.8. Autoradiogram showing cleavage of 3'-^{32}P-labeled RNA duplex (lanes 13-16) and Y•R-Y triplex (lanes 5-12) by $Rh(phen)_2phi^{3+}$ in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine, pH 5.5 . Lanes 1-3: A-, G- and U- specific reactions respectively. Lanes 4-6: cleaved RNA Y•R-Y triplex, after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 5', 10', and 20' respectively at 313 nm. Lanes 7-9: cleaved RNA Y•R-Y triplex, after incubation with 5 μM $Rh(phen)_2phi^{3+}$ and irradiation for 5', 10', and 20' respectively at 313 nm. Lane 10-12: cleaved RNA Y•R-Y triplex, after incubation with 1 μM $Rh(phen)_2phi^{3+}$ and irradiation for 5', 10', and 20' respectively at 313 nm. Lane 13: labeled RNA duplex after incubation with 1 μM $Rh(phen)_2phi^{3+}$ and irradiation for 20' at 313 nm. Lane 14: labeled RNA duplex after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 20' at 313 nm. Lane 15: labeled RNA duplex with 10 μM $Rh(phen)_2phi^{3+}$ but without irradiation. Lane 16: labeled RNA duplex upon irradiation for 10' at 313 nm.


Rh(phen)2phi3+
hν
5'GGAGAGGAGGGAUU
*3'CCUCUCCUCCCUUU
no cleavage
(A)
5'GGAGAGGAGGGAUU
*3'CCUCUCCUCCCUUU
3'GGAGAGGAGGGA 5'
Rh(phen)2phi3+, hν
3'GGAGAGGAGGGA 5'
5'GGAGAGGAGGGAUU
*3'CCUCUCCUCCCUUU
(B)
5'CCUCUCCUCCCU 3'
Rh(phen)2phi3+, hν
5'CCUCUCCUCCCU 3'
5'GGAGAGGAGGGAUU
*3'CCUCUCCUCCCUUU
(C)

experiments using 3'-^{32}P-end-labeled pyrimidine strand, a finding similar to that in the R•R-Y motif case.

Cleavage by Rh(phen)$_2$phi^{3+} of the Y•R-Y triplex is seen to depend upon rhodium concentration, as well as time and wavelength for irradiation in a manner consistent with the photoreaction. Triplex cleavage is only observed at low pH (5.5) under conditions which stabilize the Y•R-Y triplex. As noted above, the photocleavage efficiency of Rh(phen)$_2$phi^{3+} is reduced at lower pH.[41] It should be noted that a quantitative comparison of the various efficiencies of rhodium complex cleavage of the Y•Y-R motif versus the R•R-Y motif is made difficult because of the lower yield of the Y•Y-R motif and the necessity for low pH conditions. Figure 3.8 summarizes in a histogram format the cleavage results on the RNA hairpin duplex and two RNA triplexes.

3.3.4. Dependence of Cleavage on Carrier and Salt Concentrations

When experiments varying the carrier oligonucleotide were carried out on the R•R-Y motif (which was chosen because the intensity of rhodium cleavage is greater than on the Y•R-Y motif), the same cleavage pattern and cleavage intensity was seen regardless of whether carrier oligonucleotide or tRNA^{Phe} is used. This suggests that Rh(phen)$_2$phi^{3+} has the same or greater binding affinity for the triplex as for tRNA^{Phe}, which is at least 3.5 x 10^5 M^{-1}. In addition, very little dependence of cleavage on carrier oligonucleotide concentration was observed when the concentration of carrier was varied from as little as 14 μM nt. to 100 μM nt., which is expected as the thermal stability of a triplex is Na^+-dependent, but nearly independent of concentration.[42] In this system, triplex formation was better stabilized by Mg^{2+} than by Na^+, and only a modest range of Na^+ concentrations were sampled (5 - 100 mM); cleavage by Rh(phen)$_2$phi^{3+} did not show a large degree of Na^+ dependence through this range. However, there was a dependence on polycations such as Mg^{2+} and spermidine, which promoted triplex formation and hence cleavage. When a competition experiment was carried out with the

analogous DNA oligomers as carrier, it was found that the intensity of cleavage in the RNA triplex was undiminished (data not shown). This again points to a high affinity for the RNA triplex substrate by $Rh(phen)_2phi^{3+}$.

3.3.5. Cleavage of RNA Triple Helices with RNases

Results of cleavage by the synthetic metal complex may also be compared to cleavage results using enzymatic probes for RNA single-strands. Both RNase *Phy*M (A-, G- and U-specific in the absence of urea) and RNase *Bc* (C- and U-specific)[43] produce cleavage primarily in the loop region of the hairpin, both alone and when bound to a third strand. As illustrated in Figure 3.9, the purine triplex motif confers a degree of protection on the exposed loop region, compared to that for the duplex. It is interesting to note that the addition of the purine third strand produces RNase cleavage in the purine side of the duplex, even though this RNase is not specific for purines. This cleavage could be due to the widening of the major groove of the duplex to accommodate the third strand, thus rendering it more susceptible to nucleases. We find that there is 85%, 53% and 86% digestion of the parent RNA band respectively in the duplex, R•R-Y and Y•R-Y lanes with identical amounts of RNase *Bc* and digestion time. On the other hand, there is little change in the RNase digestion pattern or the percentage digestion of the pyrimidine triplex motif.

Similar results are observed when the RNA motifs are digested with RNase *PhyM*, as shown in Figure 3.10. Again, RNase digestion of the loop region is decreased in the purine triplex compared to the duplex. Cleavage is also apparent on the purine side of the duplex when bound to the purine third strand, as opposed to the lack of cleavage in the duplex alone. Cleavage in the loop region of the duplex in the Y•R-Y triplex motif is also diminished compared to the duplex alone, but less so than with the R•R-Y triplex motif.

Figure 3.9. Autoradiogram showing cleavage of 3'-^{32}P-labeled RNA duplex (lanes 1-4), R•R-Y triplex (lanes 5-8) and Y•R-Y triplex (lanes 9-12) by RNase *Bc* in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine, pH 5.5. Lanes 1 & 2: cleaved RNA duplex, after incubation with 2 units RNase *Bc* for 15' and 60' respectively. Lane 3: cleaved RNA duplex, after incubation with 1 unit RNase *Bc* for 70'. Lane 4: labeled RNA duplex. Lanes 5-6: cleaved RNA R•R-Y triplex, after incubation with 2 units RNase *Bc* for 15' and 60' respectively. Lane 7: cleaved R•R-Y RNA triplex, after incubation with 1 unit RNase *Bc* for 70'. Lane 8: labeled RNA R•R-Y triplex. Lanes 9-10: cleaved RNA Y•R-Y triplex, after incubation with 2 units RNase *Bc* for 15' and 60' respectively. Lane 11: labeled RNA Y•R-Y triplex, after incubation with 1 unit RNase *Bc* for 70'. Lane 12: labeled RNA Y•R-Y triplex. Lanes 13, 14 & 15: A-, G- and U-specific reactions on the labeled RNA duplex respectively.

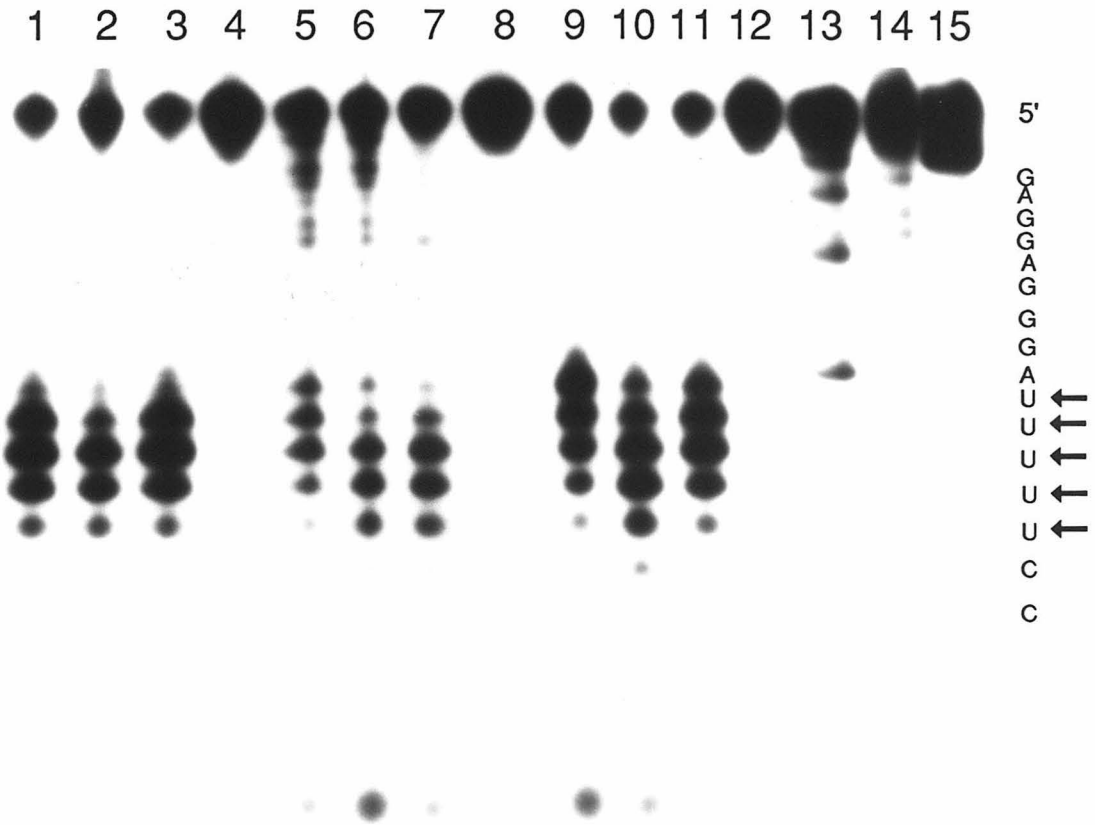
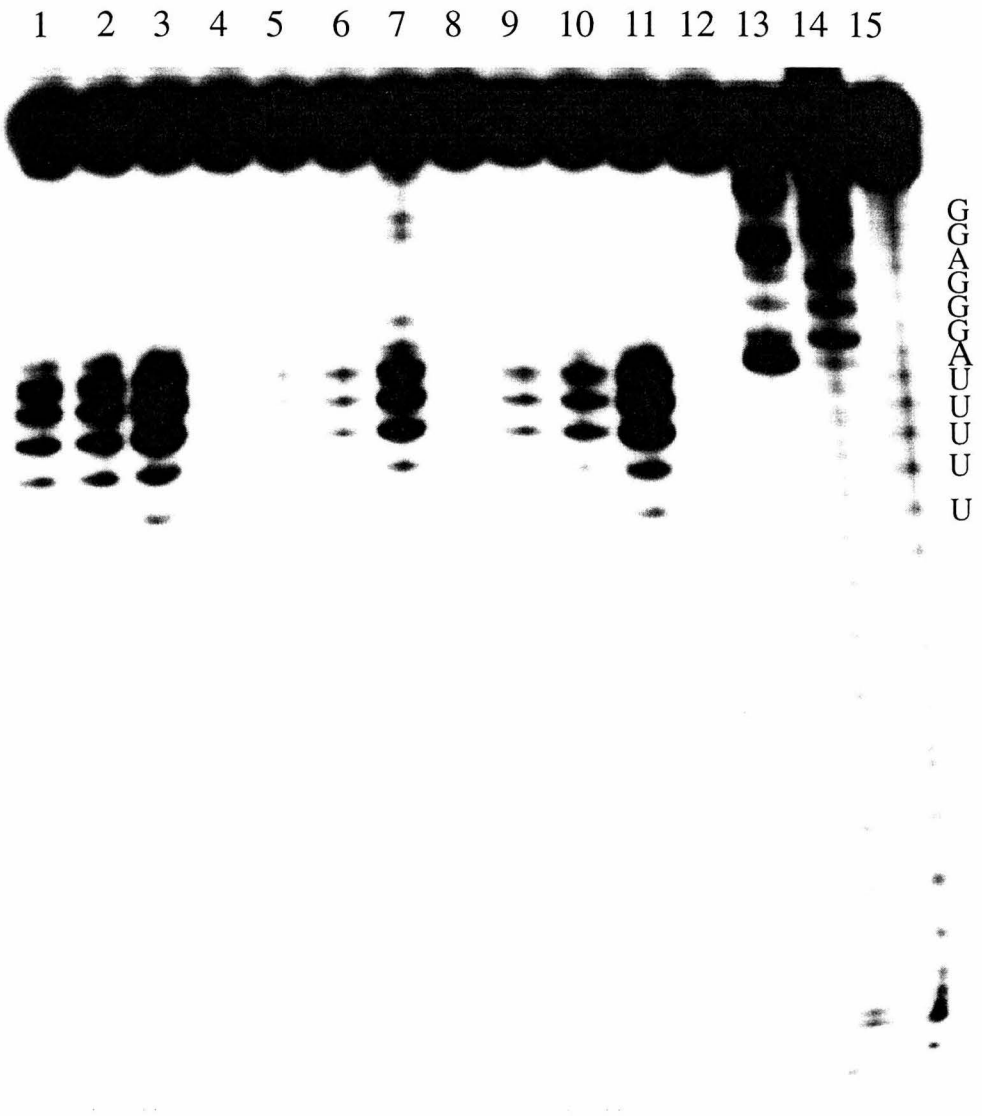

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
5'
GAGGAGGGAUUUUUCCC

Figure 3.10. Greyscale representation of an autoradiogram showing cleavage of 3'-^{32}P-labeled RNA duplex (lanes 1-4), R•R-Y triplex (lanes 5-8) and Y•R-Y triplex (lanes 9-12) by RNase *PhyM* in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine, pH 5.5. Lanes 1 & 2: cleaved RNA duplex, after incubation with 1 unit RNase *PhyM* for 15' and 60' respectively. Lane 3: cleaved RNA duplex, after incubation with 5 units RNase *PhyM* for 60'. Lane 4: labeled RNA duplex. Lanes 5-6: cleaved RNA R•R-Y triplex, after incubation with 1 unit RNase *PhyM* for 15' and 60' respectively. Lane 7: cleaved RNA R•R-Y triplex, after incubation with 5 units RNase *PhyM* for 60'. Lane 8: labeled RNA R•R-Y triplex. Lanes 9-10: cleaved RNA Y•R-Y triplex, after incubation with 1 unit RNase *PhyM* for 15' and 60' respectively. Lane 11: labeled RNA Y•R-Y triplex, after incubation with 5 units RNase *PhyM* for 60'. Lane 12: labeled RNA Y•R-Y triplex. Lanes 13, 14 & 15: A-, G- and U-specific reactions on the labeled RNA duplex respectively.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
GGAGGGAUUUU
U

This differential cleavage of the hairpin loops in the duplex compared to the triplex form complements the results seen with $Rh(phen)_2phi^{3+}$. Thus, a triplex-induced conformational change in the loop region can be detected by both metal complex and RNase cleavage.

3.3.6. Recognition of DNA Triple Helices by $Rh(phen)_2phi^{3+}$

Rhodium-promoted cleavage was also examined on the analogous DNA triplexes formed by annealing together the DNA hairpin duplex with either a pyrimidine or purine DNA third strand. When the labeled DNA duplex alone is irradiated in the presence of metal complex, one observes the expected cleavage by $Rh(phen)_2phi^{3+}$ of B-form DNA due to its open major groove.[20] Cleavage of the pyrimidine strand of the duplex is, however, seen to be somewhat greater than of the purine strand. Upon addition of the R third strand, however, there is a significant loss of cleavage on the pyrimidine side of the hairpin; cleavage on the purine side of the hairpin does not change substantially. Interestingly, addition of the Y third strand, in contrast, leads to a substantial reduction in cleavage on the purine side of the hairpin strand, while cleavage on the pyrimidine side remains largely intact. As on the RNA oligomers, there is a background of non-specific cleavage at high metal complex and irradiation times on the labeled single-stranded oligomers alone, with no change upon addition of the unlabeled hairpin duplex. These results are shown in Figure 3.11 and schematically illustrated in histogram form in Figure 3.12. These data resemble those obtained with the RNA triple helices. However, the result is less striking in comparing the duplex to triplexes owing to the high level of cleavage obtained on the DNA duplex alone. This is consistent with the normal recognition of the B-form helix by $Rh(phen)_2phi^{3+}$.

Figure 3.11. Autoradiogram showing cleavage of 3'-^{32}P-labeled DNA duplex (lanes 1-4), R•R-Y triplex (lanes 5-8) and Y•R-Y triplex (lanes 9-12) by $Rh(phen)_2phi^{3+}$ in 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM spermidine, pH 5.5. Lanes 1 & 2: cleaved DNA duplex, after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 2' and 10' at 313 nm respectively. Lane 3: labeled DNA duplex with 10 μM $Rh(phen)_2phi^{3+}$ but without irradiation. Lane 4: labeled DNA duplex upon irradiation for 10' at 313 nm. Lanes 5-6: cleaved DNA R•R-Y triplex, after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 2' and 10' at 313 nm respectively. Lane 7: labeled DNA R•R-Y triplex with 10 μM $Rh(phen)_2phi^{3+}$ but without irradiation. Lane 8: labeled DNA R•R-Y triplex upon irradiation for 10' at 313 nm. Lanes 9-10: cleaved DNA Y•R-Y triplex, after incubation with 10 μM $Rh(phen)_2phi^{3+}$ and irradiation for 2' and 10' at 313 nm respectively. Lane 11: labeled DNA Y•R-Y triplex with 10 μM $Rh(phen)_2phi^{3+}$ but without irradiation. Lane 12: labeled DNA Y•R-Y triplex upon irradiation for 10' at 313 nm. Lanes 13 & 14: A+G and C+T Maxam Gilbert sequencing reactions respectively.

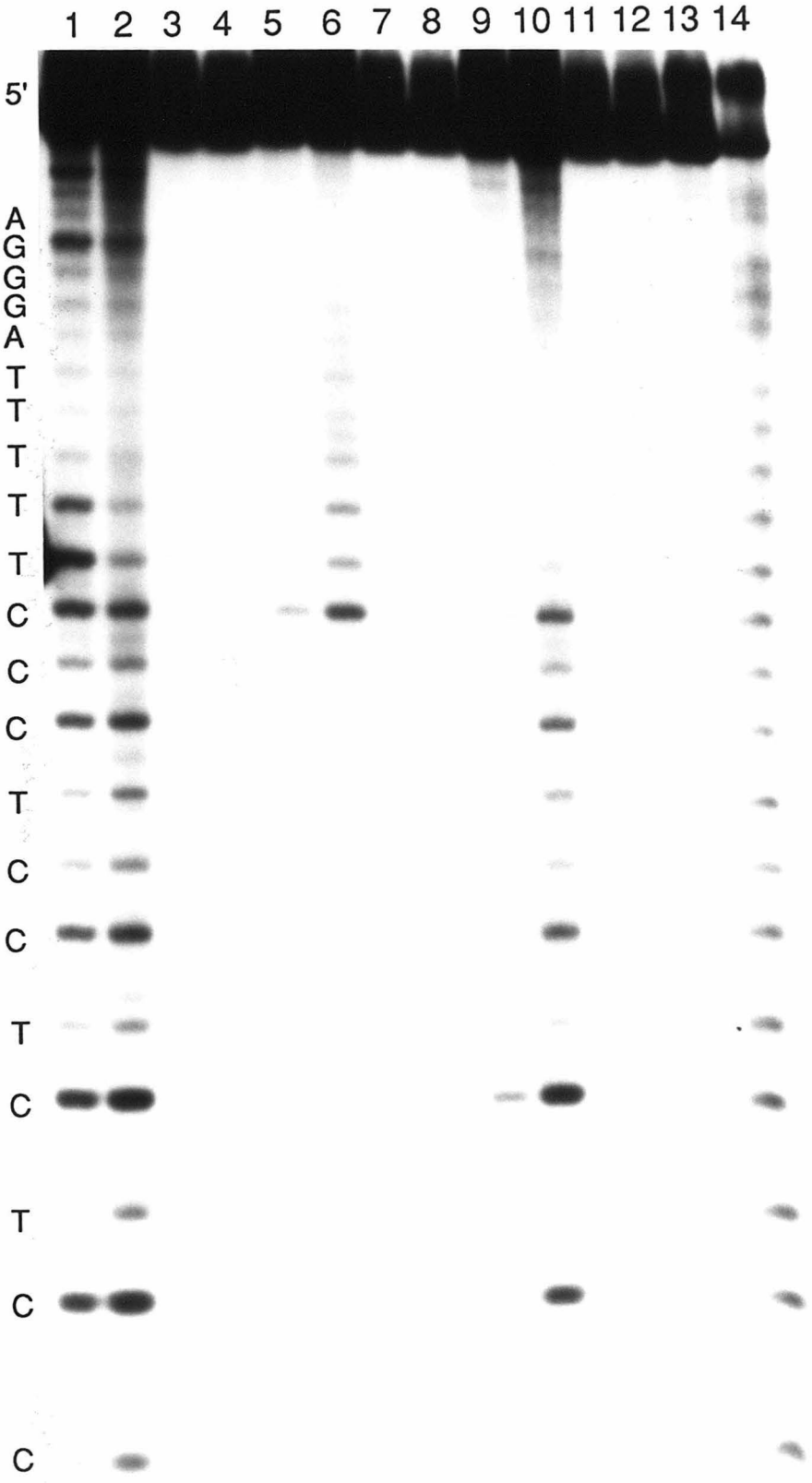
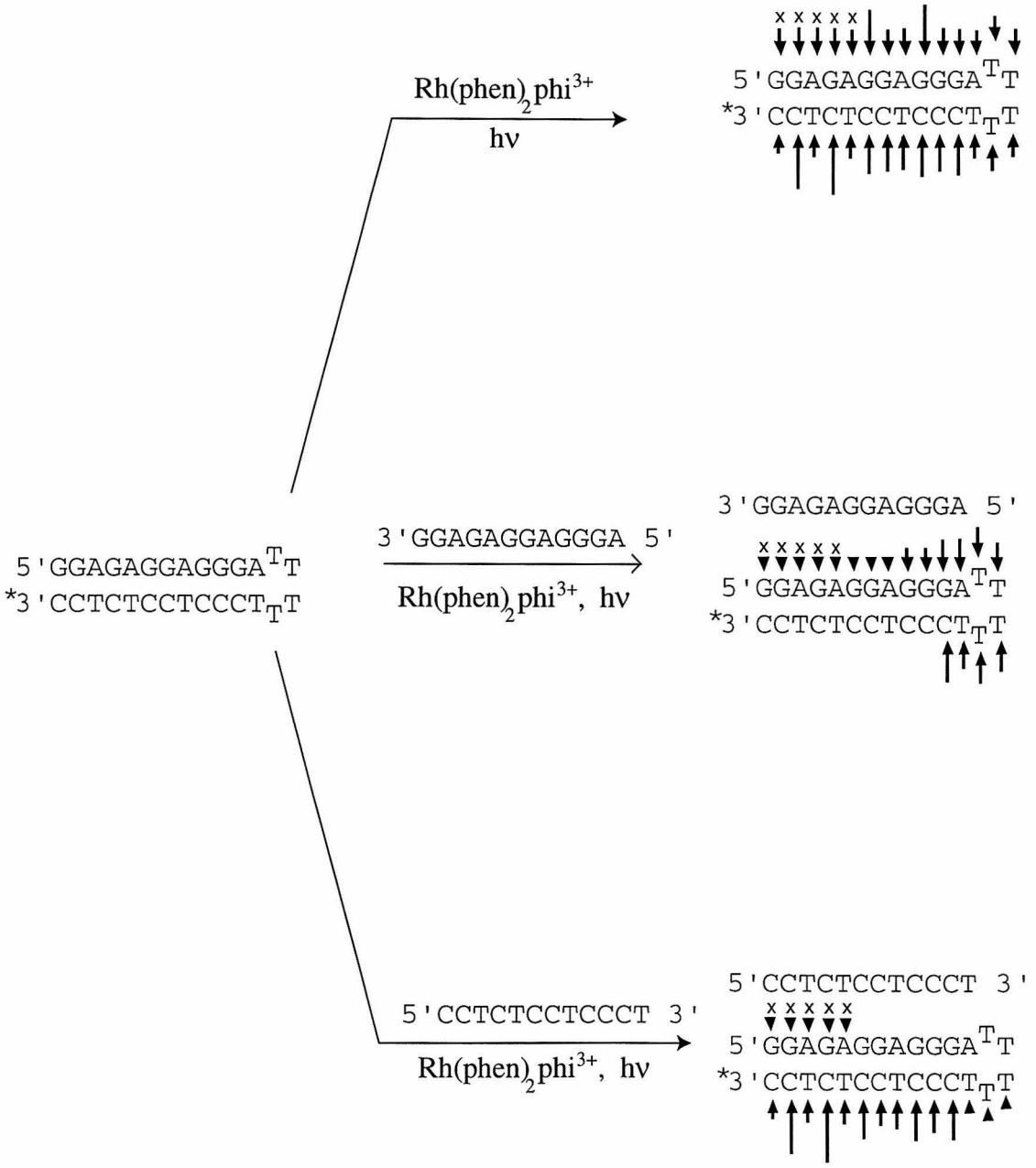

1 2 3 4 5 6 7 8 9 10 11 12 13 14
5'
A
G
G
G
A
T
T
T
T
T
C
C
C
T
C
C
T
C
T
C
C

Figure 3.12. Schematic of $Rh(phen)_2phi^{3+}$ cleavage on the DNA duplex (A), DNA R•R-Y triplex (B), and DNA Y•R-Y triplex (C). Relative cleavage intensities (as determined by integration using ImageQuant) are depicted by the size of the arrows. Arrows with asterisks are approximations only. Cleavage is evident throughout the duplex strand alone, while cleavage is seen on the purine side of the Watson-Crick duplex on the R•R-Y motif. Cleavage is seen on the pyrimidine side of the Watson-Crick duplex on the Y•R-Y motif.


Rh(phen)2phi3+
hν
5'GGAGAGGAGGGA T T
*3'CCTCTCCTCCCT T T
3'GGAGAGGAGGGA 5'
Rh(phen)2phi3+, hν
5'CCTCTCCTCCCT 3'

3.4. Discussion

3.4.1. Recognition of RNA Triple Helices by Rhodium Complexes

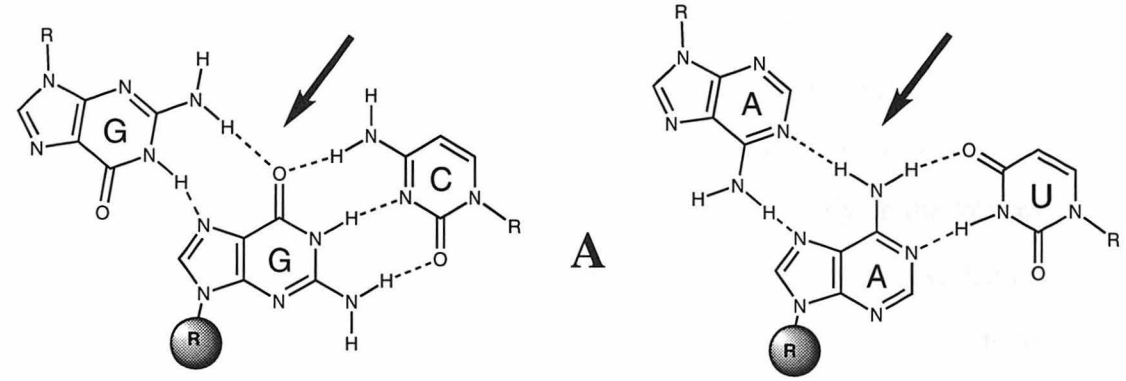
The results described here indicate that $Rh(phen)_2phi^{3+}$ targets RNA triple helices under conditions where significant cleavage is not evident on corresponding RNA duplexes. This recognition is structure-specific rather than sequence-specific, as cleavage across the triple helical segment in a sequence-neutral fashion is observed. $Rh(phen)_2phi^{3+}$ generally appears to be sensitive to the width of the major groove, and less so to the bases themselves. Hence, $Rh(phen)_2phi^{3+}$ may be particularly useful in probing RNA tertiary structures.

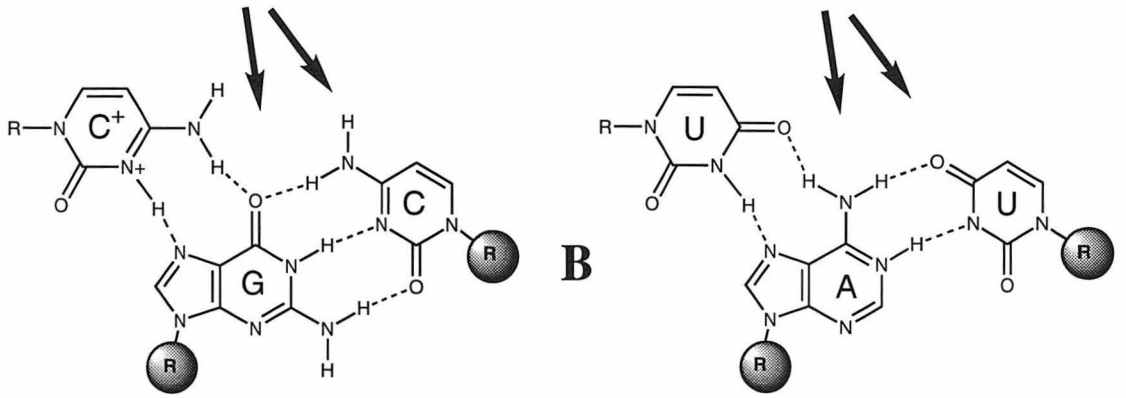
DNA triple helices are similarly targeted by the rhodium complex. In this case, however, cleavage is also evident on B-DNA duplexes, and $Rh(phen)_2phi^{3+}$ is thus less useful as a probe. Since the Watson-Crick basepaired duplex in the DNA triplex compared to in the duplex alone is still essentially B-form in nature,[14] the rhodium complex is not sensitive to this lack of change in the width of the major groove.

We attribute the novel RNA recognition characteristics of $Rh(phen)_2phi^{3+}$ to its shape-selective intercalation in the major groove. In our structural model, a triple helix could provide a platform of bases for intercalative stacking by the complex from the major groove side. In contrast, an A-form duplex would provide a poor target for intercalative stacking form the major groove side, since, in an A-form helix, the major groove is narrowed and deepened.

RNA triple helices of both motifs, R•R-Y and Y•R-Y, are targeted by $Rh(phen)_2phi^{3+}$, although the cleavage characteristics across the triple helices differ for the two motifs. For the R•R-Y motif, cleavage is observed preferentially on the purine Watson-Crick strand, while in the Y•R-Y motif, cleavage is observed on both Watson-Crick strands. The difference in cleavage pattern across the triplex between the two motifs might be viewed in the context of a simple intercalation model. Figure 3.13 illustrates the triplex hydrogen bonding pattern considered for the two motifs as well as

Figure 3.13. Schematic of cleavage by $Rh(phen)_2phi^{3+}$ of the R•R-Y motif (A) and the Y•R-Y motif (B). Stippled sugars show the nucleotides cleaved. Arrows show the probable direction of intercalation of the phi ligand.


A
G
G
C
A
A
U


B
C+
G
C
U
A
U

the direction of intercalation suggested by the cleavage results. The differential cleavage on the two motifs might best be understood in terms of how the phi intercalator stacks on the bases. In the R•R-Y motif, the four aromatic rings of the two purine bases (on the Hoogsteen and Watson Crick pair) would provide a larger aromatic surface for stacking by the phi ligand than the three aromatic rings of the Watson Crick base pair. In addition, in the case of the Y•R-Y motif, the Hoogsteen base-paired cytosine is protonated, leading to some charge repulsion with the positively charged metal complex. To a first approximation, therefore, cleavage could be distributed fairly evenly on the Watson-Crick pair, but not on the Hoogsteen strand. It is noteworthy in this context that cleavage results from a hydrogen abstraction reaction by the phi with the ribose, not with the nucleic acid base. Nonetheless, molecular modeling on a triple helix, assuming all bases to be in the *anti* configuration, reveals in three dimensions that access to the sugar correlates directly with the orientation on the base stack.

One observation with respect to the RNA triple helix which emerges from these studies involves the conformational change apparent in the hairpin loop upon triplex formation. Although the tetra-uridine hairpin loop of the RNA duplex is not cleaved by the metal complex, addition of either third strand, with resultant triple helix formation, leads to a structural change in the loop which renders the loop hypersensitive to the metal complex. Earlier studies with $Rh(phen)_2phi^{3+}$ have suggested that only tightly stacked or structured loops are targeted by the metal complex.[3] The hypersensitivity observed may then point to a helical unwinding that accompanies the formation of the triple helix. Such unwinding would generate a torque in the neighboring loop and alter how the bases are packed within the loop. These changes are also detected by probing with RNases where in a complementary experiment, the loop region of the duplex when the third strand is bound to form the triplex is found to be less susceptible to RNase digestion than is the loop in the duplex alone.

This suggests a structural resemblance between the two triplex forms. The pattern of cleavage across the triplex does differ between the two motifs, but this difference may result from a difference in how the metal complex is oriented in the helical stack rather than a structural difference between the triple helices (*vide supra*). It is difficult to make a comparison of the photoreactions of either motif at pH 5.5 because the Y•R-Y motif is formed in lower yield than the R•R-Y motif even at that pH. Without accounting for the lower yield of formation Y•R-Y motif, it appears that the Y•R-Y motif is cleaved less than half as efficiently as the R•R-Y motif, based on phosphorimager quantitation. Comparisons between RNA and DNA triple helices are more difficult to make because of the inherently strong reactivity of the metal complex with the B-form duplex alone.

3.4.2. Rh(phen)$_2$phi^{3+} as a Probe of Naturally Folded RNAs

Our results on the synthetic triple helices provide a strong foundation for understanding the cleavage patterns already obtained on different tRNAs as well as for new mapping studies on folded RNAs. For example, tRNA^{Phe} contains a central triple base interaction, m^{7}G46-G22-C13,[15] and strong cleavage is observed by Rh(phen)$_2$phi^{3+} at G22 and m^{7}G46, and with none at C13. This cleavage pattern is fully consistent with the results on the R•R-Y synthetic triplex, as cleavage is seen on the purine strand of the duplex, even though the triplex does not contain the methylated G. Interestingly, when G46 is mutated to a C, cleavage at C46 is lost,[3] while with the mutation G46A-G22A-C13U, which preserves the R•R-Y triplex motif, cleavage is maintained. In the case of the analogous triple A46•G22-Ψ13 on tRNA^{Asp}, cleavage is centered on the Hoogsteen R•R base pair. Again, this is consistent with the cleavage seen on the R•R-Y synthetic triplex. Thus generally, results obtained on the tRNAs may be understood in the context of the synthetic triple helices.

These studies therefore lay the groundwork for probing folded RNAs which have not yet been structurally characterized in sufficient detail. Rh(phen)$_2$phi^{3+} serves as a

chemical probe to delineate positions where nucleotides form triply bonded arrays, though, of course, some such arrays in the interior of larger folded RNAs would not be accessible. Nonetheless, $Rh(phen)_2phi^{3+}$ is expected to cleave at selective sites on a folded RNA which are not part of an A-form duplex. Instead, these sites would represent those in which the major groove is made accessible to intercalative stacking, either through a bulged or mismatched opening which disrupts the A-form duplex or through a tertiary interaction involving a triply bonded array. In concert with mutational studies, then, the application of $Rh(phen)_2phi^{3+}$ to probe RNA now represents a valuable strategy to delineate RNA tertiary structure.

References

1. Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839-2841.
2. Chow, C. S.; Hartmann, K. M.; Rawlings, S. L.; Huber, P. W.; Barton, J. K. *Biochemistry* **1992**, *31*, 3534-3542.
3. Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972-982.
4. Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *107*, 5528-5529.
5. Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927-7929.
6. Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.
7. Beal, P.; Dervan, P. B. *Science* **1991**, *251*, 1360.
8. Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63-80.
9. Chastain, M.; Tinoco Jr., I. *Nucleic Acids Res.* **1992**, *20*, 315-318.
10. Broitman, S. L.; Im, D. D.; Fresco, J. R. *Proc. Nat. Acad. Sci. U.S.A.* **1987**, *84*, 5120 -5124.
11. Letai, A. G.; Palladino, M. A.; Fromm, E.; Rizzo, V.; Fresco, J. R. *Biochemistry* **1988**, *27*, 9108-9112.

12. Mirkin, S. M.; Lyamichev, V. I.; Drushlyak, K. N.; Dobrynin, V. N.; Filipov, S. A.; Frank-Kamenetskii, M. D. *Nature* **1987**, *330*, 495-497.

13. de los Santos, C.; Rosen, M.; D., P. *Biochemistry* **1989**, *28*, 7282-7289.

14. Vlieghe, D.; Van Meervelt, L.; Dautand, A.; Gallois, B.; Précigoux, G.; Kennard, O. *Science* **1996**, *273*, 1702-1705.

15. Kim, S.-H.; Sussman, J. L.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Wang, A., H.; Seeman, N. C.; Rich, A. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 4970-4974.

16. Quigley, G. J.; Rich, A. *Science* **1976**, *194*, 796-806.

17. Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K. H.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678-1685.

18. David, S. S.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 2984-2985.

19. Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520-4522.

20. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303-2312.

21. Hudson, B. P.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1995**, *117*, 9379-9380.

22. Campisi, D.; Morii, T.; Barton, J. K. *Biochemistry* **1994**, *33*, 4130-4139.

23. Jenkins, Y.; Friedman, A. E.; Turro, N. J.; Barton, J. K. *Biochemistry* **1992**, *31*, 10809 -10816.

24. Scaria, P. V.; Shafer, R. H. *J. Biol. Chem.* **1991**, *266*, 5417-5423.

25. Mergny, J.-L.; Collier, D.; Rougée, M.; Montenay-Garestier, T.; Hélène, C. *Nucleic Acids Res.* **1991**, *19*, 1521-1526.

26. Tuite, E.; Norden, B. *Bioorganic & Med. Chem.* **1995**, *3*, 701-711.

27. Marchand, C.; Bailly, C.; Nguyen, C. H.; Bisagni, E.; Garestier, T.; Hélène, C.; Waring, M. J. *Biochemistry* **1996**, *35*, 5022-5032.

28. Mergny, J.-L.; Duval-Valentin, G.; Nguyen, C. H.; Perroualt, L.; Faucon, B.; Rougée, M.; Montenay-Garestier, T.; Bisagni, E.; Hélène, C. *Science* **1992**, *256*, 1681-1684.

29. Collier, D. A.; Mergny, J.-L.; Thuong, N. T.; Hélène, C. *Nucleic Acids Res.* **1991**, *19*, 4219-4224.

30. Mouscadet, J.-F.; Ketterlé, C.; Coulaouic, H.; Carteau, S.; Subra, F.; Le Bret, M.; Auclair, C. *Biochemistry* **1994**, *33*, 4187-4196.

31. Kane, S. A.; Hecht, S. M.; Sun, J. S.; Garestier, T.; Helene, C. *Biochemistry* **1995**, *34*, 16715-16724.

32. Lim, A. C.; Barton, J. K. *Biochemistry* **1993**, *32*, 11029-11034.

33. Weeks, K. M.; Crothers, D. M. *Science* **1993**, *261*, 1574-1577.

34. Roberts, R. W.; Crothers, D. M. *Proc. Nat. Acad. Sci. U.S. A.* **1991**, *88*, 9397-9401.

35. Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463-1466.

36. England, T. E.; Uhlenbeck, O. C. *Nature* **1978**, *275*, 560-561.

37. Chow, C. S. Doctoral thesis, Dept. of Chemistry, California Institute of Technology: Pasadena, 1992.

38. Peattie, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1760.

39. Sambrook, J.; Fritsch, E. F.; Manniatis, T. *Molecular Cloning: a Laboratory Manual*; 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.

40. Semerad, C. L.; Maher, L. J. *Nucleic Acids. Res.* **1994**, *22*, 5321-5325.

41. Krotz, A. H.; Kuo, L. Y.; Barton, J. K. *Inorg. Chem.* **1993**, *32*, 5963-5974.

42. Plu, G. E.; Park, Y.-W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436-9440.

43. Kuchino, Y.; Nishimura, S. *Meth. Enzym.* **1989**, *180*, 154-163.

Chapter 4:

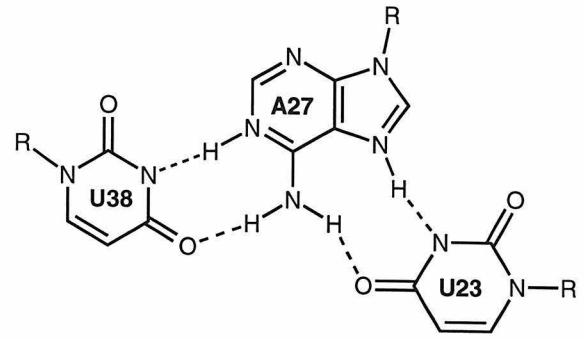
Specific Cleavage of the Human Immunodeficiency Virus TAR RNA by Rhodium Complexes

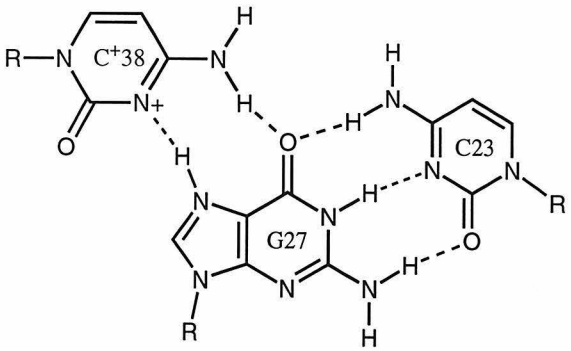
4.1. Introduction

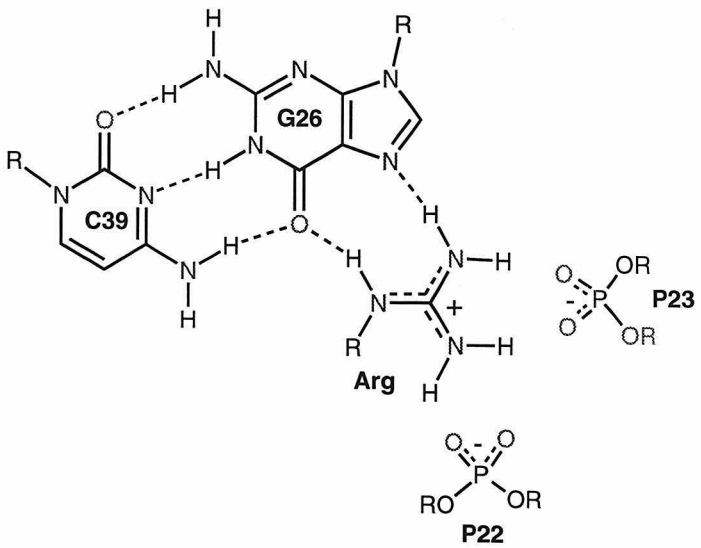
The binding of the viral Tat protein (*trans*-activating transcriptional activator) to the RNA target sequence TAR (*trans*-activation response), present at the 5'- end of all viral mRNAs, is required for the transcription of the HIV-1 retrovirus.[1-4] The RNA-binding domain of Tat consists of an arginine-rich basic domain of nine amino acids, RKKRRQRRR (residues 49 to 57).[4,5] A single arginine in this basic region of Tat binds specifically to TAR,[6] as does free arginine alone.[7] Certain nucleotides in TAR (U23, G26·C39, and A27·U38) have been shown to be involved in transactivation,[3,8] specific peptide binding,[4-6] and arginine binding,[6] suggesting that arginine recognizes TAR in the same fashion both as the free amino acid as well as when in a peptide.

NMR studies show that the three-nucleotide bulge in TAR changes conformation upon the binding of argininamide, a tight-binding arginine analog.[9] In the absence of argininamide, the two stem regions are base-paired, A-form helices, while U23 at the 5' end of the bulge stacks over A22. U25 is looped out, creating a major distortion of the phosphate backbone between C24 and G26.[10] With the addition of argininamide, the two A-form stem regions stack coaxially, while the nucleotides on the bulge region unstack. This is consistent with hydrodynamic studies that show that the addition of argininamide leads to essentially complete straightening of the previously bent helix.[11] Thus, upon binding of a single arginine group to a specific site, the bulge region changes conformation to form a U23•A27-U38 base triple, which stabilizes arginine hydrogen bonding to G26 and phosphates[9] (Figure 4.1). This putative base triple is supported by the finding that when the U•A-U triple is substituted by the isomorphous C^+•G-C base triple, the identical RNA structure is formed upon binding by arginine.[9]

Figure 4.1. Schematics of the U23•A27-U38 base triple (top), the isomorphous C^{+}23•G27-C38 triple (center), and the G26-C39 interaction with arginine (bottom).


A27
U38
U23


C+38
C23
G27


G26
C39
Arg
P23
P22

The NMR solution structure of the HIV-2 TAR-argininamide complex has also been determined.[12] As previously observed in the case of HIV-1 TAR-argininamide complex, the two A-form stems stack co-axially and the critical U23 is located in the major groove. Base-triple formation helps open the major groove to increase the accessibility of G26 to hydrogen bond donors from the guanidinium group of argininamide.[12] Thus, the formation of this base triple represents a motif which facilitates recognition of the RNA by the Tat peptide in both the case of HIV-1 and the HIV-2 TAR RNA.

Thus, the base triple in the bulge of the HIV-1 WT TAR RNA represents an attractive target for recognition by the metal complex $Rh(phen)_2phi^{3+}$ (phen = phenanthroline, phi = 9,10-phenanthrenequinone diimine). This complex binds to DNA duplexes by intercalation in the major groove,[13] and effects strand cleavage upon photoactivation.[14] As shown already in this thesis, $Rh(phen)_2phi^{3+}$ binds to RNA sites which have an open major groove. Thus, $Rh(phen)_2phi^{3+}$ recognizes areas of tertiary interaction which established the folded structure of the RNA in tRNAs and tRNA mutants,[15] and synthetic triple helices (Chapter 3). It has also been shown that $Rh(phen)_2phi^{3+}$ recognizes the bulge region of HIV[16] and BIV (Chapter 5) TAR RNA.

Another metal complex which would reveal details of the tertiary structure of this RNA molecule is $Rh(MGP)_2phi^{5+}$ (MGP=4-guanidylmethyl-1,10-phenanthroline; phi=phenanthrenequinone diimine). This complex has the advantage of containing both a guanidinium moiety and an intercalating ligand, which could make it a more selective and avidly-binding probe of RNA structure. 1Λ-$Rh(MGP)_2phi^{5+}$ recognizes the DNA sequence 5'-CATATG-3' site-specifically by unwinding the DNA helix and making direct contacts between guanine and the guanidinium functionalities.[17] These guanidinium arms could serve as mimics for the functional arginine group on the Tat peptide, thus enhancing the selectivity and affinity of the metal complex for the TAR RNA, and making it a better mimic for proteins. Since MGP is an asymmetric ligand, it also provides flexibility in the positioning of the guanidinium groups relative to the intercalative phi ligand[17] when the

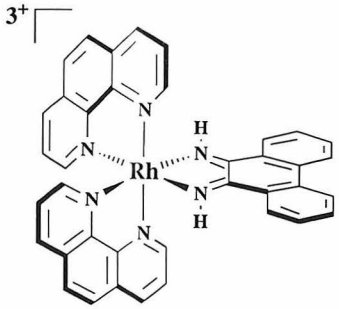
resolved enantiomers are used to probe the RNA structure. In addition, resolved enantiomers of $Rh(GEB)_2phi^{5+}$ (GEB = 4-(2-guanidylethyl)-4'-methyl-2,2'-bipyridine) can be used in order to probe the effect of different guanidinium-containing arm lengths and stereocenters on the recognition of RNA by these small molecules. These metal complexes are shown in Figure 4.2.

Here we describe the specific binding by these complexes to TAR RNA. The sites of strongest photocleavage, and thus binding, is U23 in the bulge region of the RNA and U40 opposite it. The Δ-enantiomer of $Rh(phen)_2phi^{3+}$ recognizes the wild type TAR RNA preferentially over the Λ-enantiomer. Differences in photocleavage sites between the native TAR RNA and mutants suggest these metal complexes bind within the bulge region in each case, but that the metal complex induces the formation of a U23•A27-U38 base triple in the wild type RNA, thus tightening up the RNA structure. The metal complex affinity for the RNA is enhanced slightly by the addition of guanidinium groups on the ancillary ligands of the rhodium complexes. Photocleavage, binding affinity and the circular dichroism spectra together suggest that the guanidinium groups on the derivatives provide non-specific binding affinity for the RNA by making contacts with the phosphate groups on the backbone, but do not make specific contacts with the bases. In addition, photocleavage competition studies have shown that despite their lower binding affinities, these complexes compete successfully with the native peptides for binding sites upon the RNA molecule.

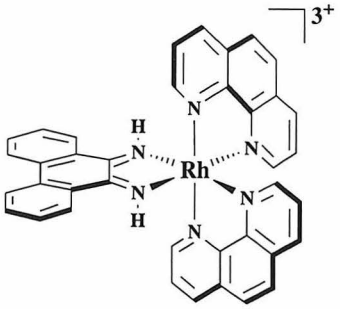
4.2. Experimental Methods

RNA preparation. The HIV TAR and TAR1 RNAs were prepared by *in vitro* transcription[18] using synthetic DNA templates and T7 RNA polymerase (Pharmacia). DNA primer and template strands were synthesized on an ABI DNA/RNA 392 synthesizer, purified by reverse phase HPLC, detritylated, and purified again by HPLC. The RNA products from the transcription reactions were precipitated and washed with EtOH. The RNA oligomers were gel purified on an 8 or 10% polyacrylamide denaturing gel, located

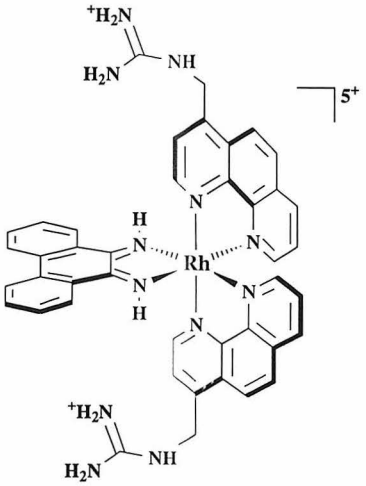
Figure 4.2. Schematics of rhodium complexes. Δ-$Rh(phen)_2phi^{3+}$, Λ-$Rh(phen)_2phi^{3+}$, 1Δ-$Rh(MGP)_2phi^{5+}$, 2Δ-$Rh(MGP)_2phi^{5+}$, 3Δ-$Rh(MGP)_2phi^{5+}$ and 3Δ-$Rh(GEB)_2phi^{5+}$.



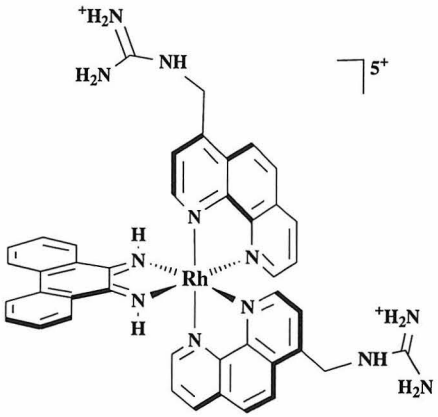
Δ–Rh(phen)$_2$phi^{3+}



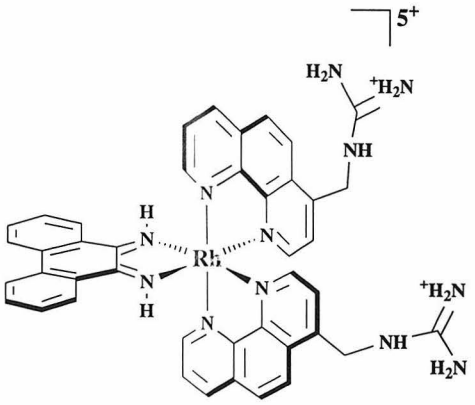
Λ–Rh(phen)$_2$phi^{3+}



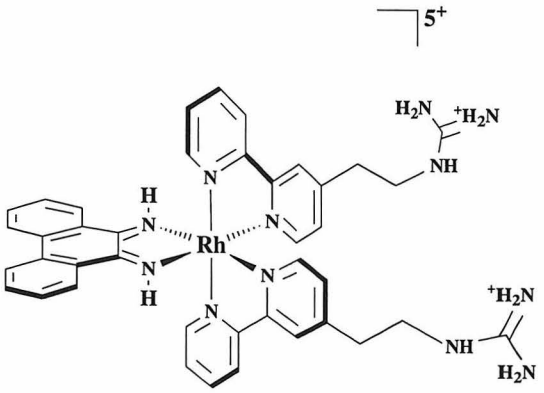
1 -Rh(MGP)$_2$phi^{3+}



2 -Rh(MGP)$_2$phi^{3+}



3 -Rh(MGP)$_2$phi^{5+}



3 -Rh(GEB)$_2$phi^{5+}

with UV shadowing, excised and then eluted from the gel using an Elutrap device (Schleicher & Schuell). The other RNA oligomers were synthesized on the DNA/RNA synthesizer, and gel purified in the same manner. After precipitation, the RNA was stored frozen in 10 mM Tris HCl buffer pH 7.0 at -20°C. The RNA oligomers were quantitated using UV-visible spectroscopy. RNA extinction coefficients were calculated to be 8900 $M^{-1}cm^{-1}$ /nucleotide for TAR WT, TAR1, TAR2 and TAR4. The extinction coefficients for TAR3 and TAR5 were 8700 $M^{-1}cm^{-1}$ /nucleotide. The RNA oligomers were 3'-end-labeled with cytidine 3',5'-[5'-^{32}P]-bisphosphate using T4 RNA ligase[19] or 5'-end-labeled with [γ-^{32}P]-ATP using T4 polynucleotide kinase.[20] They were purified and recovered as described for the unlabeled RNA. The eluted RNA oligomers were ethanol-precipitated twice and stored frozen in 10 mM Tris-HCl, pH 7.5. $Rh(phen)_2phi^{3+}$ solutions were prepared fresh in 10 mM Tris-HCl pH 7.0.

Rhodium photocleavage. $[Rh(phen)_2phi]Cl_3$ was prepared as described earlier.[21] $[Rh(DIP)_3]Cl_3$ was obtained from I. Lee. $[Rh(MGP)_2phi]Cl_5$ and $[Rh(GEB)_2phi]Cl_5$ were obtained from R. Terbrueggen as resolved enantiomers.[22] The enantiomers had been resolved by cation-exchange chromatography on a Sephadex SPC-25/potassium antimonyl tartrate column or by HPLC separation on a Chiracel column (Chiral Technologies).[17] $Rh(phen)_2phi^{3+}$ and the other metal stock solutions were freshly prepared in 10 mM Tris-HCl, pH 7.5 and pH 5.5 respectively. Photocleavage samples were prepared in Tris buffer (50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0 or 5.5) or phosphate buffer (10 mM NaH_2PO_4, 50 mM NaCl, pH 6.5). A typical 20 µL irradiation sample consisted of 20000-40000 cpm renatured ^{32}P-labeled RNA, 0.001-10 µM rhodium complex, 50 or 100 µM nucleotides carrier TAR or $tRNA^{Phe}$ in aqueous buffer. The photocleavage was also done in the presence of 10 - 50 µM argininamide or 10 mM $MgCl_2$. Rhodium complex was added to the sample 5 - 20 minutes before irradiation at ambient temperature at 365 nm on a 1000-W Hg/Xe lamp and monochrometer (Oriel model 77250) for 5-20 minutes. The samples were precipitated, washed and dried, and then eluted through a 20% denaturing

polyacrylamide gel. The full-length RNA oligomers and cleavage products were identified by coelectrophoresing with $Ru(phen)_3^{2+}$ (G-specific) reactions,[23] diethyl pyrocarbonate (DEPC) (A-specific) and hydrazine (U-specific) reactions.[24] The fragments produced by the metal complex cleavage possess 3' and 5' phosphate termini, and thus could be directly compared with the chemical sequencing lanes.[25]

Determination of Affinity Constants. Binding constants were obtained through quantitative affinity cleavage titrations according to established procedures[26,27] and experimental conditions as described above. The rhodium/RNA nucleotide ratio was held constant at 1:20. The bulk RNA used was tRNA from baker's yeast (Boehringer Mannheim); previous experiments indicated that there is no difference in photocleavage sites and intensities between reactions done in the presence of tRNA or HIV RNA carrier (*vide infra*). The range of RNA concentrations used was 0.1 μM to 1000 μM, while the rhodium complex ranged from 5 nM to 50 μM. Photocleavage reactions were carried out at ambient temperature or on ice.

Affinity constants were determined in photocleavage experiments under single hit conditions, where

$$I_{site} = I_{sat}\frac{K_b[RNA]_{site}}{1 + K_b[RNA]_{site}}$$

and I_{site} = the intensity of photocleavage as measured by phosphorimagery, I_{sat} = the intensity of photocleavage at saturating rhodium concentration, K_b = affinity constant, and $[RNA]_{site}$ = the concentration of RNA hairpins.

Quantitation was accomplished using photostimulable storage phosphorimaging Kodak screens S0230 from Molecular Dynamics. A Molecular Dynamics 400S PhosphorImager was used to scan the screens, and Imagequant version 3.3 was used to analyze the data.

Peptide preparation. The 9-mer RKKRRQRRC (Tat-9) derived from residues 49 - 57 of the HIV Tat peptide and the 11-mers YRRRRRRRRRA (Arg-11) and YKKKKKKKKKA

(Lys-11) were chemically synthesized and purified by the Biopolymer Synthesis and Analysis Resource Center at Caltech. The peptides were prepared as their COOH-termini amides, and the HIV Tat peptide was acetylated on the NH_2 terminus. The peptides were analyzed by capillary electrophoresis and mass spectrometry. The peptides were suspended in 10 mM Tris HCl pH 7.0 and stored frozen until use.

Competition experiments with Tat peptide. Rhodium competition experiments were carried under the same conditions for photocleavage, with the addition of various peptides to the photocleavage solutions. Experiments were carried out with either a fixed concentration of rhodium complex with the peptide concentrations varied, or with a fixed concentration of peptide and varying the rhodium concentration. In both cases, all components except for the rhodium complexes were allowed to equilibrate for 1 hour, before the addition of the rhodium complex and subsequent irradiation. All competition experiments were carried out at ambient temperature.

Circular dichroism spectroscopy. Circular dichroism spectra were taken on a Jasco J500a spectrometer. Samples were prepared in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0, and the spectra taken at room temperature. Spectra were recorded from 400 to 200 nm using a 1-cm path-length cuvette. Scans were repeated 4 times and averaged, with a 1 s averaging time at each wavelength. Half an equivalent at a time of a total of 4 equivalents racemic 3-$Rh(MGP)_2phi^{5+}$ or $Rh(phen)_2phi^{3+}$ was titrated into a 300 μL solution of 2 μM ends renatured WT TAR or TAR1 RNA. Argininamide was then added to a total concentration of 10 mM. Argininamide was also added first to the free RNA solution, followed by the rhodium complex. 9,10-Diaminophenanthrene (DAP) in ethanol was also added to the RNA solution alone.

Ethyl nitrosourea (EtNU) cleavage. EtNU (Sigma) was freshly prepared in EtOH to 300 mM. A solution of ^{32}P WT TAR, tRNA carrier, metal complex and buffer was prepared. All components were added together except for the EtNU, and incubated for 1 hour at room temperature. The EtNU was then added to a final volume of 20 μL, and the

solution was then incubated for 5 mins - 3 hr. at 4°C, room temperature, 37°C, or 60°C. The samples were then precipitated, washed and dried. When dry, the samples were resuspended in 20 μL each 100 mM triethylbicarbonate (pH 8.5) (from Sigma) and incubated at 55°C for 30 - 60 min. The samples were dried, resuspended in loading dye, and run out on a 20% polyacrylamide denaturing gel.

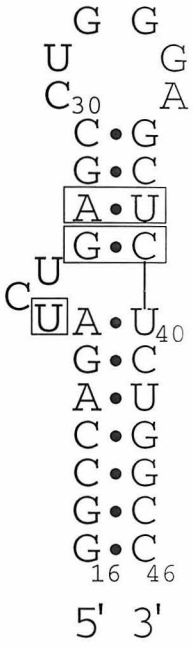
Gel shift assays for Tat-TAR complex formation. A variety of metal complexes were preincubated with ^{32}P labeled TAR RNA in the presence and absence of tRNA carrier. The metal complexes assayed were racemic, Δ- and Λ-$Rh(phen)_2phi^{3+}$, racemic, Δ- and Λ-3-$Rh(MGP)_2phi^{5+}$ and 3Δ-$Rh(GEB)_2phi^{5+}$. A variety of different conditions were used, with constant Rh:RNA carrier nt. ratios of 20:1, with the concentration of Rh ranging from 1 - 200 μM. A constant RNA carrier concentration of 100 μM nt. was also used, and the rhodium complex concentration was varied from 10 to 500 μM. In addition, these concentrations were used in the absence of any carrier. Solutions totaling 10 μM containing 1x Running buffer (50 mM Tris HCl, 20 mM NaOAc, 18 mM NaCl, pH 7.0), metal complex, ^{32}-P labeled TAR RNA, carrier RNA, and H_2O were preincubated for 1 - 24 hr at 4°C. In addition, some samples contained 20% glycerol. 2 μL 10x nondenaturing loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, 25% Ficoll) was then added to these samples, and they were incubated at 4°C for another hour. Non-denaturing mini-preparatory gels of 20% polyacrylamide (20:1 acrylamide:bisacrylamide) or 5% polyacrylamide (30:1 acrylamide:bisacrylamide) were prepared and preincubated at 4°C. Samples were then run on mini-prep gels at 4°C at 125 V for between 3-6 hr.

4.3. Results

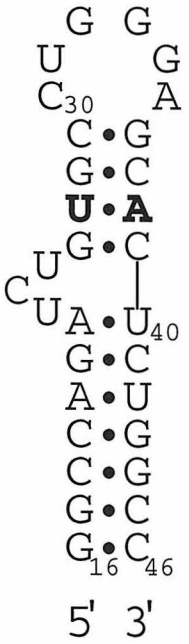
4.3.1. Photocleavage of TAR RNA and Mutants by Rhodium Complexes.

Site specific photocleavage of TAR RNA by Rh(phen)phi^{3+} and derivatives. Figure 4.3 shows the sequences and secondary structure of the WT TAR RNA and several mutants. Figure 4.4 shows racemic $Rh(phen)_2phi^{3+}$ cleavage sites on HIV TAR RNA.

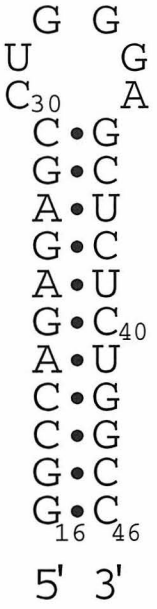
Figure 4.3. Sequence and secondary structures of wild-type TAR RNA and mutants. The boxed uridine in the bulge and the boxed A-U and G-C basepairs in the WT RNA are essential for Tat peptide binding. Residues in bold in the mutants are those mutated from wildtype.



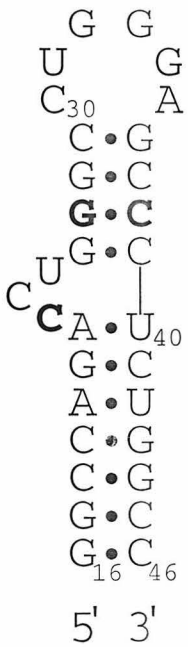
Wild-type TAR



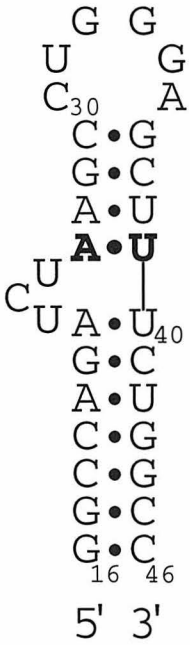
TAR1



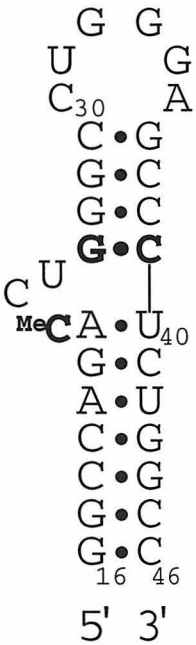
TAR2



TAR3

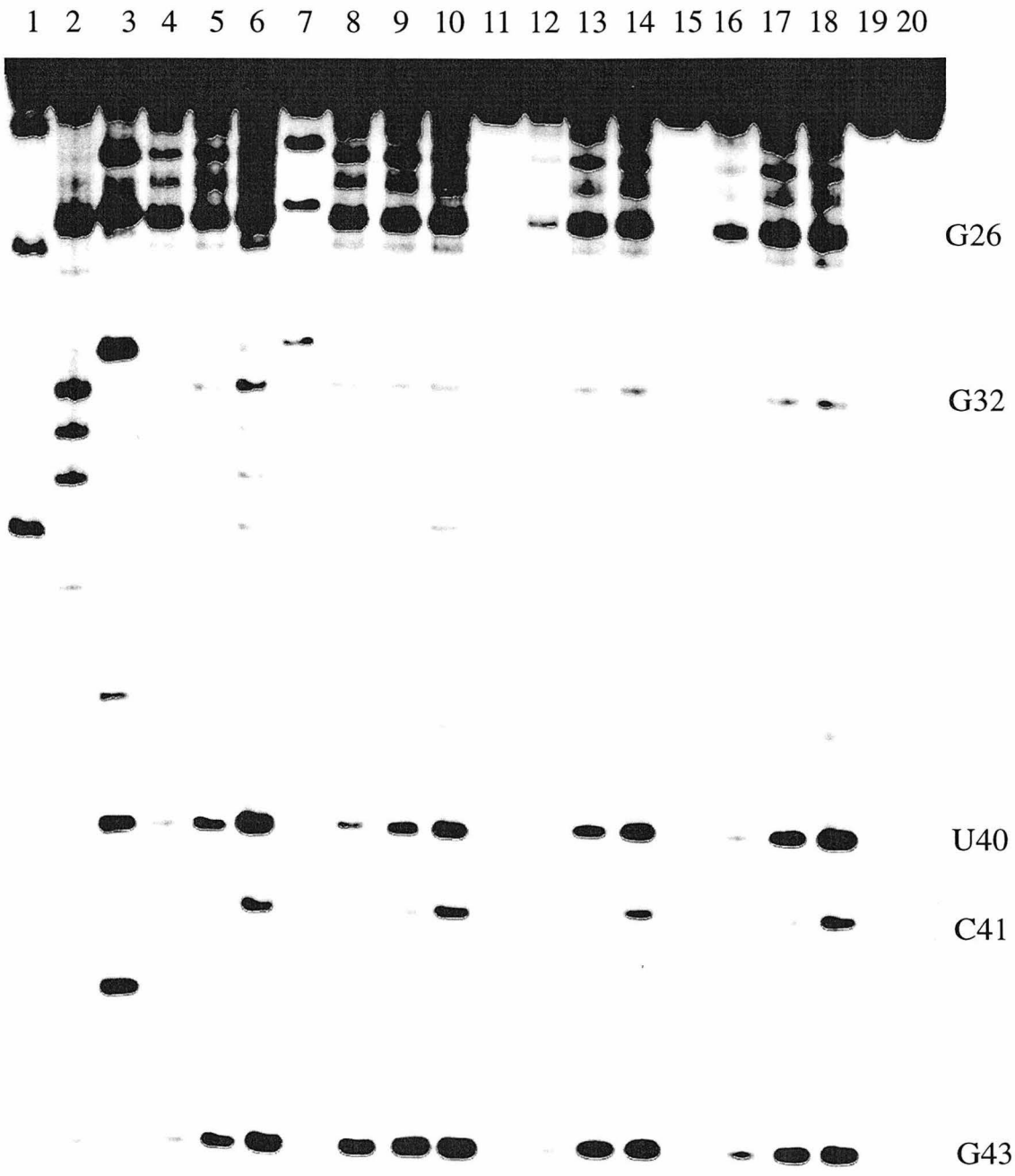


TAR4



TAR5

Figure 4.4. Grey-scale phosphorimager representation of a 20% polyacrylamide gel showing sites of racemic $Rh(phen)_2phi^{3+}$ and racemic $Rh(MGP)_2phi^{5+}$ isomers 1, 2 and 3 cleavage on ^{32}P 3'-end-labeled WT TAR RNA. All samples contained 100 μM tRNA. Lanes 1, 2 & 3: A-, G- and U-sequencing lanes respectively. Lanes 4-6: 2, 5 & 10 μM $Rh(phen)_2phi^{3+}$ and 10' irradiation at 365 nm. Lane 7: 10 μM $Rh(phen)_2phi^{3+}$. Lanes 8-10: 2, 5 & 10 μM 1-$Rh(MGP)_2phi^{5+}$ and 10' irradiation at 365 nm. Lane 11: 10 μM 1-$Rh(MGP)_2phi^{5+}$. Lanes 12-14: 2, 5 & 10 μM 2-$Rh(MGP)_2phi^{5+}$ and 10' irradiation at 365 nm. Lane 15: 10 μM 2-$Rh(MGP)_2phi^{5+}$. Lanes 16-18: 2, 5 & 10 μM 3--$Rh(MGP)_2phi^{5+}$ and 10' irradiation at 365 nm. Lane 19: 10 μM 3-$Rh(MGP)_2phi^{5+}$. Lane 20: light control: 10' irradiation at 365 nm.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
G26
G32
U40
C41
G43

The cleavage sites are identical to those reported previously, the strongest sites being U23, G26, U40 and C41.[16] Cleavage studies (data not shown) with the separated enantiomers show that the Δ-enantiomer is responsible for most of the cleavage, although the Λ-enantiomer does cleave at U40 and C41.

$Rh(MGP)_2phi^{5+}$ cleaves at essentially the same sites as the parent complex. Figure 4.4 also shows a comparison of the cleavage sites of the parent racemic $Rh(phen)_2phi^{3+}$ and the 3 isomers of $Rh(MGP)_2phi^{5+}$. The main sites of cleavage for the 3 different isomers are U23, G26, U40 and C41, and are very similar to the $Rh(phen)_2phi^{3+}$ sites. These cleavage sites are summarized in Table 4.1.

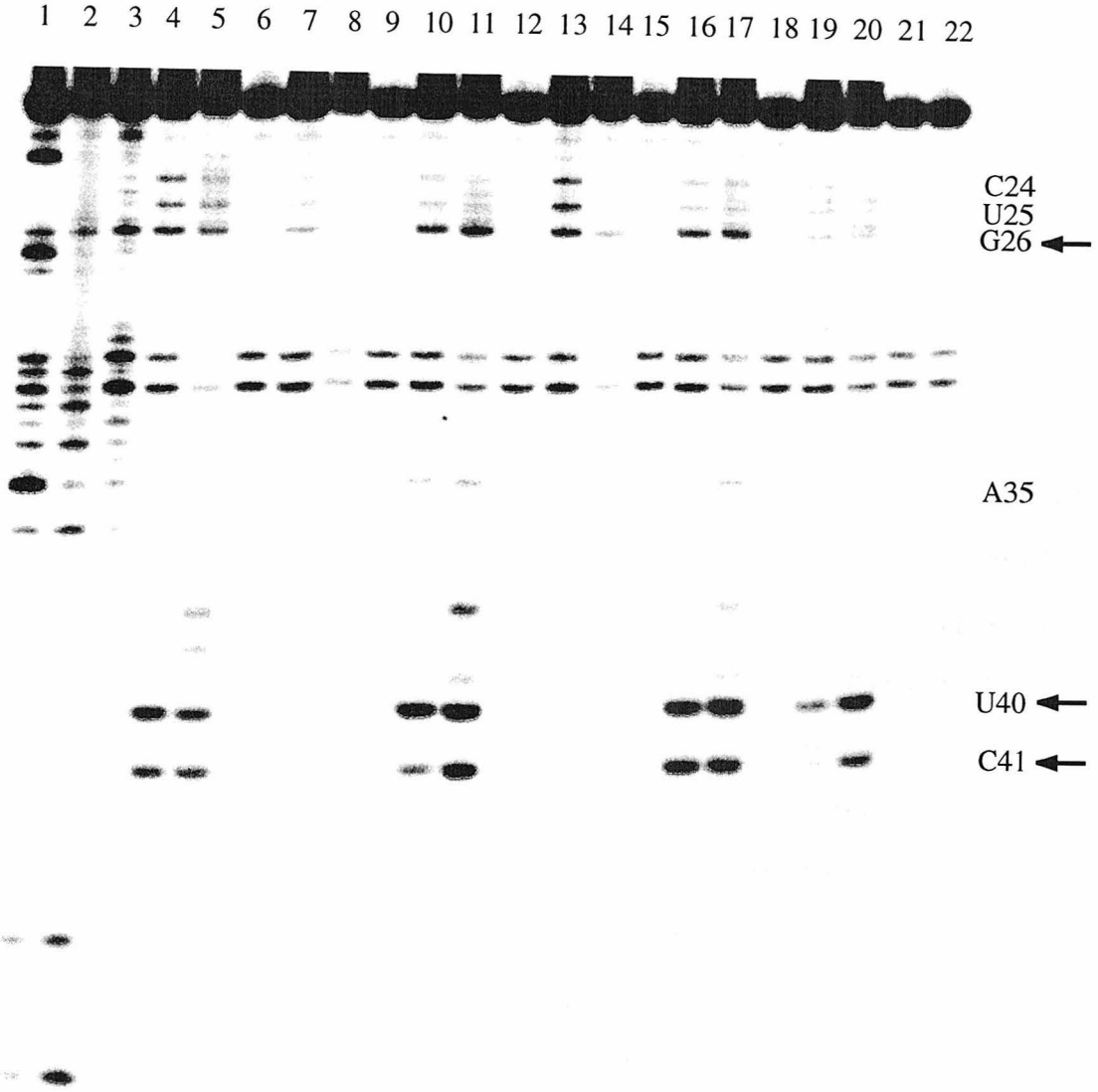
When the resolved enantiomers of the positional isomers of $Rh(MGP)_2phi^{5+}$ were used, the complex with the greatest degree of cleavage is 3Δ. The same cleavage patterns were seen with the various separated enantiomers of the 3 isomers of $Rh(GEB)_2phi^{5+}$. Figure 4.5 shows a polyacrylamide gel with a comparison of the various isomers and enantiomers of $Rh(GEB)_2phi^{5+}$, showing that the Δ enantiomers are responsible for the primary cleavage, and the isomer 3 has the strongest binding sites. Table 2 includes a summary of the cleavage sites of the resolved enantiomers of $Rh(GEB)_2phi^{5+}$. The strongest sites for all 3 Δ-enantiomers are U40 and C41. In addition, G26 is recognized strongly. The cleavage sites for all the Δ complexes are very similar. However, the Λ enantiomers of the 3 isomers show differences in their recognition sites. The 1Λ-$Rh(GEB)_2phi^{5+}$ does not appear to recognize the RNA molecule, while the 2Λ-$Rh(GEB)_2phi^{5+}$ cleaves weakly at C24, U25 and G26. On the other hand, 3Λ-$Rh(GEB)_2phi^{5+}$ cleaves at a subset of the 3Δ sites, cleaving at U40 and C41. Thus, the placement of the guanidinium arms does not appear to affect the recognition characteristics significantly when the shape of the metal complex is better matched to the helix, in the case of the Δ enantiomers, but has more of an effect when the poorly matched Λ-complex is used. Since the same trend is seen in the cleavage by the resolved isomers of the $Rh(MGP)_2phi^{5+}$ enantiomers, this indicates that the addition of the extra methylene group

Table 4.1. A comparison of the cleavage sites of rac-Rh(MGP)$_2$phi^{5+} (isomer 3) on TAR RNA and its mutants.

RNA species	Cleavage sites[a]
WT	A22, U23, G26, *G32*, U38, **U40**, **C41**, G43
TAR1 (A27U-U38A)	A22, U23, U25, G26, **U40**, G43
TAR2 (U23Δ-C24Δ-U25Δ)	G43
TAR3 (U23C- A27G-U38C)	U40
TAR4 (G26A-C39U	A26, C39, **U40** and C41

[a] Strong sites are in bold, very weak sites are in italics.

Figure 4.5. Grey-scale phosphorimager representation of a 20% polyacrylamide gel showing sites of cleavage by Δ- and Λ- 1,2,3 $Rh(GEB)_2phi^{5+}$ isomers on ^{32}P 3'-end-labeled WT TAR RNA. All samples contained 100 μM tRNA. Lanes 1, 2 & 3: A-, G- and U-sequencing lanes respectively. Lanes 4 & 5: 5 & 10 μM 1Δ-$Rh(GEB)_2phi^{5+}$ and 10' irradiation at 365 nm. Lanes 7 & 8: 5 & 10 μM 1Λ-$Rh(GEB)_2phi^{5+}$ and 10' irradiation at 365 nm. Lanes 10 & 11: 5 & 10 μM 2Δ-$Rh(GEB)_2phi^{5+}$ and 10' irradiation at 365 nm. Lanes 13 & 14: 5 & 10 μM 2Λ-$Rh(GEB)_2phi^{5+}$ and 10' irradiation at 365 nm. Lanes 16 & 17: 5 & 10 μM 3Δ-$Rh(GEB)_2phi^{5+}$ and 10' irradiation at 365 nm. Lanes 19 & 20: 5 & 10 μM 3Λ-$Rh(GEB)_2phi^{5+}$ and 10' irradiation at 365 nm. Lanes 6, 9, 12, 15, 18, 21: dark controls: 10 μM 1Δ, 1Λ, 2Δ-, 2Λ, 3Δ- and 3Λ-$Rh(GEB)_2phi^{5+}$ respectively. Lane 22: light control: 10' irradiation at 365 nm.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
C24
U25
G26
A35
U40
C41

and the additional flexibility of the bpy ligand on $Rh(GEB)_2phi^{5+}$ do not substantially affect the recognition characteristics.

We rationalize that the primary site of cleavage by all the Δ-enantiomers, U40, is recognized because it is directly opposite the bulge. Another site of cleavage is U23, which lies within the bulge. This bulge region is not canonical A-form and the major groove is more open. DEPC[28] and $Rh(phen)_2phi^{3+}$ [16] both recognize the bulge region, and thus this opening in the major groove represents a recognition element for both the Tat peptide and chemical cleavage agents. The cleavage site at G43 is harder to rationalize from a structural standpoint, but as described in the next section, it is also recognized in a canonical A-form helix. Therefore, this site has a recognition element which is not apparent from the secondary structure and can not be explained by the structural data on this RNA molecule.

In addition, very weak cleavage at the three guanine residues, G32, G33, and G34, on the loop region and at A35 is observed, but not reproducibly. This effect appears with a variety of metal complexes. The photocleavage could be due to the loop becoming structured and forming a binding pocket for the metal complex. This effect is not due to nucleases, even though this loop is susceptible to nuclease cleavage, because the metal-induced cleavage comigrates with the Maxam Gilbert sequencing lanes.

As the primary sites of cleavage are at the bulge region and the region opposite it, we hypothesize that the primary site of intercalation is right in the bulge itself, between A22-U40 and G26-C39. Intercalation, or at least a close intimate association, is required for productive cleavage, given the short excited state lifetime of the rhodium complex. When a triple is formed by U23 swinging up to interact with the A27-U38 basepair, this should tighten the binding pocket and bring G26 and U40 in close proximity, hence these would together represent the main sites of cleavage. In the absence of the triple, as in TAR1 as described in the next section, the metal complex still may intercalate into the bulge as well, but the pattern of photocleavage is slightly different, as the structure is looser.

Photocleavage of TAR RNA mutants by rhodium complexes. Table 4.1 shows a comparison of cleavage sites by the 3Δ-Rh(MGP)$_2$phi^{5+} complex on WT TAR RNA and its various mutants. TAR2 (bulge mutant deletion) shows only one strong site of cleavage, at G43. It is unclear why this site is targeted, as it should lie within a canonical A-form helix with a deep and narrow major groove, and thus be inaccessible. This site is also targeted in each of the RNA molecules. Other sites of cleavage with the bulge present are not recognized underscoring the fact that Rh(phen)$_2$phi^{3+} primarily targets non-canonical RNA.

There are subtle differences between the sites of rhodium complex cleavage in the WT TAR molecule, and the TAR1 mutant. G26 and U40 are recognized strongly in the WT TAR, but the intensity of cleavage at G26 in TAR1 is diminished, as shown in Figure 4.6. A schematic illustrating the main differences in TAR WT and TAR1 rhodium photocleavage is shown in Figure 4.7. We attribute these differences to a tightening in the bulge region in TAR WT vs. the TAR1 bulge.

The TAR3 mutant was expected to form a C$^+$23•G27-C38 base triple under acidic conditions. It was found that there was enhanced cleavage by GEB3Δ on TAR3 at pH 5.5 over pH 7.0. The opposite effect in TAR WT was observed, where cleavage decreased with decreasing pH. The TAR WT observation is consistent with diminished rhodium complex photocleavage efficiency at acidic pHs.[29] However, in the case of TAR3, the enhanced photocleavage by the metal complex could indicate that the triple structure is tightened at lower pH's where the C23 base is protonated, thus forming a better binding pocket for the metal complex. This would negate the effects of diminished photoefficiency.

The TAR4 mutant is still competent to form the U23•A27-U38 base triple, but lacks the G26-C39 base pair which an arginine group on the Tat peptide has been proposed to contact .[9] A wide screen of complexes were assayed against this mutant, and the rhodium complexes which bound the tightest are Δ-Rh(phen)$_2$phi^{3+}, 3Δ-Rh(MGP)$_2$phi^{5+}, and 2Δ- and 3Δ-Rh(GEB)$_2$phi^{5+}, which was similar to the range of complexes which recognized

Figure 4.6. Grey-scale phosphorimager representation of a 20% polyacrylamide gel showing sites of racemic $Rh(phen)_2phi^{3+}$ and racemic $Rh(MGP)_2phi^{5+}$ isomers 1, 2 and 3 cleavage on ^{32}P 3'-end-labeled TAR1 RNA. All samples contained 100 μM tRNA. Lanes 1, 3, 5, 7: 5 μM $Rh(phen)_2phi^{3+}$, 5 μM $Rh(MGP)_2phi^{5+}$ isomers 1, 2, and 3 respectively, 10' irradiation at 313 nm. Lanes 2, 4, 6, 8: 5 μM $Rh(phen)_2phi^{3+}$, 5 μM $Rh(MGP)_2phi^{5+}$ isomers 1, 2, and 3 respectively. Lane 11: light control: 10' irradiation at 365 nm. Lanes 12, 13 & 14: A-, G- and U-sequencing lanes respectively.

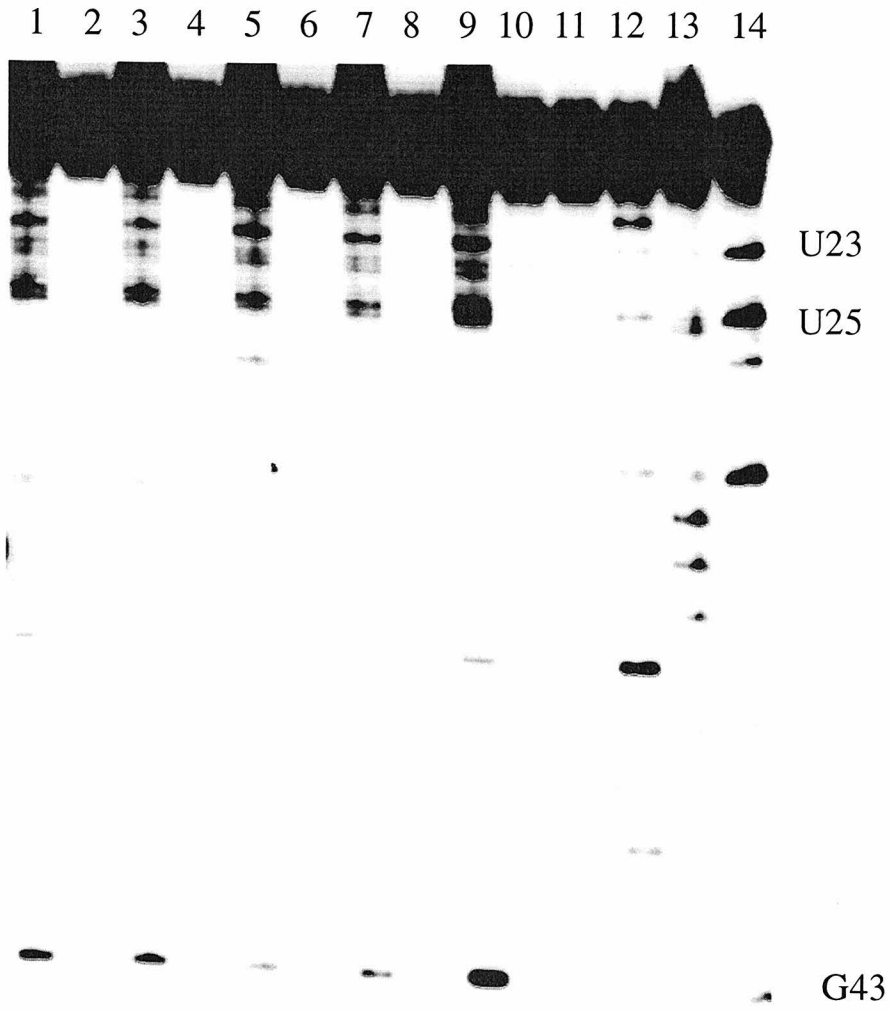
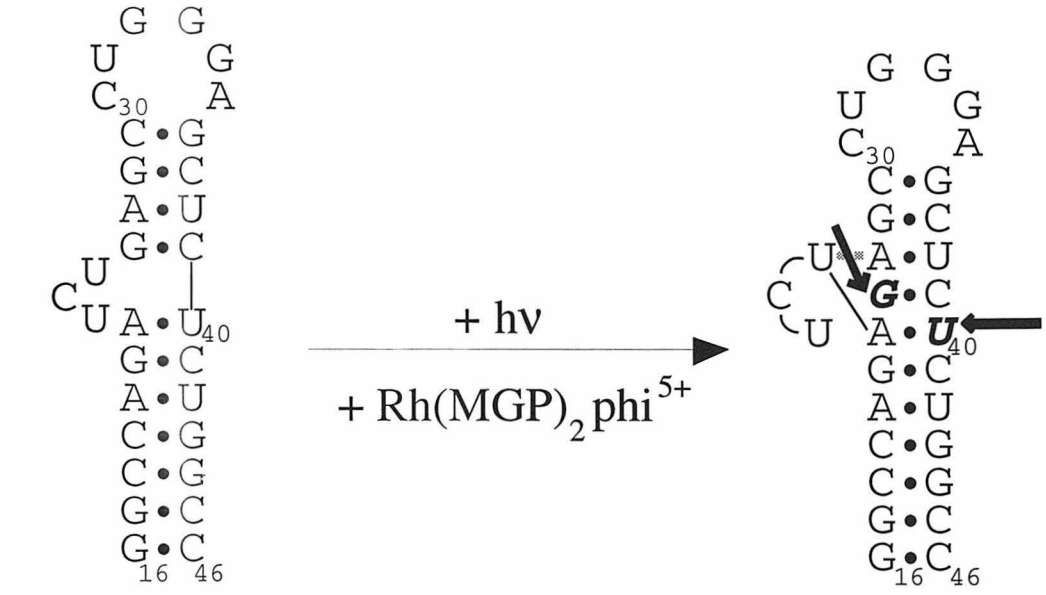
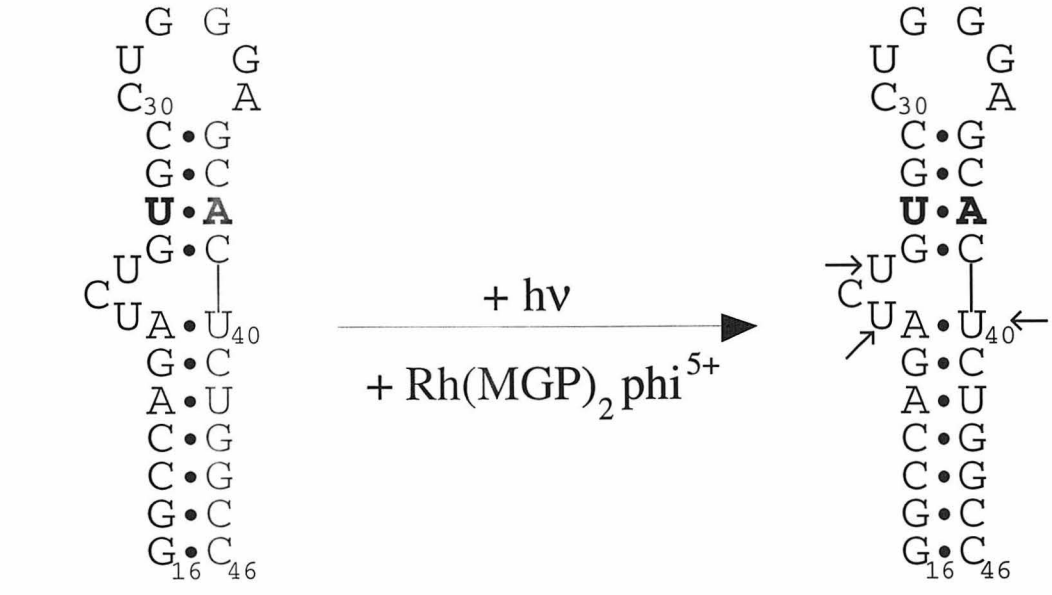

1 2 3 4 5 6 7 8 9 10 11 12 13 14
U23
U25
G43

Figure 4.7. Schematic of sites of cleavage on WT TAR RNA vs. on TAR1, showing that G26 and U40 are in close proximity in WT TAR if the base triple is formed, while the bulge is more loosely structure in TAR1 in the absence of the base triple.



Wild-type TAR



A27U-U38A TAR

the parent TAR WT molecule. These complexes all yielded identical cleavage patterns, at A26, C39, U40 and C41. These sites were similar to those recognized on the WT TAR. Therefore, it appears that the binding and cleavage of the rhodium complexes are not affected by the lack of the G26-C39 basepair. In particular, the rhodium complexes containing guanidinium moieties ($Rh(MGP)_2phi^{3+}$ and $Rh(GEB)_2phi^{5+}$) show no difference in cleavage patterns on TAR4 compared with WT TAR. 3Δ-$Rh(MGP)_2phi^{5+}$ has the highest binding affinity to this molecule, in the micromolar range, which is consistent with its binding affinity for WT TAR. However, the similarity in cleavage sites between $Rh(phen)_2phi^{3+}$ and $Rh(MGP)_2phi^{3+}$ indicate that the guaninidium arms are unlikely to be mimicking the arginine group of the Tat peptide in contacting the C39 base. Therefore, most of the impetus for rhodium complex-RNA affinity comes from the intercalating phi ligand.

The U23^{5}Me-C•A27G-U38C TAR RNA mutant (TAR5) was also probed. The ^{5}Me-C stabilizes the base triple and renders it stable at neutral pH, as opposed to the C23•G27-C38 triple which is only stable at acidic pH where the metal complex cleavage efficiencies are lower.[29] The photocleavage gels done on this mutant are very highly nucleased, as the ^{5}Me-C is not very stable. However, there did not appear to be a pH dependence in photocleavage.

4.2.2. Effect of Other Factors on Rhodium Photocleavage

Effect of argininamide on rhodium photocleavage of TAR RNA. No difference in sites and intensity of cleavage by $Rh(phen)_2phi^{3+}$ at either the region opposite the bulge region (Figure 4.8), or bulge region itself (Figure 4.9), was seen upon the addition of argininamide to the photocleavage samples. No change in $Rh(MGP)_2phi^{5+}$ photocleavage was seen as well (data not shown). Thus, the change induced by argininamide was either not detected by the metal complex, or that the metal complex was already inducing a conformational change. The latter explanation is the more likely one, and is supported by

Figure 4.8. Grey scale representation of a 20% polyacrylamide gel showing the effect of argininamide on $Rh(phen)_2phi^{3+}$ photocleavage of WT TAR RNA. Lanes 1, 2 & 3: A, G and U sequencing lanes. Lane 3: 1 μM $Rh(phen)_2phi^{3+}$, 10 μM argininamide, 30' irradiation at 313 nm. Lanes 4 & 5: 5 & 10 μM $Rh(phen)_2phi^{3+}$ respectively, 10 μM argininamide, 10' irradiation at 313 nm. Lanes 6, 7 & 8: identical to 3-5, but with 50 μM argininamide. Lane 9: 10 μM $Rh(phen)_2phi^{3+}$, 50 μM argininamide. Lane 10: 50 μM argininamide, 30' irradiation at 313 nm. Lane 11: 1 μM $Rh(phen)_2phi^{3+}$, 30' irradiation. Lanes 12 & 13: 5 & 10 μM $Rh(phen)_2phi^{3+}$ respectively, 10' irradiation at 313 nm. Lane 14: 10 μM $Rh(phen)_2phi^{3+}$. Lane 15: 10' irradiation at 313 nm.

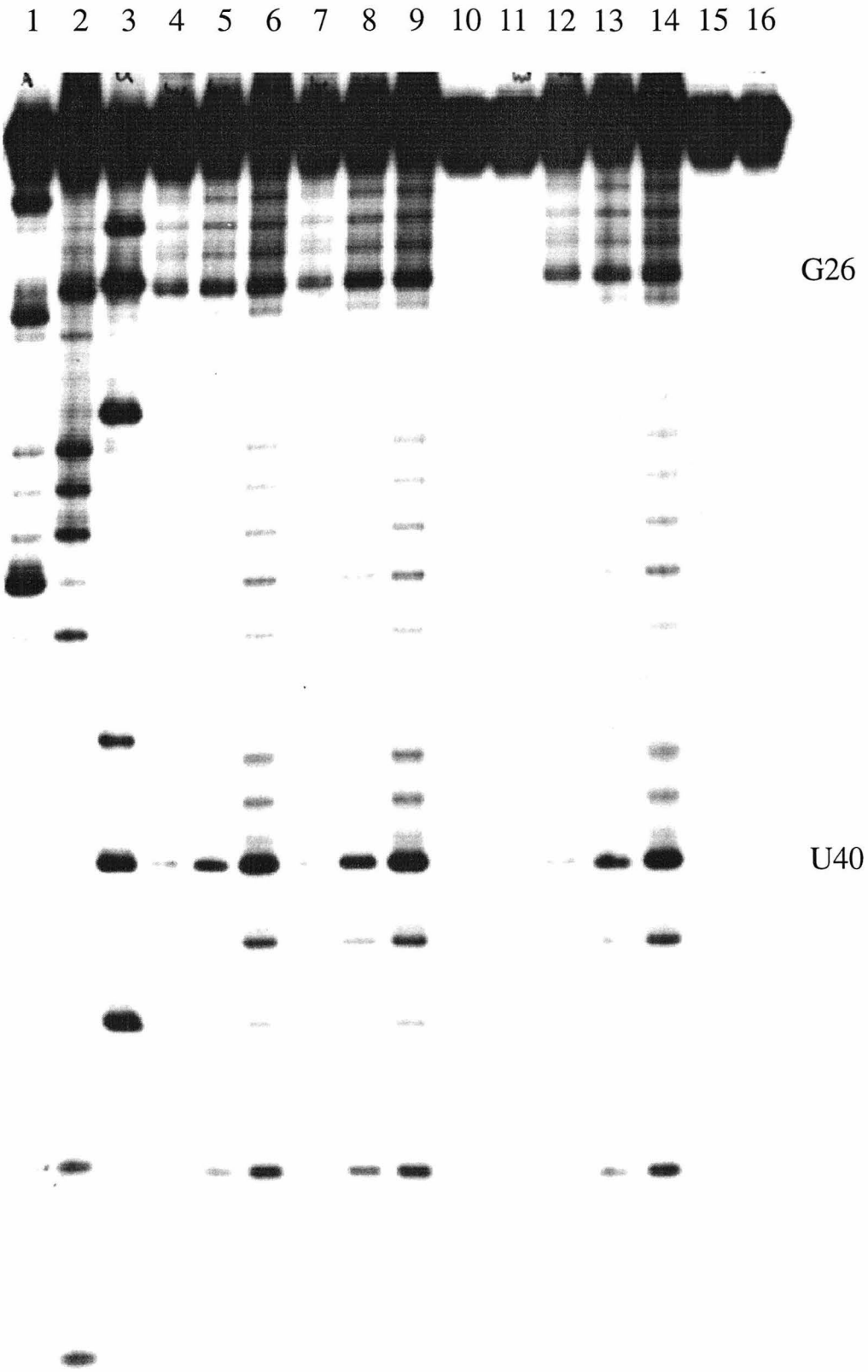
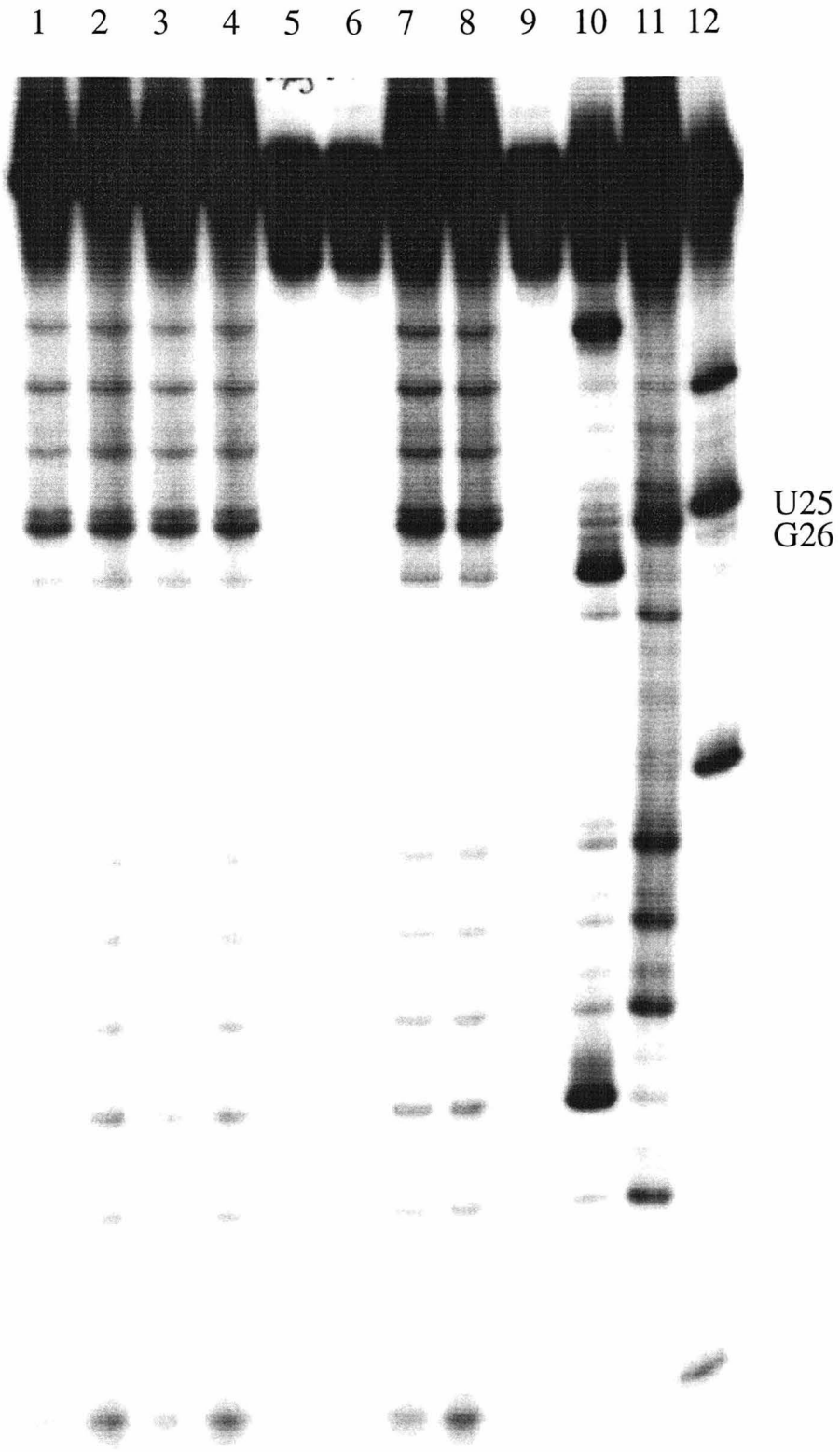

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
G26
U40

Figure 4.9. Grey scale representation of a 20% polyacrylamide gel showing the effect of argininamide on $Rh(phen)_2phi^{3+}$ photocleavage of WT TAR RNA. Lanes 1 & 2: 5 & 10 μM $Rh(phen)_2phi^{3+}$ respectively, 10 μM argininamide, 10' irradiation at 313 nm. Lanes 3 & 4: 5 & 10 μM $Rh(phen)_2phi^{3+}$ respectively, 50 μM argininamide, 10' irradiation at 313 nm. Lane 5: 10 μM $Rh(phen)_2phi^{3+}$, 50 μM argininamide. Lane 6: 50 μM argininamide, 30' irradiation at 313 nm. Lanes 7 & 8: 5 & 10 μM $Rh(phen)_2phi^{3+}$ respectively, 10' irradiation at 313 nm. Lane 9: 10 μM $Rh(phen)_2phi^{3+}$. Lanes 10, 11 & 12: A, G & U sequencing respectively.


1 2 3 4 5 6 7 8 9 10 11 12
U25
G26

the circular dichroism studies (*vide infra*). In addition, this indicates that the argininamide does not displace the metal complex from its binding site.

Effect of wavelength on rhodium photocleavage on TAR. The TAR RNA was irradiated at 313 nm in the presence of the metal complexes racemic, Δ- and Λ-$Rh(phen)_2phi^{3+}$, and racemic, Δ- and Λ- 3-$Rh(MGP)_2phi^{5+}$. Sites of photocleavage were identical to those observed at 365 nm.

Effect of buffer and metal ions on rhodium complex photocleavage on TAR RNA. Effect of buffer on photocleavage. The TAR photocleavage reactions were run in either 50 mM NaCl, 10 mM NaH_2PO_4, pH 6.5 or 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. Both buffers gave identical cleavage sites and relative intensities. The effect of magnesium ions on the rhodium complex cleavage of TAR was also studied. There was no effect on the cleavage pattern of 3Δ- and 3Λ-$Rh(MGP)_2phi^{5+}$ in the presence of 10 mM $MgCl_2$.

Effect of carrier on rhodium complex photocleavage. Rhodium complex photocleavage on WT TAR RNA was carried out in the presence of both 100 μM nt. WT TAR RNA and 100 μM nt. $tRNA^{Phe}$. As can be seen in Figure 4.10, there was no difference in both the sites of cleavage and the intensity of cleavage by racemic $Rh(phen)_2phi^{3+}$ with either buffer. The cleavage by the 3 isomers of $Rh(MGP)_2phi^{5+}$ showed a similar lack of dependence on carrier (Figure 4.11).

Effect of $Rh(DIP)_3^{3+}$ on TAR RNA. TAR RNA was irradiated in the presence of 5 μM $Rh(DIP)_3^{3+}$ at 313 and 365 nm for 10 and 20 minutes. No sites of cleavage were observed.

4.3.3. Binding Constants of Metal Complexes to TAR RNA

Binding constants for the metal complexes to WT RNA were obtained by affinity titrations.[26,27] A typical gel is shown in Figure 4.12, and a typical binding affinity curve is shown in Figure 4.13. The binding constants for a series of metal complexes to WT TAR

Figure 4.10. Grey scale representation of a 20% polyacrylamide gel showing the effect of carrier on $Rh(phen)_2phi^{3+}$ photocleavage of WT TAR RNA. Lanes 1-6 contain 100 μM nt. TAR RNA carrier. Lanes 7-12 contain 100 μM nt. $tRNA^{Phe}$ carrier. Lanes 1 & 2: 2 μM $Rh(phen)_2phi^{3+}$, 20 μM argininamide, irradiated for 20' at 313 nm and no irradiation respective. Lane 3: 20 μM argininamide, irradiated for 20' at 313 nm. Lanes 4 - 6: identical to lanes 1-3, but without argininamide. Lanes 7-12: identical to lanes 1-6 respectively, but with tRNA carrier.

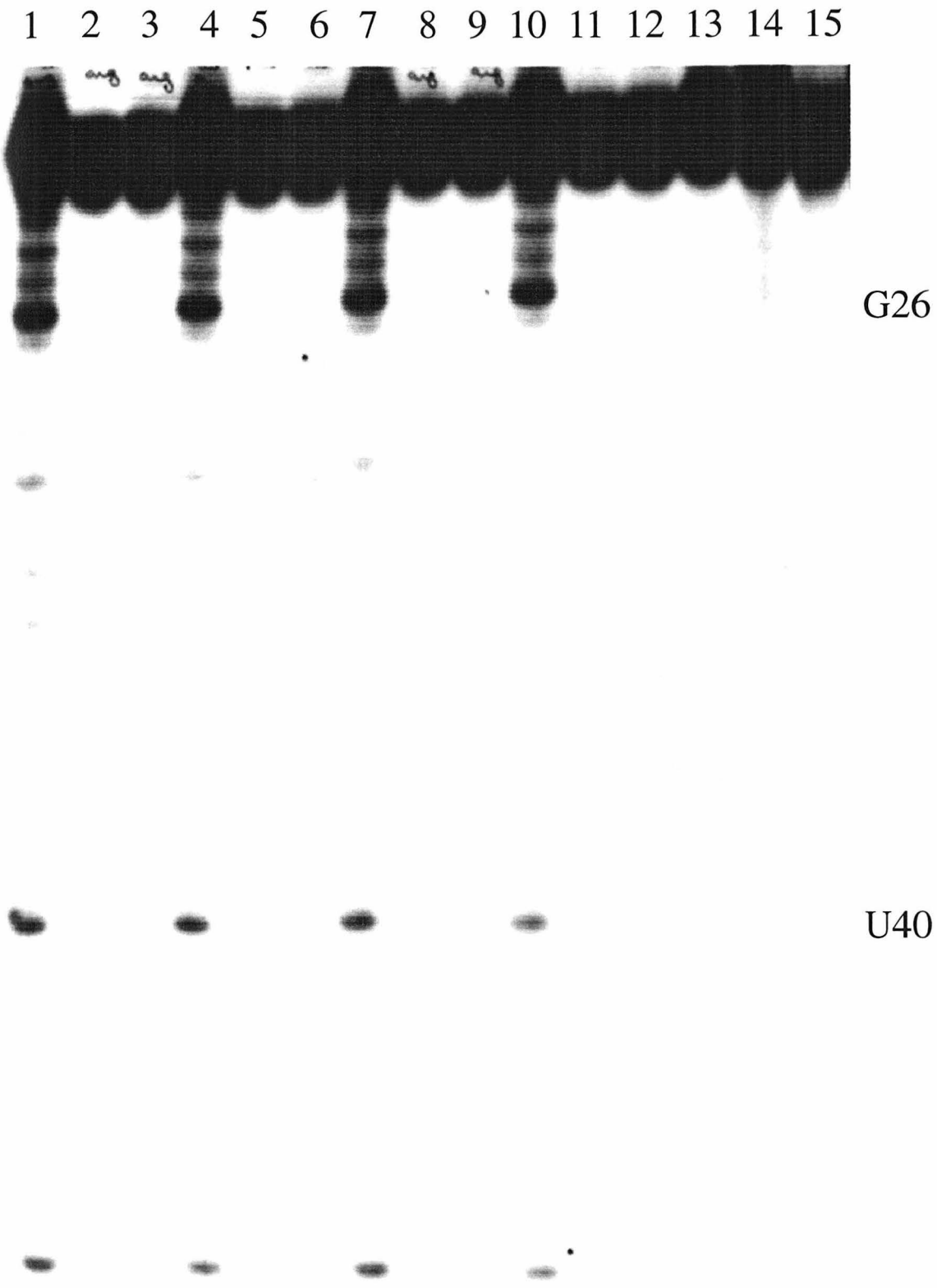

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
G26
U40

Figure 4.11. Grey scale representation of a 20% polyacrylamide gel showing the effect of carrier on racemic $Rh(MGP)_2phi^{5+}$ photocleavage of WT TAR RNA. Lanes 1-3: 10 μM $Rh(MGP)_2phi^{5+}$ isomers 1, 2 & 3 respectively, 10' irradiation at 313 nm, 100 μM TAR RNA carrier. Lanes 4-6: 10 μM $Rh(MGP)_2phi^{5+}$ isomers 1, 2 & 3 respectively, 100 μM TAR RNA carrier. Lanes 7-9: 10 μM $Rh(MGP)_2phi^{5+}$ isomers 1, 2 & 3 respectively, 10' irradiation at 313 nm, 100 μM $tRNA^{Phe}$ carrier. Lanes 10-12: 10 μM $Rh(MGP)_2phi^{5+}$ isomers 1, 2 & 3 respectively, 100 μM $tRNA^{Phe}$ carrier.

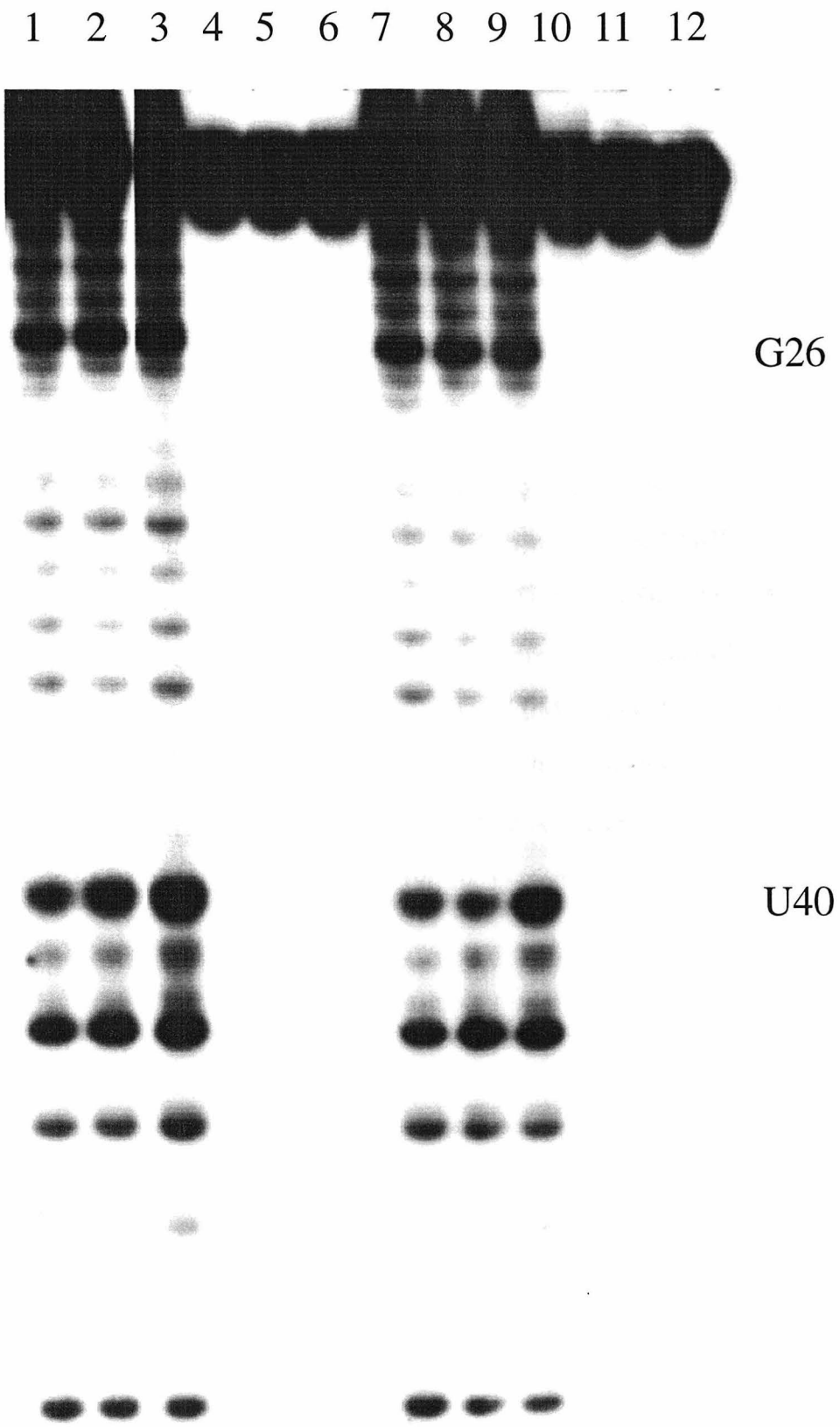

1 2 3 4 5 6 7 8 9 10 11 12
G26
U40

Figure 4.12. Greyscale phosphorimager representation of a 20% polyacrylamide gel showing the quantitative affinity titration of 3Δ-Rh(MGP)$_2$phi^{5+} to WT TAR RNA. Experiments were conducted at ambient temperature in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. buffer. All Rh/TAR RNA ratios were at 1:20. Lanes 1-13: 1 nM, 3 nM, 6 nM, 10 nM, 30 nM, 60 nM, 100 nM, 0.3 μM. 0.6 μM, 1 μM, 3 μM, 6 μM and 10 μM 3Δ-Rh(MGP)$_2$phi^{5+} respectively upon irradiation for 20' at 365 nm. Lane 14: 10 μM 3Δ-Rh(MGP)$_2$phi^{5+}, no irradiation. Lane 15: labeled RNA upon irradiation for 20' at 365 nm. Lane 16: labeled RNA. Lanes 17, 18 & 19: A-, G- and U-sequencing lanes respectively.

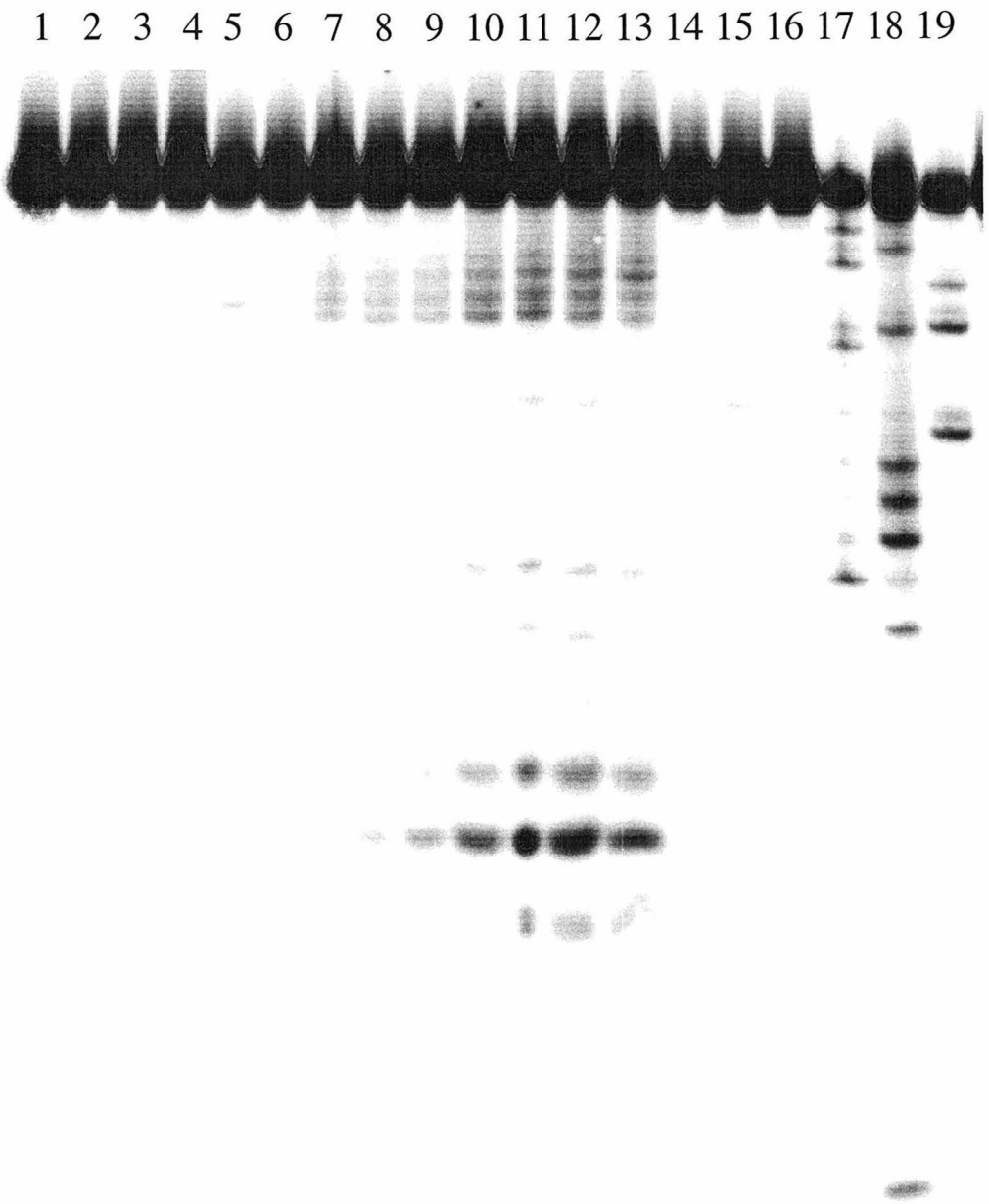
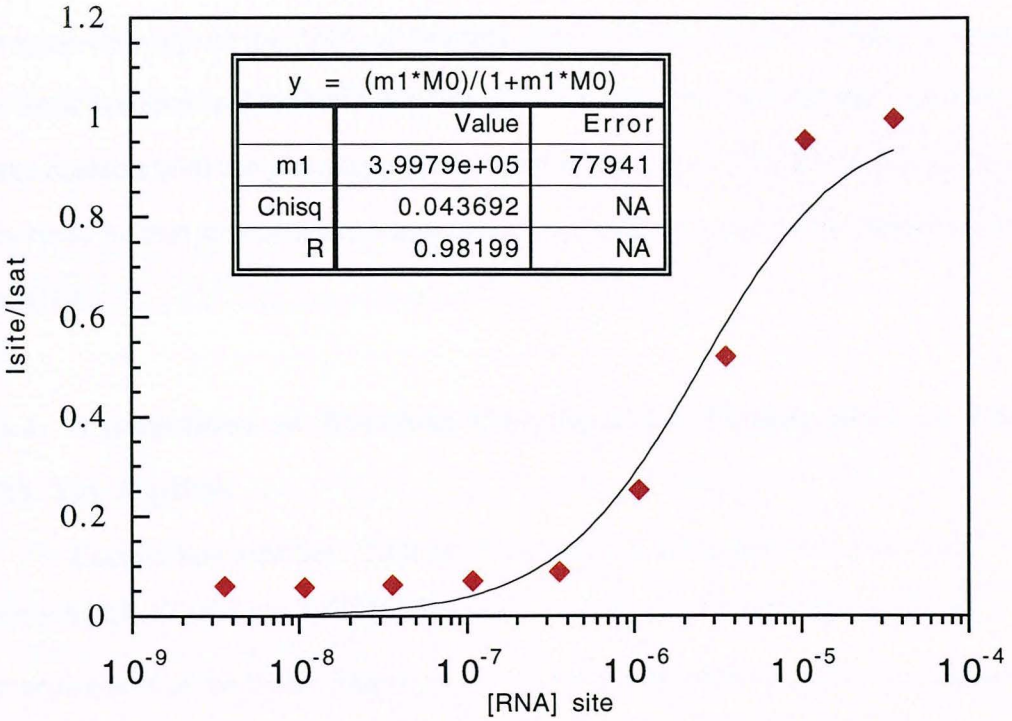

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
U40

Figure 4.13. Quantitative affinity cleavage titration for rac-Rh(phen)$_2$phi^{3+} on WT TAR RNA. Plot of cleavage intensity, I_{site}, relative to the intensity at saturation, I_{sat}, as a function of RNA concentration. These data represent the site-specific cleavage signal intensity at U40 with a Rh/TAR RNA ratio of 1:20. Experiments were conducted at ambient temperature in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. buffer, with irradiation at 365 nm.

Binding curve for racemic $Rh(phen)_2phi^{3+}$ to TAR RNA



RNA are summarized in Table 4.2. As can be seen, the Δ-enantiomer of each rhodium complex binds tighter to the TAR RNA than does the Λ-enantiomer. However, the difference in binding affinity is not substantial, even though in most cases, the Δ-enantiomer shows stronger photocleavage than the Λ-enantiomer does.

When the various rhodium complexes are compared, it can be seen that the addition of guanidinium arms provides the metal complex with greater affinity for the RNA, on the order of 0.5-1 kcal/mol. The $Rh(MGP)_2phi^{3+}$ isomer with the greatest affinity is isomer 3, where both arms are pointing out. This was consistent with this isomer having the strongest cleavage on the RNA, as detailed in the previous section. When the interaction of the metal complex and the RNA is modeled (*vide infra*), we find that the arms are poised to make contacts with the phosphate groups on the backbone of the RNA, due to the non-canonical A-form structure and major groove opening provided by the three base bulge in the RNA.

4.3.4. Competition of Rhodium Complexes for Binding Sites on TAR RNA with Tat Peptides

Competition with Tat. TAR photocleavage in the presence of a constant concentration of racemic 3-$Rh(MGP)_2phi^{5+}$ was done in the presence of various concentrations of the 9-mer fragment of Tat (Tat-9). In addition, the lysine- (Lys-11) and arginine-rich (Arg-11) 11mer peptides were used as controls, based on previous studies which showed that Arg-11 binds specifically to TAR, while Lys-11 binds poorly .[6] In addition, peptides containing Arg-11 were active in transactivation, but those containing Lys-11 were not.[6] Figure 4.14 shows a gel where rhodium complex photocleavage is inhibited by Arg-11, Lys-11 and Tat-9. Figure 4.15 shows the average percentage cleavage by 5 μM $Rh(MGP)_2phi^{5+}$ in the presence of increasing concentrations of peptide, compared to 5 μM rhodium complex photocleavage with no peptide. As can be seen from the graph, 5 μM Tat-9 was necessary to inhibit 5 μM 3Δ-$Rh(MGP)_2phi^{5+}$ photocleavage

Table 4.2. Cleavage sites of $Rh(GEB)_2phi^{5+}$ enantiomers and isomers on WT TAR RNA.

$Rh(GEB)_2phi^{5+}$ species	Cleavage sites[a]
1Δ	C24, U25, G26, *A35*, **U40**, **C41**
1Λ	*U40*
2Δ	G26, *A35*, **U40**, **C41**
2Λ	C24, U25, G26
3Δ	G26, *A35*, **U40**, **C41**
3Λ	U40, C41

[a] Strong sites are in bold, very weak sites are in italics.

Figure 4.14. Greyscale phosphorimager representation of a 20% polyacrylamide gel showing peptide inhibition of 3Δ-Rh(MGP)$_2$phi^{5+} photocleavage on 3'-^{32}P-labeled WT TAR RNA in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. Lanes 1 - 3: 1, 10 & 100 μM Tat-9 respectively, 5 μM 3Δ-Rh(MGP)$_2$phi^{5+}, 10' irradiation at 365 nm. Lane 4: 100 μM Tat-9, 10' irradiation. Lanes 5 - 7: 1, 10 & 100 μM Arg-11 respectively, 5 μM 3Δ-Rh(MGP)$_2$phi^{5+}, 10' irradiation at 365 nm. Lane 8: 100 μM Arg-11, 10' irradiation. Lanes 9 - 11: 1, 10 & 100 μM Lys-11 respectively, 5 μM 3Δ-Rh(MGP)$_2$phi^{5+}, 10' irradiation at 365 nm. Lane 12: 100 μM Lys-11, 10' irradiation. Lane 13: 5 μM 3Δ-Rh(MGP)$_2$phi^{5+}, 10' irradiation at 365 nm. Lane 14: 5 μM 3Δ-Rh(MGP)$_2$phi^{5+}, no irradiation. Lanes 15, 16 & 17: A-, G- and U-sequencing lanes respectively.

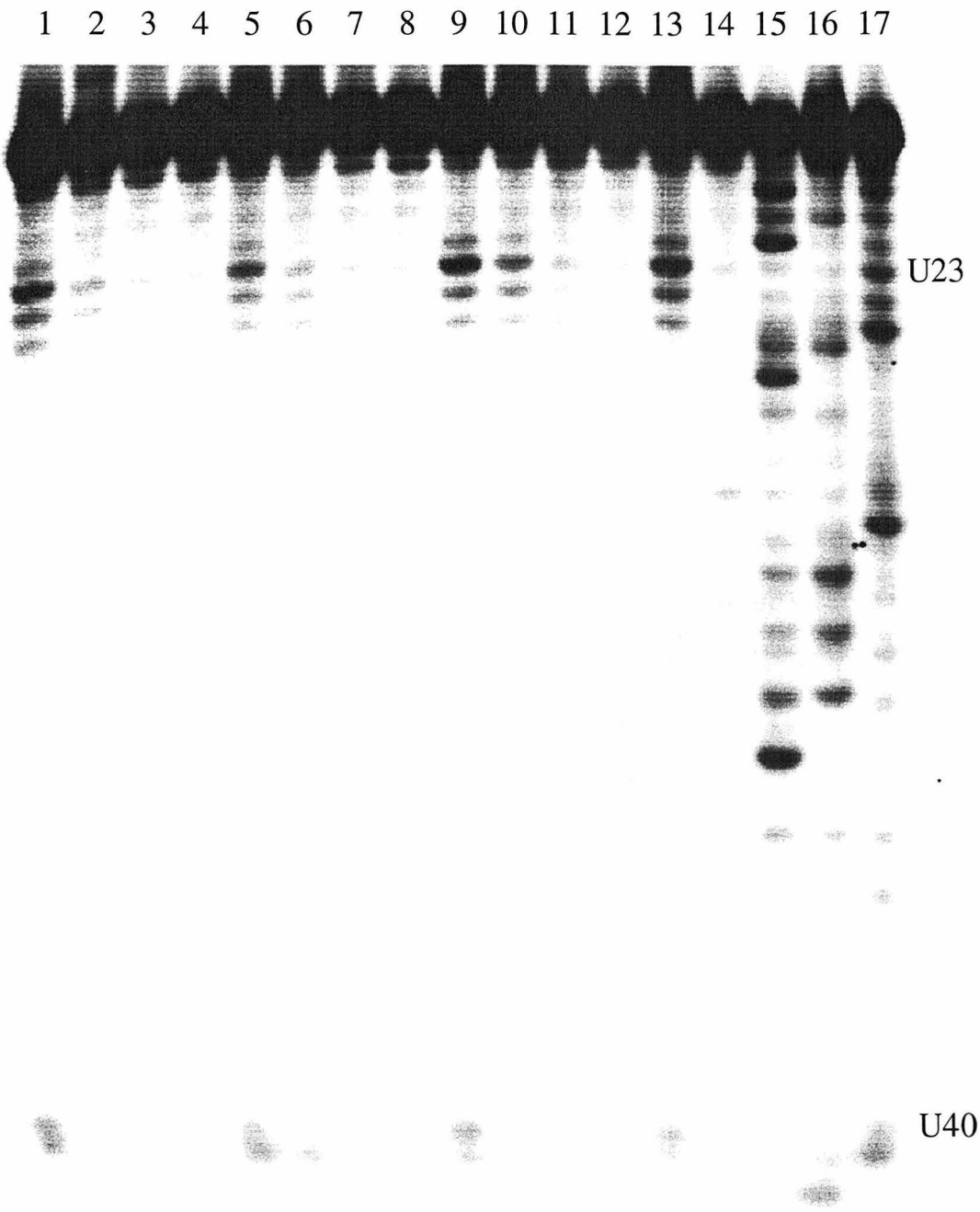
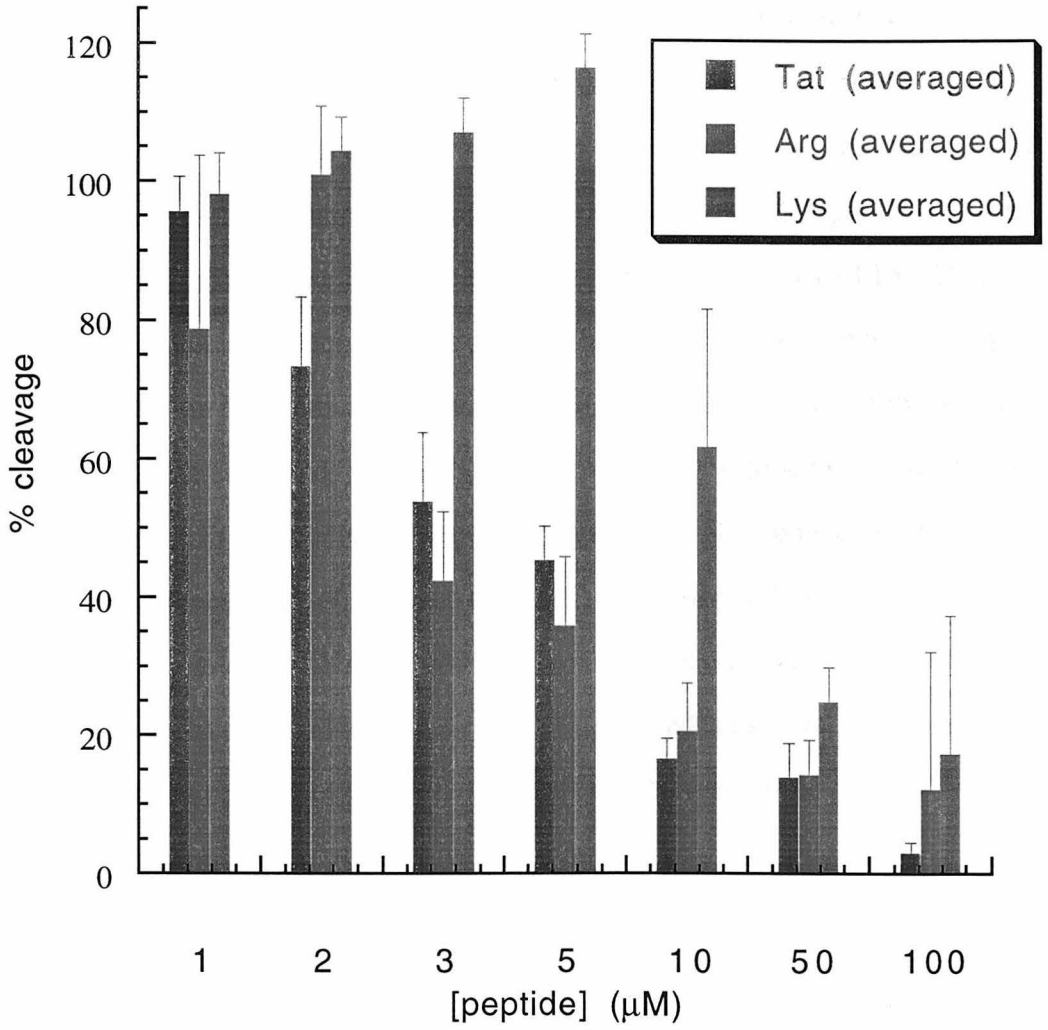

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
U23
U40

Figure 4.15. Inhibition of racemic 3-Rh(MGP)$_2$phi^{5+} photocleavage induced by the Tat-9, Arg-11 and Lys-11 peptides. Photocleavage at G26 by 5 μM 3-Rh(MGP)$_2$phi^{5+} in the presence of increasing concentrations of peptide was quantitated by phosphorimager, and normalized to photocleavage at that site with 5 μM 3-Rh(MGP)$_2$phi^{5+} alone.


120
100
80
60
40
20
0
% cleavage
Tat (averaged)
Arg (averaged)
Lys (averaged)
1
2
3
5
10
50
100
[peptide] (μM)

by 50%, while the same concentration of Lys-11 did not inhibit photocleavage at all. At 10 μM Lys-11, twice the concentration of metal complex, cleavage was reduced by less than 40%. In contrast, Arg-11 appeared to inhibit cleavage to a similar degree as the Tat-9 peptide.

The specific inhibition of rhodium photocleavage by the Tat-9 and Arg-11 peptides indicate that these metal complexes are competing for the same binding sites on the RNA that the peptides are. Thus, these metal complexes could be recognizing or inducing the same structural change peptides do, albeit by making different contacts with the RNA.

The different charge of the various peptides means that they could interact with the RNA or the metal complex in different fashions. The metal complex is highly positively charged, and thus should interact more strongly with the Lys-11 peptide than with Tat-9 or Arg-11. Thus, the decrease in photocleavage noted in the presence of Lys-11 could reflect rhodium complex interaction with the peptide, instead of competition for binding sites on the RNA between the rhodium complex and the peptide. Incubation of the peptide and the RNA for an hour before addition of the metal complex was done to discourage this from happening. The lifetime of the wild type peptide-TAR RNA complex is on the order of 10 min in the presence of 20-60 nM competitor;[30] in addition, it has been shown that the preformed species of an RNA-Tat 24mer complex containing the Tat-binding region is stable to competition by a 100-fold-excess of unlabeled competitor, with a halflife of ≥ 2 min.[4] Thus, the hour of preincubation should be sufficient for the peptide-TAR complexes to reach equilibration.

4.3.5. EtNU Cleavage on TAR RNA

It has been shown that the ethylation of two phosphates, between G21 and A22, and A22 and U23, located at the junction of the double-stranded stem and bulge, interferes strongly with peptide and arginine binding.[6,7] Thus, if the guaninidium arms on

$Rh(MGP)_2phi^{5+}$ and $Rh(GEB)_2phi^{5+}$ are contacting these phosphates specifically, the metal complexes could protect the RNA from EtNU cleavage.

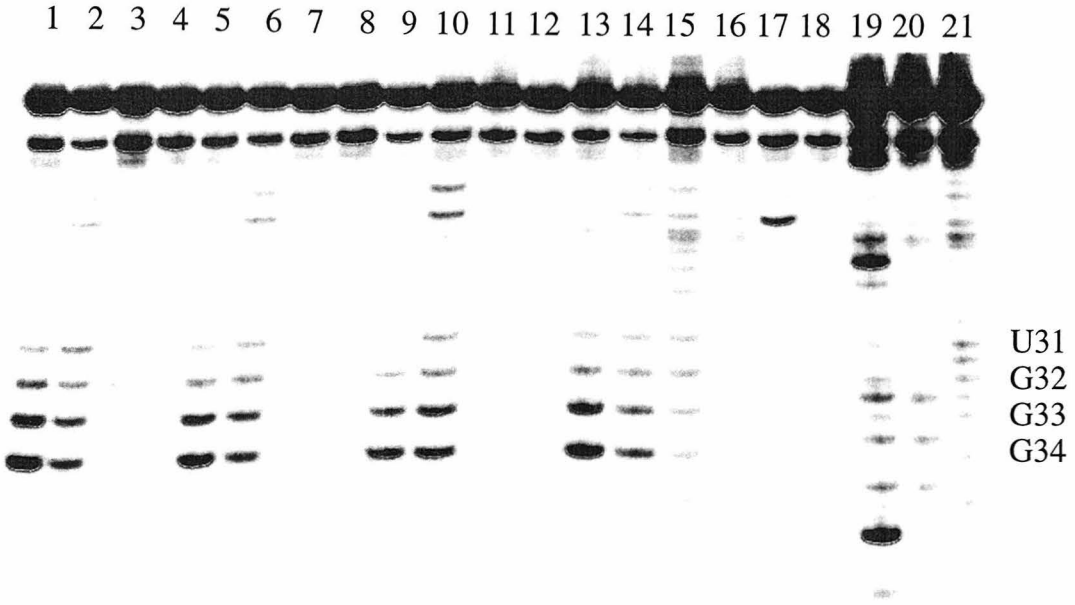
As shown in Figure 4.16, there is faint EtNU cleavage on the RNA in the absence of rhodium complex. Cleavage is non-specific throughout the RNA, and increases with increasing temperature and incubation time. However, when EtNU cleavage is carried out in the presence of $Rh(MGP)_2phi^{5+}$, there is increased cleavage in the loop region, at residues C30, U31, G32 and G33, compared to EtNU alone. Cleavage is strongest at G32 and G33. This effect is seen both in the presence and the absence of carrier RNA. This effect could be caused by the rhodium complex binding to the loop region and inducing a conformational change which sensitizes the loop region to EtNU cleavage. However, the loop region is not the strongest rhodium photocleavage site. An intriguing possibility for this increase in photocleavage is that the binding of the rhodium complex to the bulge region could result from a long-range conformational change in the RNA making the loop region more susceptible to EtNU cleavage. This is reminiscent of our observation of enhanced RNase cleavage on triple helices, where addition of a third strand to a duplex hairpin sensitizes the hairpin loop region to RNase cleavage (Chapter 3).

In the absence of carrier RNA, cleavage is also seen at the residues in the bulge region, at U26 and C25. Thus, it appears that the rhodium complex sensitizes the TAR WT to EtNU cleavage in the non-canonical A-form RNA regions.

4.3.6. Circular Dichroism Spectroscopy of Metal-RNA Complexes

The circular dichroism spectrum of TAR RNA exhibits a maximum at 265 nm, and minima at 240 and 210 nm.[31] It has been shown that the binding of the Tat peptide[9] or arginine-rich peptides arginine[31] produce a 15% decrease in the CD spectrum near the maximum of 265 nm. The nature of this conformational change is unknown, but it has been suggested that the loss in signal is associated with bases in the bulge becoming unstacked upon the binding of arginine.[31] Mutants of TAR which cannot form a base triple

Figure 4.16. Grey-scale phosphorimager representation of 20% polyacrylamide gel showing the effect of racemic 3-Rh(MGP)$_2$phi^{5+} on EtNU cleavage in loop region of 3'-^{32}P-labeled TAR WT RNA. Lane 1: 5 μM rhodium complex, 100 μM EtNU, 100 μM tRNA carrier, incubated for 5' at room temperature. Lane 2: 5 μM rhodium complex, 100 μM EtNU, no carrier, incubated for 5' at room temperature. Lane 3: 100 μM EtNU, 100 μM tRNA carrier, incubated for 5' at room temperature. Lane 4: 100 μM EtNU, no carrier, incubated for 5' at room temperature. Lanes 5-8: identical to lanes 1-4 respectively, but incubated for 5' at 37°. Lanes 9-12: identical to lanes 1-4 respectively, but incubated for 5' at 55°. Lanes 13-16: identical to lanes 1-4 respectively, but incubated for 60' at 55°. Lanes 19 & 20: 5 μM rhodium complex, incubated at 55° for 60', 100 μM tRNA and no carrier respectively. Lanes 19, 20, 21: A-, G- and U-sequencing lanes respectively.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
U31
G32
G33
G34

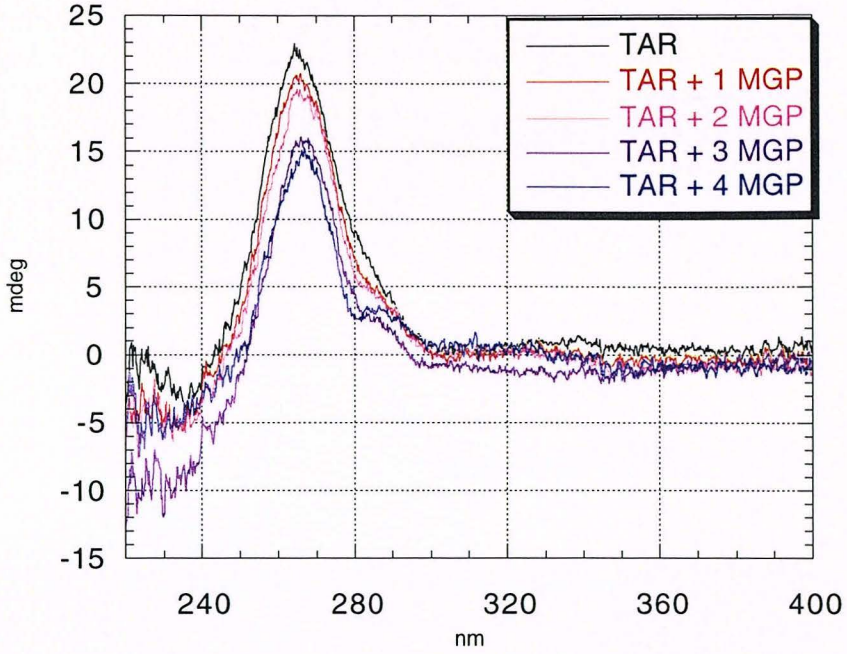
do not exhibit this feature.[31]

The binding of racemic $Rh(MGP)_2phi^{5+}$ to WT TAR was studied by circular dichroism spectroscopy. The free spectrum of TAR RNA was identical to that reported in the literature, exhibiting a maximum at 265 nm, and minima at 240 and 210 nm.[31] There was a significant decrease in the CD maximum at 265 nm upon the addition of rhodium complex. This decrease leveled off at approximately 30% at 4 equivalents of metal complex, as shown in Figure 4.17. When 10 mM argininamide was added, a further 10% decrease in ellipticity was observed. This decrease, when observed upon the addition of arginine, guanidine or Tat peptide in the literature, was attributed to a change in the conformation of the RNA upon the specific binding of guanidinium group.[5,31] When the argininamide was added first, and then the rhodium complex, a decrease (approximately 18%) was seen in the ellipticity upon addition of 10 mM argininamide, and a negligible further decrease upon the addition of rac-3-$Rh(MGP)_2phi^{5+}$ (Figure 4.18).

No decrease in the CD spectrum of TAR1 was seen upon the addition of metal complex (Figure 4.19), consistent with there being no change in the CD spectrum of this mutant upon binding of peptides, arginine or guanidine.[5,31] In addition, no change was noted in the CD spectrum of WT TAR upon the addition of diaminophenanthrene, indicating that the ligand had to be appended to the rhodium complex.

To resolve whether the decrease in ellipticity noted upon addition of rac-3-$Rh(MGP)_2phi^{5+}$ was due to the guanidinium arms on the complex contacting the G26-C39 basepair, or due to intercalation of the phi ligand, or both, $Rh(phen)_2phi^{3+}$ was titrated into TAR RNA. A total of 4 equivalents of racemic $Rh(phen)_2phi^{3+}$ was titrated into a solution of TAR RNA. Titration with racemic $Rh(phen)_2phi^{3+}$ produced less of a decrease in ellipticity than the $Rh(MGP)_2phi^{5+}$ complex, but nevertheless still produced approximately a 10% decrease in signal (Figure 4.20). Argininamide to 10 mM was then titrated in, upon which the ellipticity decreased further. Thus, these results are consistent with intercalation

Figure 4.17. Circular dichroism spectra of WT TAR RNA showing A) Racemic 3-$Rh(MGP)_2phi^{5+}$ titration from 1 equivalents to 4 equivalents, and rhodium complex induced decrease in ellipticity at 265 nm B) Further decrease in ellipticity upon addition of 10 mM argininamide.


TAR
TAR + 1 MGP
TAR + 2 MGP
TAR + 3 MGP
TAR + 4 MGP
mdeg
25
20
15
10
5
0
-5
-10
-15
240
280
320
360
400
nm

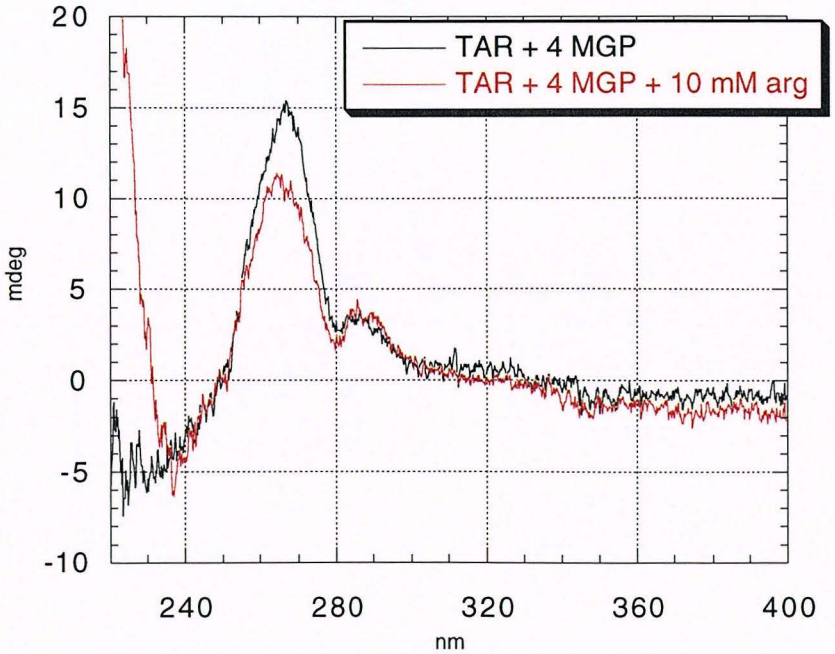
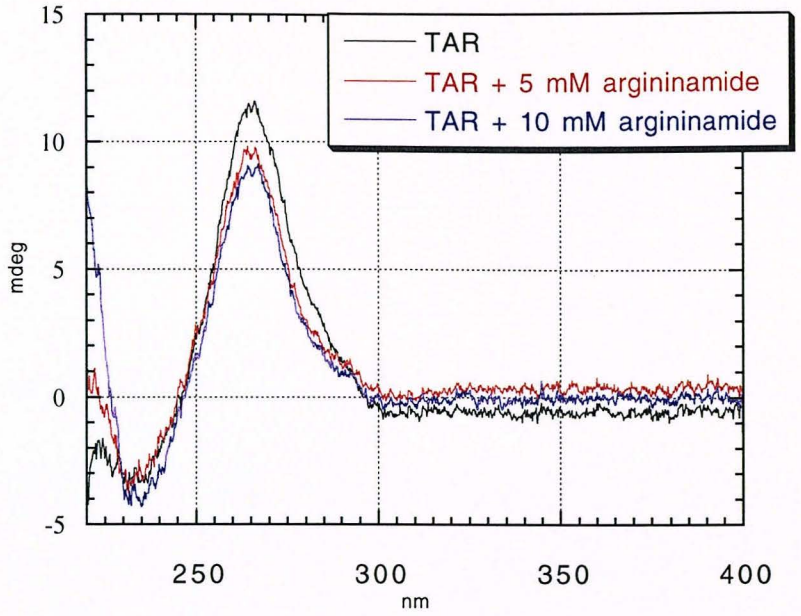

TAR + 4 MGP
TAR + 4 MGP + 10 mM arg
mdeg
20
15
10
5
0
-5
-10
240
280
320
360
400
nm

Figure 4.18. Circular dichroism spectra of WT TAR RNA showing A) Decrease in ellipticity at 265 nm upon the addition of 5 and 10 mM argininamide respectively. B) Further decrease in intensity upon the addition of 1, 2 and 3 equivalents racemic 3-$Rh(MGP)_2phi^{5+}$ to WT TAR RNA and 10 mM argininamide.


TAR
TAR + 5 mM argininamide
TAR + 10 mM argininamide
15
10
5
0
-5
mdeg
250
300
350
400
nm

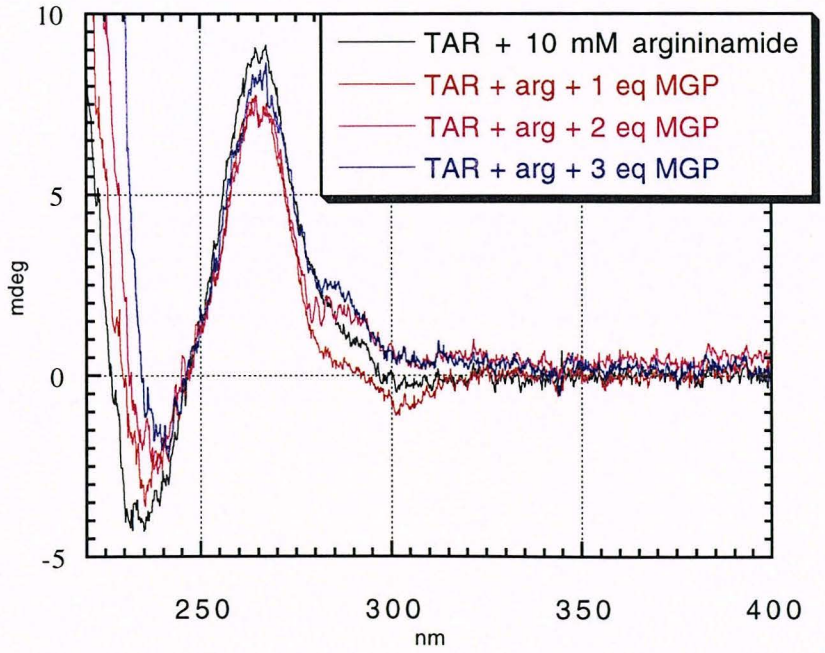
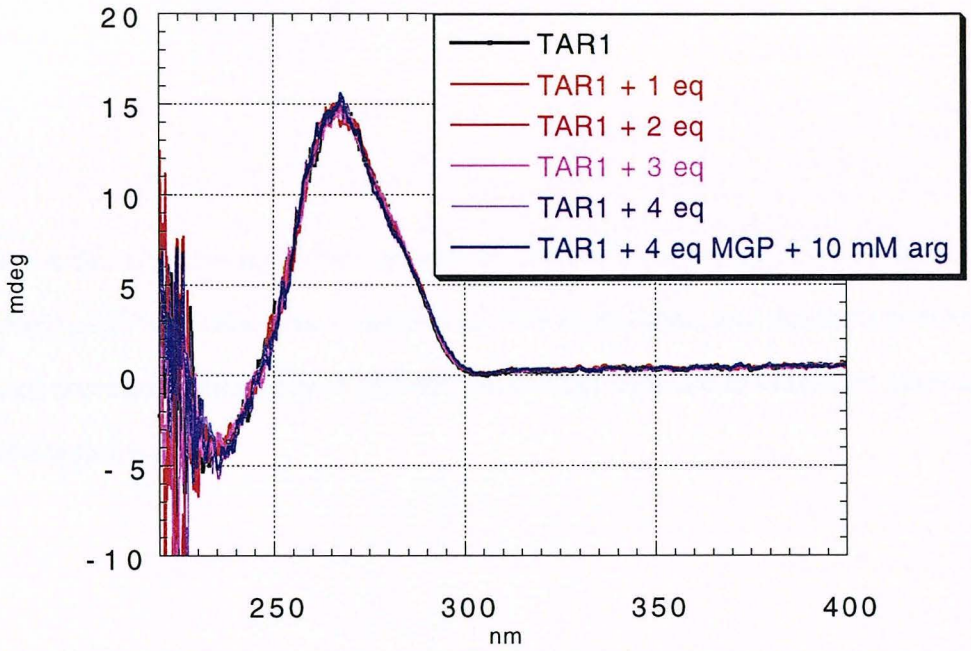

TAR + 10 mM argininamide
TAR + arg + 1 eq MGP
TAR + arg + 2 eq MGP
TAR + arg + 3 eq MGP
10
5
0
-5
mdeg
250
300
350
400
nm

Figure 4.19. Circular dichroism spectra of TAR1 RNA showing no change in ellipticity at 365 upon the addition of racemic 3-Rh(MGP)$_2$phi^{5+}from 1 equivalents to 4 equivalents, and further addition of 10 mM argininamide.


20
15
10
5
0
- 5
-10
mdeg
250
300
350
400
nm
TAR1
TAR1 + 1 eq
TAR1 + 2 eq
TAR1 + 3 eq
TAR1 + 4 eq
TAR1 + 4 eq MGP + 10 mM arg

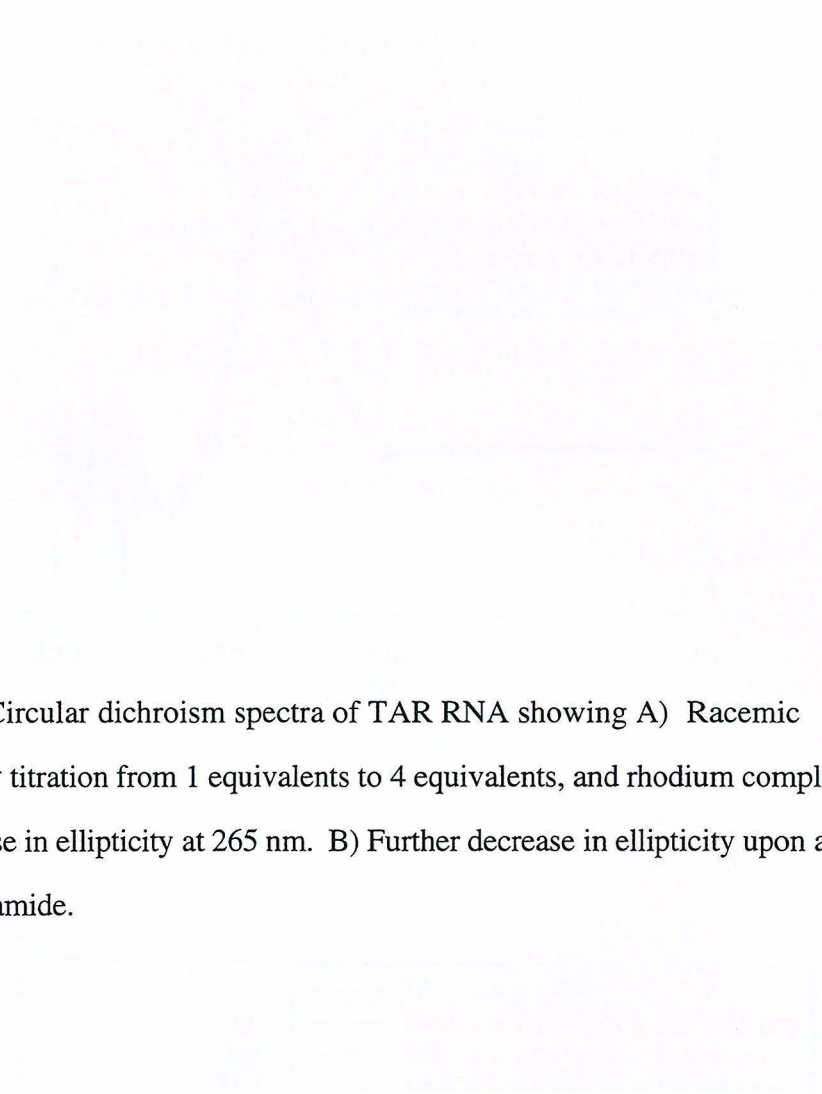
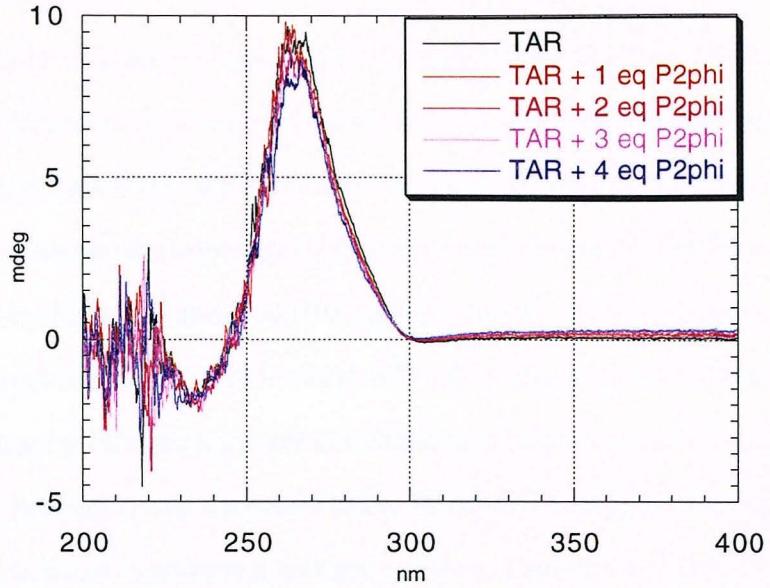
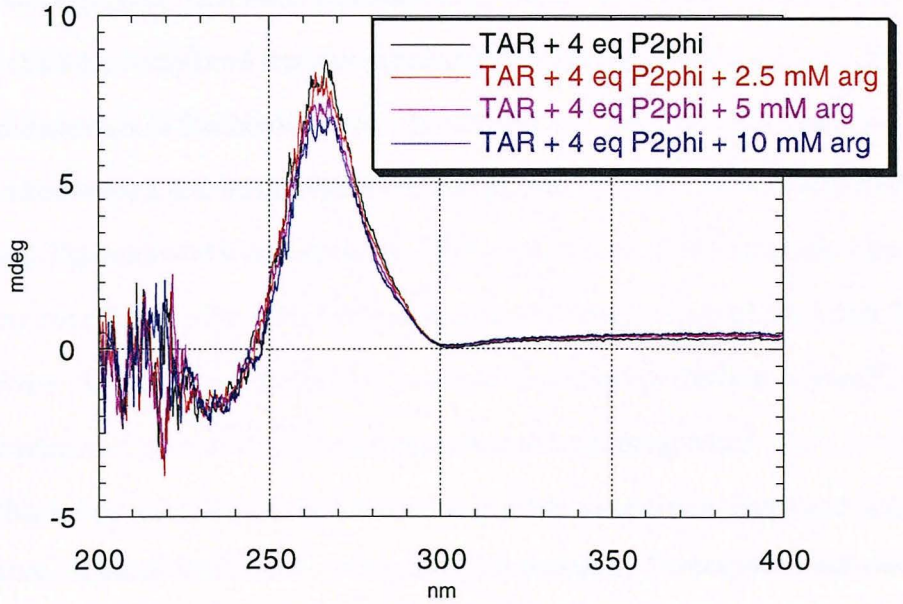


Figure 4.20. Circular dichroism spectra of TAR RNA showing A) Racemic $Rh(phen)_2phi^{3+}$ titration from 1 equivalents to 4 equivalents, and rhodium complex induced decrease in ellipticity at 265 nm. B) Further decrease in ellipticity upon addition of 10 mM argininamide.


10
5
0
-5
mdeg
200
250
300
350
400
nm
TAR
TAR + 1 eq P2phi
TAR + 2 eq P2phi
TAR + 3 eq P2phi
TAR + 4 eq P2phi


10
5
0
-5
mdeg
200
250
300
350
400
nm
TAR + 4 eq P2phi
TAR + 4 eq P2phi + 2.5 mM arg
TAR + 4 eq P2phi + 5 mM arg
TAR + 4 eq P2phi + 10 mM arg

of the phi ligand on the metal complex inducing the same change in TAR RNA structure, as probed by circular dichroism, as does the Tat peptide.

4.2.7 Gel-shift Assay of Rhodium Complex Binding to TAR RNA.

Gel shift assays have been successfully used to monitor the binding of Tat fragments to the TAR RNA and to obtain a dissociation constant for the Tat-TAR complex.[30] However, analogous gel shift assays with the metal complexes pose a problem, as they have less mass than the peptides and are positively charged; hence, obtaining an appreciable gel shift is made difficult. Previous work in these laboratories has shown that these metal complexes are not effective in retarding small fragments or DNA oligomers.[32] Instead, it was necessary to use larger DNA fragments, of approximately kilobase size, in order to observe a retardation effect. Even at very high metal complex to RNA ratios, there was only a slight retardation effect of the metal complex on RNA.

4.4. Discussion

Recognition of TAR RNA by metal complexes. This study has shown that synthetic rhodium complexes can recognize and specifically cleave areas on TAR RNA which are important in Tat binding and transactivation. The bases U23 and U40 are the strongest sites recognized by this family of complexes. U23 is crucially required in Tat binding and Tat-mediated transactivation.[3,4,33] In addition, G26 is strongly cleaved by these metal complexes. The G26-C39 basepair is absolutely required for intact Tat,[8] Tat peptide[34] and arginine[7] binding *in vitro*, as well as for transactivation *in vivo*.[8] G26 has also been proposed to interact with the arginine guanidinium group.[9]

This bulge region is a target of the Tat peptide and various drugs including intercalators. It has already been shown that $Rh(phen)_2phi^{3+}$ recognizes and cleaves this region.[16] The nonprotein chromophore of the antitumor drug neocarzinostatin (NCS-chrom) generates a distinct but weak band due to cleavage at U24 in the bulge.[35] Electric

linear dichroism data indicate that the mode of binding of the drugs Hoechst 33258, berenil and DAPI to the TAR RNA likely involves intercalation into the A-form TAR RNA helix where the major groove is widened.[36] However, the photoreactivity of these metal complexes and their ability to mark their position of intercalation upon irradiation with UV light make them uniquely useful tools of RNA structure.

Correlation between cleavage and structure on TAR mutants. Studies with mutants of TAR RNA have shown that these metal complexes can sensitively detect the presence of a base triple, the formation of which is important for transactivation *in vivo*. It is noteworthy that the metal complexes do not recognize the RNA double helix upon the deletion of the bulge. Importantly, these complexes are also sensitive to mutants such as TAR1 which should not affect the secondary structure of the RNA, suggesting that they recognize a specific three-dimensional structure formed by the intact, functional RNA. The difference in rhodium complex cleavage between TAR1 and TAR WT suggests that not only does the metal complex recognize a bulge structure, but it also recognizes the structured and unstructured bulge sequences differently.

Binding affinity of metal complexes for TAR RNA. These metal complexes bind with high affinity to TAR RNA, with K_d's in the micromolar regime. The binding constants reported in Table 4.3 were obtained at the strongest cleavage site, U40, but similar binding constants were obtained when quantitation was carried out based on U23 or G26 as well. Thus, these metal complexes bind in the micromolar range to sites which have been shown to be crucial for transactivation.

Conformational change in CD spectrum induced by metal complexes. It is noteworthy that the gross change in conformation in the CD spectrum attributed to specific arginine binding to the base triple is also induced by the addition of rhodium complex to TAR RNA. $Rh(phen)_2phi^{3+}$ alone induces the same change in the RNA CD spectrum that $Rh(MGP)_2phi^{5+}$ does, though to a lesser extent. Therefore, this suggests that the metal

Table 4.3. Binding constants of rhodium complexes to TAR WT RNA as determined by photocleavage affinity titrations at U40.

Metal complex species	K_b (M^{-1}) x 10^{-5}
racemic $Rh(phen)_2phi^{3+}$	4.0 (0.8)
Δ-$Rh(phen)_2phi^{3+}$	6.1 (1.3)
Λ-$Rh(phen)_2phi^{3+}$	4.6 (1.8)
1Δ-$Rh(MGP)_2phi^{5+}$	7.8 (2.2)
2Δ-$Rh(MGP)_2phi^{5+}$	8.0 (2.7)
3Δ-$Rh(MGP)_2phi^{5+}$	9.9 (1.0)
3Λ-$Rh(MGP)_2phi^{5+}$	5.6 (1.2)
1Δ-$Rh(GEB)_2phi^{5+}$	7.4 (2.0)
1Λ-$Rh(GEB)_2phi^{5+}$	3.9 (1.4)
2Δ-$Rh(GEB)_2phi^{5+}$	6.0 (1.2)
2Λ-$Rh(GEB)_2phi^{5+}$	2.1 (0.9)
3Δ-$Rh(GEB)_2phi^{5+}$	8.1 (3.6)
3Λ-$Rh(GEB)_2phi^{5+}$	4.9 (1.7)

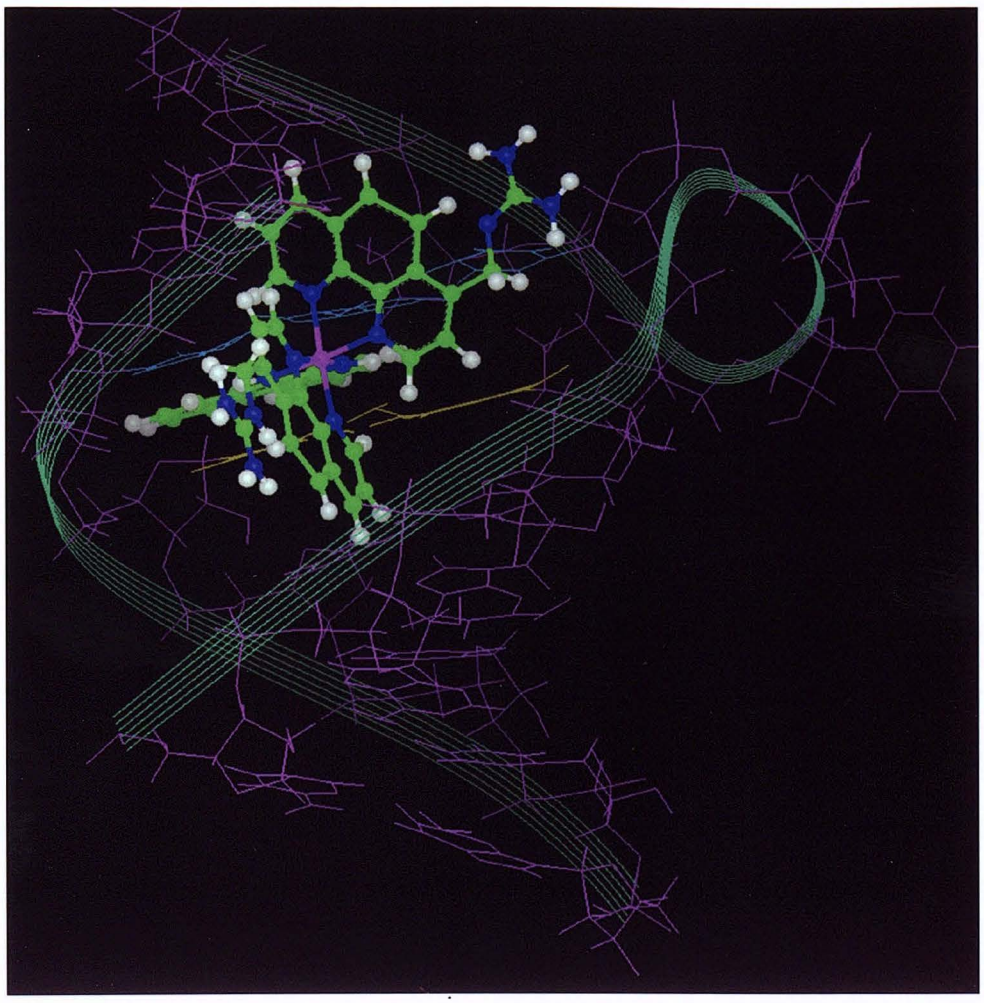
complex, even in the absence of the guanidinium arms, is capable of inducing the same structural change that arginine and the Tat peptide does.

Role of the guanidinium groups in metal complex recognition of TAR RNA. The guanidinium groups on the $Rh(MGP)_2phi^{5+}$ and $Rh(GEB)_2phi^{5+}$ complexes appear to confer additional binding affinity by the metal complex for TAR RNA by making non-specific interactions with the phosphates on the RNA backbone. There are several lines of evidence which support this, one being the similarity between the $Rh(phen)_2phi^{3+}$ and $Rh(MGP)_2phi^{5+}$ photocleavage sites on TAR WT. The addition of the guanidinium groups increases the binding constant of the rhodium complex to the TAR RNA by less than an order of magnitude. This corresponds to a free energy of approximately 0.5-1 kcal/mol, depending on the complex. The guanidinium groups do not appear to be involved in specific base-hydrogen bonding interactions, as $Rh(MGP)_2phi^{5+}$ photocleavage sites on TAR WT and TAR4 are similar. Since TAR4 lacks the G26-C39 base pair which an arginine group on the Tat peptide has been proposed to contact, this would indicate that this base pair is not involved in specific metal complex-RNA recognition.

The changes observed in the CD spectra are also consistent with there being non-specific interactions between the guanidinium groups and the RNA. As mentioned in the previous section, $Rh(phen)_2phi^{3+}$ alone induces the same change in the RNA CD spectrum that $Rh(MGP)_2phi^{5+}$ does, although there is a greater decrease in ellipticity with the latter complex due to the additional affinity that the guanidinium groups provide. The electrostatic interactions of the 9mer basic domain with the phosphates on the RNA have been calculated to account for 80% of the total binding energy of this basic region to TAR.[37] Therefore, in the case of these guanidinium-containing metal complexes, it is likely that the guanidinium arms are incorrectly positioned to make the specific contact with G26, but are still able to interact with the phosphate backbone.

Model of TAR RNA interaction with 3Δ-$Rh(MGP)_2phi^{5+}$ A model of the interaction of 3Δ-$Rh(MGP)_2phi^{5+}$ with TAR RNA is shown in Figure 4.21.

Figure 4.21. Model of interaction of 3Δ-Rh(MGP)$_2$phi^{5+} with TAR RNA, showing the phi ligand intercalated within the bulge between G26-C39 and U38•A27-U23, and non-specific contacts between the guanidinium arms and the phosphate groups on the backbone of the RNA. Coordinates for RNA courtesy of J. Williamson.



The features represented by this model include intercalation from the major groove by the rhodium complex in the bulge, between G26-C39 and U38•A27-U23. This close intercalation would permit cleavage at G26, C39, and U23 in the metal bound complex. The guanidinium arms of $Rh(MGP)_2phi^{5+}$ make non-specific contacts to the phosphate backbone. Because of the flexibility of the guanidinium arms, all three isomers should be able to make such non-specific contacts. It can also be seen that the Δ-complex provides better shape-complementarity to the groove than would the Λ-isomer. Thus, in all cases, we observe Δ-selectivity in the cleavage patterns.

This model provides a starting point reflecting relative the sizes of the RNA and the metal complex, as well as the orientation of the metal complex. Since strong cleavage is evident at U40, it is likely that the intercalation site could be one basepair lower. The same structural constraints would apply, however.

It is also likely that the metal complex is binding at sites one basepair above and below this binding site. This would be consistent with the photocleavage sites at C41 and U23 respectively.

Recognition of the loop region of TAR RNA by rhodium complexes, and by EtNU in the presence of rhodium complex. The role of the loop region in the TAR-Tat complex is less defined than that of the bulge region. There is evidence that this region functions by making non-specific contacts with the peptide residues in the 59-72 region of the Tat protein. RNase studies have shown that the intact Tat protein protects both the bulge and loop regions of TAR.[38] Loop residue mutations impede the binding of a longer Tat peptide, but not a short Tat 11-mer containing the basic recognition sequence.[39] In addition, mutations in the loop region[40] decrease Tat function *in vivo*. Thus, the loop region is important for transactivation *in vivo* but is not contacted specifically by the Tat protein.[40,41]

This largely sequence-independent interaction of the loop region with Tat is mimicked by the metal complexes. The weak photocleavage of the bases in the loop by the

metal complexes means that they could be contacting the loop in a non-specific manner. Perhaps more intriguing is the possibility that these metal complexes could be inducing a conformational change in the loop region upon binding to the bulge, as evidenced by the sensitization of the loop to EtNU cleavage in the presence of 3Δ-Rh(MGP)$_2$phi^{5+}. Whether such a conformational change arises upon probable binding needs to be explored.

Comparison of peptide and metal complex binding to RNA. As can be seen from inhibition studies of metal complex photocleavage by the peptides, these metal complexes are competitors of Tat peptide binding to RNA. In order to inhibit 3Δ-Rh(MGP)$_2$phi^{5+} photocleavage by 50%, approximately equal concentrations of Tat-9 or Arg-11 peptide as metal complex are required. A range of different binding affinities have been obtained for the Tat peptide to its target RNA; it has been estimated that the K_d for the Tat peptide is in the micromolar[42] or nanomolar[31] regime by gel-shift assays, or even in the picomolar regime by circular dichroism studies.[30] However, it was not possible to carry out the analogous gel shift assay on the RNA in the presence of the metal complexes (*vide infra*), and so a direct comparison between the respective metal complex and peptide binding constants can not be done. Nevertheless, the binding constants obtained for these metal complexes via photoaffinity cleavage titrations are comparable to the lower estimate of binding affinities (K_d = 0.78 μM)[42] obtained via gel-shift assays. Thus, not only are their binding affinities comparable, but photocleavage competition studies indicate that the rhodium complex does indeed compete successfully for the Tat binding site on TAR RNA. Importantly, the metal complex and Tat target the same site on the RNA polymer and promote analogous conformational changes.

Conclusions. The Tat-peptide-TAR RNA interaction makes use of several features that are seen in other protein-RNA interactions: the use of direct readout by the guanidinium contacting the G26 base, and an induced fit, by the formation of the U23•A27-U38 triple base.[9] We find that our small metal complexes can functionally mimic some of the interactions of the peptide-RNA complex, and in fact, serve to inhibit specific peptide-RNA

binding. Importantly, metal complex cleavage is seen at several sites in the RNA which are important for transactivation. The metal complex-RNA interaction is enantioselective, underscoring the importance of matching the shape of the metal complex to the right-handed helix of the RNA. The affinity titrations permit a careful dissection of the elements which contribute to metal complex-RNA recognition, and we find that the intercalation of the phi ligand is the main driving force for this interaction. The addition of guanidinium groups to the ancillary ligands provide 0.5 - 1 kcal/mol additional binding affinity through what is likely a non-specific interaction with the phosphate backbone.

Nevertheless, these rhodium complexes, especially $Rh(MGP)_2phi^{5+}$, are highly specific probes of the folded TAR RNA structure, and inhibit Tat binding efficiently. This inhibition of Tat binding is likely the result of the functional targeting of sites on the RNA by the metal complex. These metal complexes provide a scaffold upon which different ancillary ligands can be appended, and the functional groups on them varied. Thus, there is scope for the design and application of metal complexes with higher specificity and tighter binding affinity to biologically important RNA molecules. The specificity of these metal complexes and their ability to disrupt protein-RNA interactions make them potential pharmaceuticals.

References

1. Frankel, A. D. *Curr. Opin. Genet. Dev.* **1992**, *2*, 293-298.
2. Cullen, B. *Cell* **1990**, *63*, 655-657.
3. Roy, S.; Delling, U.; Chen, C. H.; Rosen, C. A.; Sonenberg, N. *Genes & Development* **1990**, *4*, 1365-1373.
4. Weeks, K. M.; C., A.; Schultz, S. C.; Steitz, T. A.; Crothers, D. M. *Science* **1990**, *249*, 1281-1285.

5. Calnan, B. J.; Biancalana, S.; Hudson, D.; Frankel, A. D. *Genes & Development* **1991**, *5*, 201-210.

6. Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D. *Science* **1991**, *252*, 1167-1171.

7. Tao, J.; Frankel, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2723-2786.

8. Delling, U.; Reid, L. S.; Barnett, R. W.; Ma, M. Y. X.; Climie, S.; Sumnersmith, M.; Sonenberg, N. *J. Virology* **1992**, *66*, 3018-3025.

9. Puglisi, J. D.; Tan, R.; J., C. B.; Frankel, A. D.; Williamson, J. R. *Science* **1992**, *257*, 76-80.

10. Aboulela, F.; Karn, J.; Varani, G. *Nucleic Acids Res.* **1996**, *24*, 3974-3981.

11. Zacharias, M.; Hagerman, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6042-6056.

12. Brodsky, A. S.; Williamson, J. R. *J. Mol. Biol.* **1997**, *267*, 624-639.

13. David, S. D.; K., B. J. *J. Am. Chem. Soc.* **1993**, *115*, 2984-2985.

14. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303-2312.

15. Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972-982.

16. Neenhold, H. R.; Rana, T. M. *Biochemistry* **1995**, *34*, 6303-6309.

17. Terbrueggen, R. H.; Barton, J. K. *Biochemistry* **1995**, *34*, 8227-8234.

18. Miligan, J. F.; Groebe, D. R.; Witherall, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783-8798.

19. England, T. E.; Bruce, A. G.; Uhlenbeck, O. C. *Meth. Enzym.* **1980**, *65*, 65-74.

20. Maxam, A. M.; Gilbert, W. *Meth. Enzym.* **1988**, *65*, 499.

21. Chow, C. S.; Barton, J. K. *Meth. Enzym.* **1992**, *212*, 219-242.

22. Terbrueggen, R. H.; Johann, T. W.; Barton, J. K. manuscript in preparation **1997**.

23. Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839-2841.

24. Peattie, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1760.

25. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303-2312.

26. Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995-11003.

27. Jackson, B. A.; Hastings, C. A.; Hudson, B. P.; Johann, T. W. manuscript in preparation **1997**.

28. Weeks, K. M.; Crothers, D. M. *Science* **1993**, *261*, 1574-1577.

29. Krotz, A. H.; Kuo, L. Y.; Barton, J. K. *Inorg. Chem.* **1993**, *32*, 5963- 5974.

30. Long, K. S.; Crothers, D. M. *Biochemistry* **1995**, *34*, 8885-8895.

31. Tan, R.; Frankel, A. *Biochemistry* **1992**, *31*, 10288-10294.

32. Campisi, D., Doctoral Thesis, Dept. of Chemistry, California Institute of Technology: Pasadena, 1995.

33. Sumner-Smith, M.; Roy, S.; Barnett, R.; Reid, L. S.; Kuperman, R. A.; Delling, U.; Sonenberg, N. *J. Virology* **1991**, *65*, 5196-5202.

34. Weeks, K. M.; Crothers, D. M. *Cell* **1991**, *66*, 577-588.

35. Kappen, L. S.; Goldberg, I. H. *Biochemistry* **1995**, *34*, 5997-6002.

36. Bailly, C.; Colson, P.; Houssier, C.; Hamy, F. *Nucleic Acids Res.* **1996**, *24*, 1460-1464.

37. Loret, E. P.; Georgel, P.; Johnson Jr., W. C.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9734-9738.

38. Harper, J. W.; Logsdon, N. J. *Biochemistry* **1991**, *30*, 8060-8066.

39. Kamine, J.; Loewenstein, P.; Green, M. *Virology* **1991**, *182*, 570-577.

40. Feng, S.; Holland, E. C. *Nature* **1988**, *334*, 165-167.

41. Berkhout, B.; Jean, K.-T. *J. Virol.* **1989**, *63*, 5501-5504.

42. Wang, X.; Huq, I.; Rana, T. M. *J. Am. Chem. Soc.* **1997**, *119*, 6444-6445.

Chapter 5.

Targeting the Tat-binding Site of Bovine Immunodeficiency Virus TAR RNA with a Shape-Selective Rhodium Complex†

5.1 Introduction

The *trans*-activating (Tat) protein of bovine immunodeficiency virus (BIV) binds the *trans*-activation response (TAR) RNA hairpin element located at the 5' end of the viral mRNA and thereby activates transcription.[1] Recognition of this RNA site has been found to depend upon the folding of the RNA into a remarkable structure in which an intramolecular base triple is formed and the RNA major groove is widened to accept the Tat β-hairpin peptide. Recent high resolution NMR studies[1-3] have served to elucidate specific contacts between the RNA bases and amino acids of the Tat peptide in the opened major groove of RNA. Additionally, NMR data have been used to establish the intramolecular base triple, U10•A13-U24, in the presence of the specifically bound BIV Tat peptide SGPRPRGTRGKGRRIR.[3] Bases involved in maintaining this fold have furthermore been found to be essential for *in vivo* expression.[1] This interesting RNA fold resembles that found by NMR[4] in the recognition of human immunodeficiency viral (HIV) TAR RNA by its associated Tat protein, despite differences in the structure and sequences of the RNA bulge and loop region.[1]

Our interest is in the design of octahedral transition metal complexes which bind nucleic acids with site-selectivity.[5] On tRNAs, $Rh(phen)_2phi^{3+}$ has been shown specifically to target only sites involved in tertiary interactions.[6] Mutations which disrupt RNA folding disrupt targeting by the metal complex, but base changes which maintain folding also maintain site-specific reaction. Based upon the unique shape-selective targeting by $Rh(phen)_2phi^{3+}$, the metal complex has already been applied in probing HIV

† Adapted from Lim, A. & Barton, J. K. *Bioorganic Med. Chem.* **1997**, 1131-1136.

TAR RNA.[7] In addition, the application of $Rh(phen)_2phi^{3+}$, $Rh(MGP)_2phi^{5+}$, and $Rh(GEB)_2phi^{5+}$ to target TAR RNA has also been detailed in Chapter 4. Here, as well, site-specific cleavage is evident neighboring the site of RNA tertiary interaction within the opened major groove. Interestingly, for this RNA, adoption of the folded base triple has been found to be induced by binding of HIV Tat peptide; site-specific cleavage by the metal complex does not, however, require Tat peptide and may instead serve to drive the conformational change to the folded RNA element.

Although the BIV Tat is closely related to the HIV TAR, there are differences in the secondary structures of the RNA molecules as shown in Figure 5.1. Nevertheless, the folded structure of BIV Tat contains an intramolecular base triple, U10•A13-U24,[2] which is similar to that found in HIV TAR RNA.[4] In addition, the formation of this base triple facilitates hydrogen bonding of an arginine residue to a guanine residue and to a backbone phosphate,[3] suggesting a similar mode of molecular recognition as in the HIV Tat-TAR interaction. Thus, a comparison of the recognition characteristics of these metal complexes on BIV compared to HIV TAR RNA should provide insight into both the RNA structures themselves, and the mode of their recognition by the metal complexes.

Here we describe the site-specific targeting of BIV TAR RNA by $Rh(phen)_2phi^{3+}$ and $Rh(MGP)_2phi^{5+}$, and show that specific cleavage is induced by the rhodium complex with photocleavage at U24. In addition, BIV1 TAR RNA, which does not adopt the folded conformation involving a base triple, was also used as a substrate for these metal complexes. A comparison of the structures of these RNA molecules is shown in Figure 5.1. Furthermore, we demonstrate that binding to this site by the metal complex can compete with specific binding of the Tat peptide. Hence, this shape-selective targeting may provide a route not only to probe RNA structures but also to inhibit RNA function.

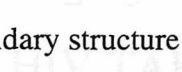
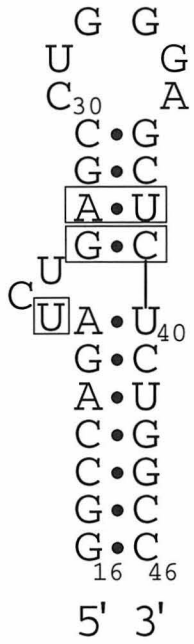
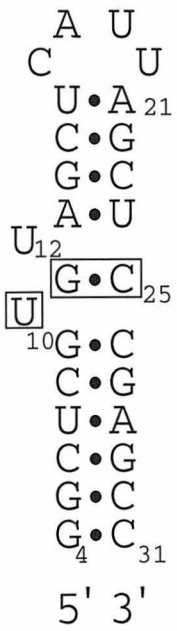


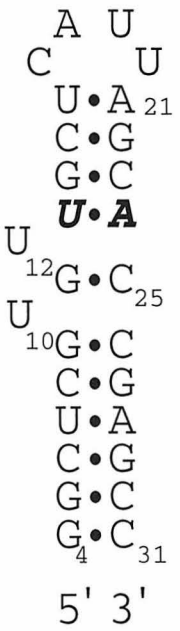
Figure 5.1. Sequence and secondary structure of HIV, BIV and BIV1 TAR RNA. The numbering of the BIV RNA corresponds to nucleotide positions in BIV mRNA. Nucleotides that are important for protein binding and function are boxed in the BIV TAR RNA sequence.



HIV TAR



BIV TAR



BIV1 TAR

5.2. Experimental

RNA preparation. The BIV and BIV1 TAR RNAs were prepared by *in vitro* transcription[8] using synthetic DNA templates and T7 RNA polymerase (Pharmacia). DNA primer and template strands were synthesized on an ABI DNA 392 synthesizer, purified by reverse phase HPLC, detritylated, and purified again by HPLC. The RNA products from the transcription reactions were precipitated and washed with EtOH. The RNA oligomers were gel purified on an 8 or 10% polyacrylamide denaturing gel, located with UV shadowing, excised and then eluted from the gel using an Elutrap device (Schleicher & Schuell). After precipitation, the RNA was stored frozen in 10 mM Tris HCl buffer pH 7.0 at -20°C. The RNA oligomers were quantitated using UV-vis spectroscopy. Extinction coefficients at 260 nm were 8850 $M^{-1}cm^{-1}$ /nucleotide based on nucleotide composition for BIV and BIV1 TAR RNA respectively. The oligomers were then 3'-end-labeled with cytidine 3',5'-[5'-^{32}P]-bisphosphate using T4 RNA ligase.[9] They were purified and recovered as described for the unlabeled RNA. The eluted RNA oligomers were ethanol-precipitated twice and stored frozen in 10 mM Tris-HCl, pH 7.5. $Rh(phen)_2phi^{3+}$ solutions were prepared fresh in 10 mM Tris-HCl pH 7.0.

Rhodium photocleavage. $[Rh(phen)_2phi]Cl_3$ was prepared as described earlier.[10] The enantiomers were resolved by cation-exchange chromatography on a Sephadex SPC-25/potassium antimonyl tartrate column.[11] Separated isomers of $[Rh(MGP)_2phi]Cl_5$ and $[Rh(GEB)_2phi]Cl_5$ were obtained from R. Terbrueggen. $Rh(phen)_2phi^{3+}$ stock solutions were freshly prepared in either ethanol or 10 mM Tris-HCl, pH 7.5, while the other metal complex stock solutions were prepared in 10 mM Tris-HCl, pH 5.5. Photocleavage samples were prepared in buffer (50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0). A typical 20 μL irradiation sample consisted of 50000 cpm renatured ^{32}P-labeled RNA, 0.001-10 μM rhodium complex, 50 or 100 μM nucleotides carrier TAR or $tRNA^{Phe}$ in aqueous buffer. Rhodium complex was added to the sample 5 - 20 minutes before irradiation at ambient temperature at 365 nm on a 1000-W Hg/Xe lamp and monochrometer

(Oriel model 77250) for 5-20 minutes. The samples were precipitated, washed and dried, and then eluted through a 20% denaturing polyacrylamide gel. The full-length RNA oligomers and cleavage products were identified by coelectrophoresing with $Ru(phen)_3^{2+}$ (G-specific) reactions[12] diethyl pyrocarbonate (DEPC) (A-specific) and hydrazine (U-specific) reactions.[13] The fragments produced were compared with the chemical sequencing lanes.[14]

Determination of Affinity Constants. Binding constants were obtained through quantitative affinity cleavage titrations according to published procedures.[15] Experimental conditions as described for TAR RNA in the previous chapter were used.

Peptide preparation. The 9-mer RKKRRQRRC and the 14-mer RPRGTRGKGRRIR peptides derived from the HIV Tat and BIV Tat peptides respectively were chemically synthesized and purified by the Biopolymer Synthesis and Analysis Resource Center at Caltech. The peptides were perpared as their COOH-termini amides, and the HIV Tat peptide was acetylated on the NH_2 terminus. The peptides were analyzed by capillary electrophoresis and mass spectrometry. The peptides were suspended in 10 mM Tris HCl pH 7.0 and stored frozen until use.

Competition experiments with Tat peptide. Rhodium competition experiments were carried under the same conditions for photocleavage, with the addition of either BIV or HIV Tat peptide to the photocleavage solutions. Experiments were carried out with either a fixed concentration of rhodium complex with the peptide concentrations varied, or with a fixed concentration of peptide and varying the rhodium concentration. In both cases, all components except for the rhodium complexes were allowed to equilibrate for 1 hour, before the addition of the rhodium complex and subsequent irradiation. All competition experiments were carried out at ambient temperature.

5.3. Results

5.3.1. Site-specific Photocleavage of BIV TAR RNA by $Rh(phen)_2phi^{3+}$

Figure 5.2 shows the $Rh(phen)_2phi^{3+}$-induced cleavage with photoactivation of the BIV TAR RNA. At micromolar concentrations of rhodium, primary cleavage is evident at U24, a base involved in the triple interaction in the presence of the BIV Tat peptide.[3] With increasing concentration of rhodium, some cleavage is also found over the controls at G9, neighboring the open bulge site. Still weaker cleavage is evident in the hairpin loop at U20 and C17 at high rhodium concentration (> 10 μM). It should be noted in these studies with synthetic RNAs that nuclease activity could not be eliminated. Nonetheless, data taken as a function of rhodium concentration make quite apparent the rhodium-induced sites of cleavage.

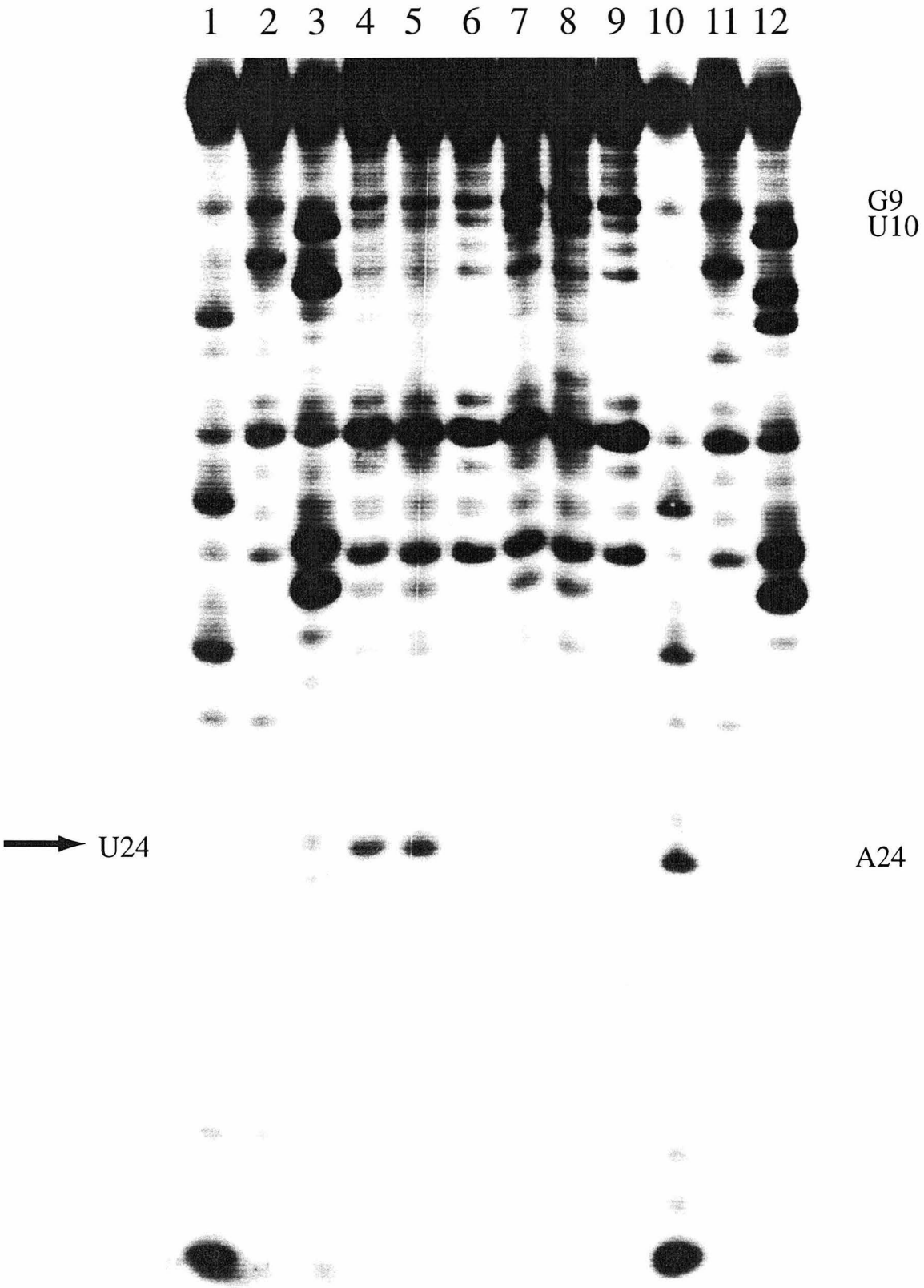
This specific targeting of the site of tertiary interaction in the RNA is enantiospecific. No similar site-specific cleavage is found with Λ-$Rh(phen)_2phi^{3+}$. It is interesting that the targeting of HIV TAR RNA by enantiomers of $Rh(phen)_2phi^{3+}$ do not show this strong degree of enantioselectivity in cleavage, as described in the previous chapter. Perhaps this reflects a tighter, more restricted site for the rhodium complex on BIV TAR RNA compared to HIV TAR RNA.

No difference in cleavage by these rhodium complexes was seen in the presence of argininamide (data not shown). This was similar to the finding in the previous chapter that argininamide has no effect on the cleavage sites on HIV TAR RNA.

5.3.2. Determination of the Binding Affinity of $Rh(phen)_2phi^{3+}$ for BIV RNA

Measurements of photocleavage as a function of absolute concentrations of rhodium/BIV TAR RNA allow the determination of the binding affinity of racemic $Rh(phen)_2phi^{3+}$ to its target site. Figure 5.3 shows this binding titration for $Rh(phen)_2phi^{3+}$ based upon quantitative photocleavage at U24. The data indicate an

Figure 5.2. Gray-scale representation of a phosphor autoradiogram of a 20% denaturing polyacrylamide gel showing Δ-Rh(phen)$_2$phi^{3+} cleavage on 3'-^{32}P-labeled BIV (lanes 1-6) and BIV1 RNA (lanes 7-12) in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. Lanes 1, 2 & 3: A-, G- and U-specific reactions on BIV, respectively. Lanes 4 & 5: labeled BIV after incubation with 1 and 5 μM Δ-Rh(phen)$_2$phi^{3+}, respectively, and irradiation for 15 minutes at 365 nm. Lane 6: labeled BIV upon irradiation for 15 minutes at 365 nm without metal complex. Lanes 7 & 8: labeled BIV1 after incubation with 1 and 5 μM Δ-Rh(phen)$_2$phi^{3+}, respectively, and irradiation for 15 minutes at 365 nm. Lane 9: labeled BIV1 upon irradiation for 15 minutes at 365 nm without metal complex. Lanes 10, 11 & 12: A-, G- and U-specific reactions on BIV1, respectively.


1 2 3 4 5 6 7 8 9 10 11 12
G9
U10
U24
A24

y = (m1*M0)/(1+m1*M0)		
	Value	Error
m1	2.1741e+06	3.2486e+05
Chisq	0.02591	NA
R	0.98988	NA

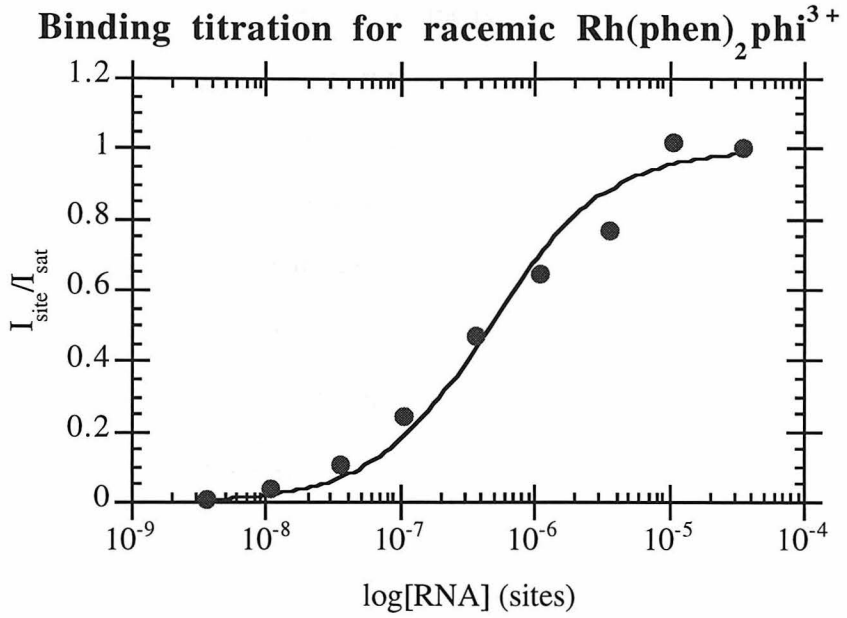


Figure 5.3. Quantitative affinity cleavage titration for $Rh(phen)_2phi^{3+}$ on BIV RNA. Plot of cleavage intensity, I_{site}, relative to the intensity at saturation, I_{sat}, as a function of RNA concentration. These data represent the site-specific cleavage signal intensity at U24 with a Rh/BIV TAR RNA ratio of 1:1. Experiments were conducted at ambient temperature in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. buffer, with irradiation at 365 nm.

affinity constant, $K_b = 2.2 \pm 0.3 \times 10^6\ M^{-1}$ for $Rh(phen)_2phi^{3+}$ to its target site on BIV TAR RNA. Affinity constants of Δ-$Rh(phen)_2phi^{3+}$ and 3Δ-$Rh(MGP)_2phi^{5+}$ were of similar magnitute.

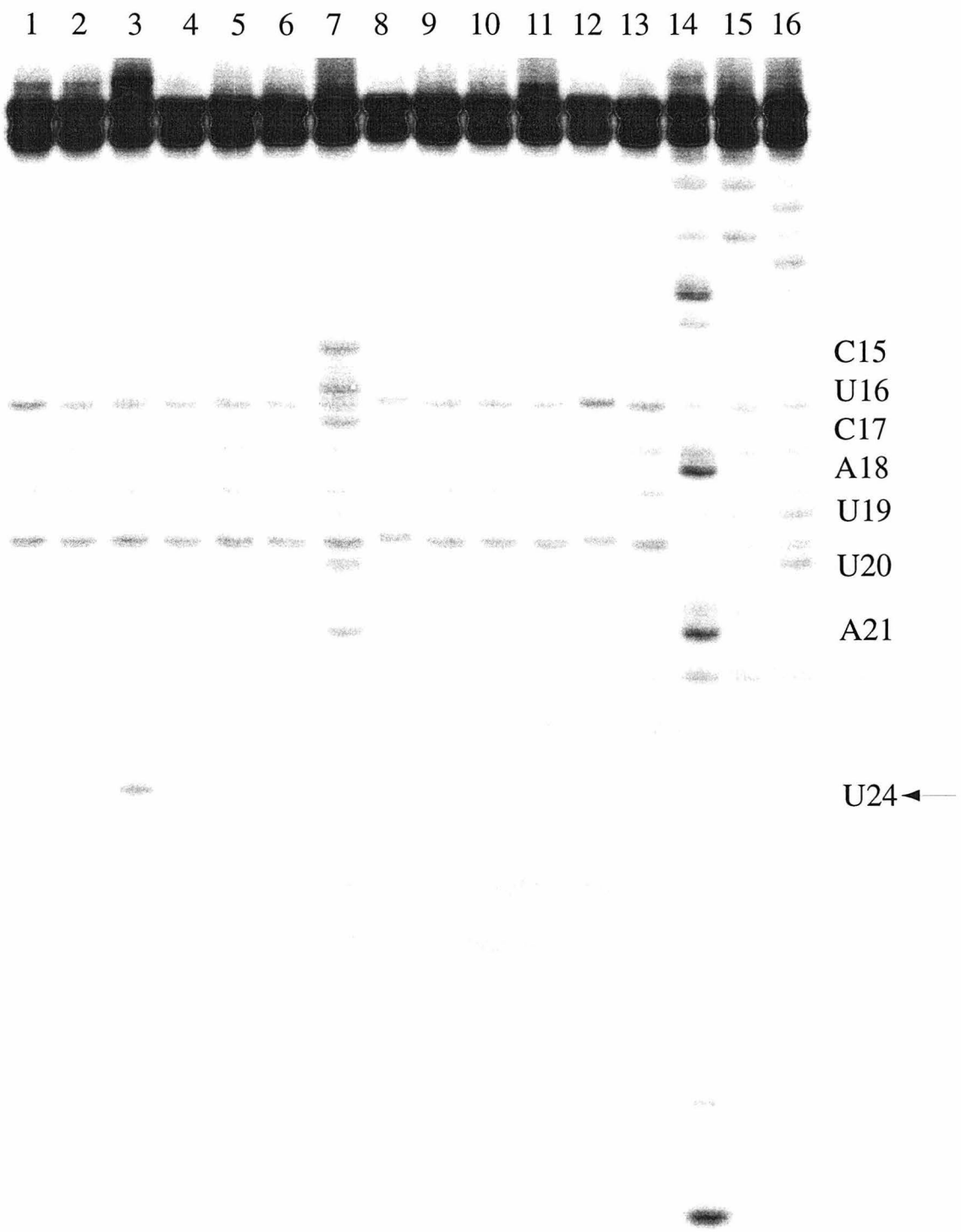
5.3.3. Photocleavage of BIV1 TAR RNA by $Rh(phen)_2phi^{3+}$

A direct demonstration that targeting by the rhodium complex of the site of tertiary interaction depends upon the formation of the base triple can be seen in comparing photocleavage on BIV TAR RNA to that on BIV1 TAR RNA, in which base triple formation is precluded. This mutation yields a 75% loss of binding affinity to the Tat peptide and a 93% loss of activation *in vivo.* [1] As evident in Figure 5.4 (lanes 7 and 8), incubation and irradiation of Δ-$Rh(phen)_2phi^{3+}$ with BIV1 yields no specific cleavage at A24. Interestingly, some cleavage is still evident on this RNA hairpin at G9 and to a lesser extent at U10. It is likely that the bulged RNA is sufficiently open to allow some access by the rhodium complex into a major groove site.

5.3.4. Photocleavage by Other Rhodium Complexes on BIV and BIV1 TAR RNA

Derivatives of $Rh(phen)_2phi^{3+}$ with guanidinium arms on the ancillary ligands, $Rh(MGP)_2phi^{5+}$ and $Rh(GEB)_2phi^{5+}$, as described and schematically illustrated in the previous chapter, were used to photocleave BIV RNA. Their recognition properties are similar but not identical to the parent $Rh(phen)_2phi^{3+}$ complex. Of these complexes, only the Δ-enantiomers cleaved the RNA (data not shown). The complex which has the strongest cleavage sites is 3Δ-$Rh(MGP)_2phi^{5+}$, but is less specific than the Δ-$Rh(phen)_2phi^{3+}$ complex (Figure 5.4). Cleavage is seen at U24, the main $Rh(phen)_2phi^{3+}$ recognition site, but is also seen at C15, U16, C17, U19, U20 and A21. Thus, this complex cleaves in the loop region as well as at the bulge. Cleavage is enantioselective, as

Figure 5.4. Gray-scale representation of a phosphor autoradiogram of a 20% denaturing polyacrylamide gel showing rac-Rh(phen)$_2$phi^{3+}, 3Δ-Rh(MGP)$_2$phi^{5+} and 3Λ-Rh(MGP)$_2$phi^{5+} cleavage on 3'-^{32}P-labeled BIV TAR RNA in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. Lanes 1 & 2: labeled BIV after incubation with 1 μM rac-Rh(phen)$_2$phi^{3+}, and irradiation for 10 and 20 minutes respectively at 365 nm. Lane 3: labeled BIV after incubation with 5 μM rac-Rh(phen)$_2$phi^{3+}, and irradiation for 10 minutes at 365 nm. Lane 4: labeled BIV after incubation with 5 μM rac-Rh(phen)$_2$phi^{3+}. Lanes 5-8: identical to lanes 1-4 respectively, but with 3Δ-Rh(MGP)$_2$phi^{5+}. Lanes 9-12: identical to lanes 1-4 respectively, but with 3Λ-Rh(MGP)$_2$phi^{5+}. Lane 13: labeled BIV after irradiation for 20 minutes at 365 nm. Lanes 14, 15 & 16: A-, G- and U-specific reactions on BIV, respectively.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
C15
U16
C17
A18
U19
U20
A21
U24

the 3Λ-enantiomer does not recognize the RNA molecule. A list of the cleavage sites by the various metal complexes on BIV TAR RNA is shown in Table 5.1.

The recognition properties of the $Rh(phen)_2phi^{3+}$ derivatives on BIV1 RNA were also similar to that of the parent complex. As shown in Figure 5.5, only 3Δ-$Rh(MGP)_2phi^{5+}$ and 3Δ-$Rh(GEB)_2phi^{5+}$ cleaved the RNA polymer, when a wide screen of complexes was assayed against this RNA. Weak cleavage by these complexes was observed at only one site, U10, one of the bulged nucleotides. This was similar to the $Rh(phen)_2phi^{3+}$ cleavage observed at this site. Cleavage was enantiospecific, as none of the Λ-enantiomers recognized BIV1. When cleavage on this mutant was compared to that on the wildtype BIV, it was striking that none of the complexes recognized A24, part of the basepair which had been mutated so that the RNA was no longer competent to form a base triple. Instead, weak cleavage was observed at only one site, U10, one of the bulged nucleotides, by the 3Δ enantiomers of both $Rh(MGP)_2phi^{5+}$ and $Rh(GEB)_2phi^{5+}$. This nucleotide was not cleaved by these complexes in the wildtype BIV, indicating that the BIV and BIV1 molecules are folded differently. A list of the cleavage site by the various metal complexes on BIV TAR RNA is shown in Table 5.2.

5.3.5. Competition Between $Rh(phen)_2phi^{3+}$ and BIV Tat-14

Arginine-rich 14-mer and 9-mer peptides derived from BIV Tat and HIV Tat respectively were chemically synthesized in order to carry out competition experiments between these Tat peptides and $Rh(phen)_2phi^{3+}$ for BIV TAR RNA. Figure 5.6 shows the results of photocleavage by $Rh(phen)_2phi^{3+}$ on BIV TAR RNA in the presence and absence of either HIV Tat peptide and BIV Tat peptide. As can be seen in Figure 5.6, despite its high arginine content, the HIV Tat peptide does not compete effectively with $Rh(phen)_2phi^{3+}$. Specific photocleavage of U24 is evident by the rhodium complex in the presence of micromolar concentrations of rhodium complex. This is despite the fact that the HIV Tat peptide has an affinity for BIV TAR RNA in the micromolar regime.[1]

Table 5.1. Rhodium complex cleavage sites on BIV TAR RNA.

Rhodium species	Cleavage sites[a]
Δ-Rh(phen)$_2$phi^{3+}	U24
Λ-Rh(phen)$_2$phi^{3+}	no cleavage
1Δ-Rh(MGP)$_2$phi^{5+}	U24
1Λ-Rh(MGP)$_2$phi^{5+}	no cleavage
racemic 2-Rh(MGP)$_2$phi^{5+}	U24
3Δ-Rh(MGP)$_2$phi^{5+}	C15, U16, C17, A18, U20, A21, U24
3Λ-Rh(MGP)$_2$phi^{5+}	no cleavage
1Δ-Rh(GEB)$_2$phi^{5+}	A21, U24
1Λ-Rh(GEB)$_2$phi^{5+}	no cleavage
2Δ-Rh(GEB)$_2$phi^{5+}	A21, U24, C25, C26
2Λ-Rh(GEB)$_2$phi^{5+}	no cleavage
3Δ-Rh(GEB)$_2$phi^{5+}	A21, U24, C25, C26
3Λ-Rh(GEB)$_2$phi^{5+}	no cleavage

[a] Strong sites are in bold, weak site are italicized.

Figure 5.5. Gray-scale representation of a phosphor autoradiogram of a 20% denaturing polyacrylamide gel showing photocleavage by a variety of rhodium complexes on 3'-^{32}P-labeled BIV1 TAR RNA in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. Lanes 1 - 8: labeled BIV1 after irradiation for 20 minutes at 365 nm in the presence of Δ-Rh(phen)$_2$phi^{3+}, rac-Rh(phen)$_2$phi^{3+}, racemic 2-Rh(MGP)$_2$phi^{5+}, 3Δ-Rh(MGP)$_2$phi^{5+}, 3Λ-Rh(MGP)$_2$phi^{5+}, racemic 3-Rh(MGP)$_2$phi^{5+}, 3Δ-Rh(GEB)$_2$phi^{5+} and 3Λ-Rh(GEB)$_2$phi^{5+} respectively. Lane 9: labeled BIV1 after irradiation for 20 minutes at 365 nm. Lanes 10, 11 & 12: A-, G- and U-specific reactions on BIV, respectively.

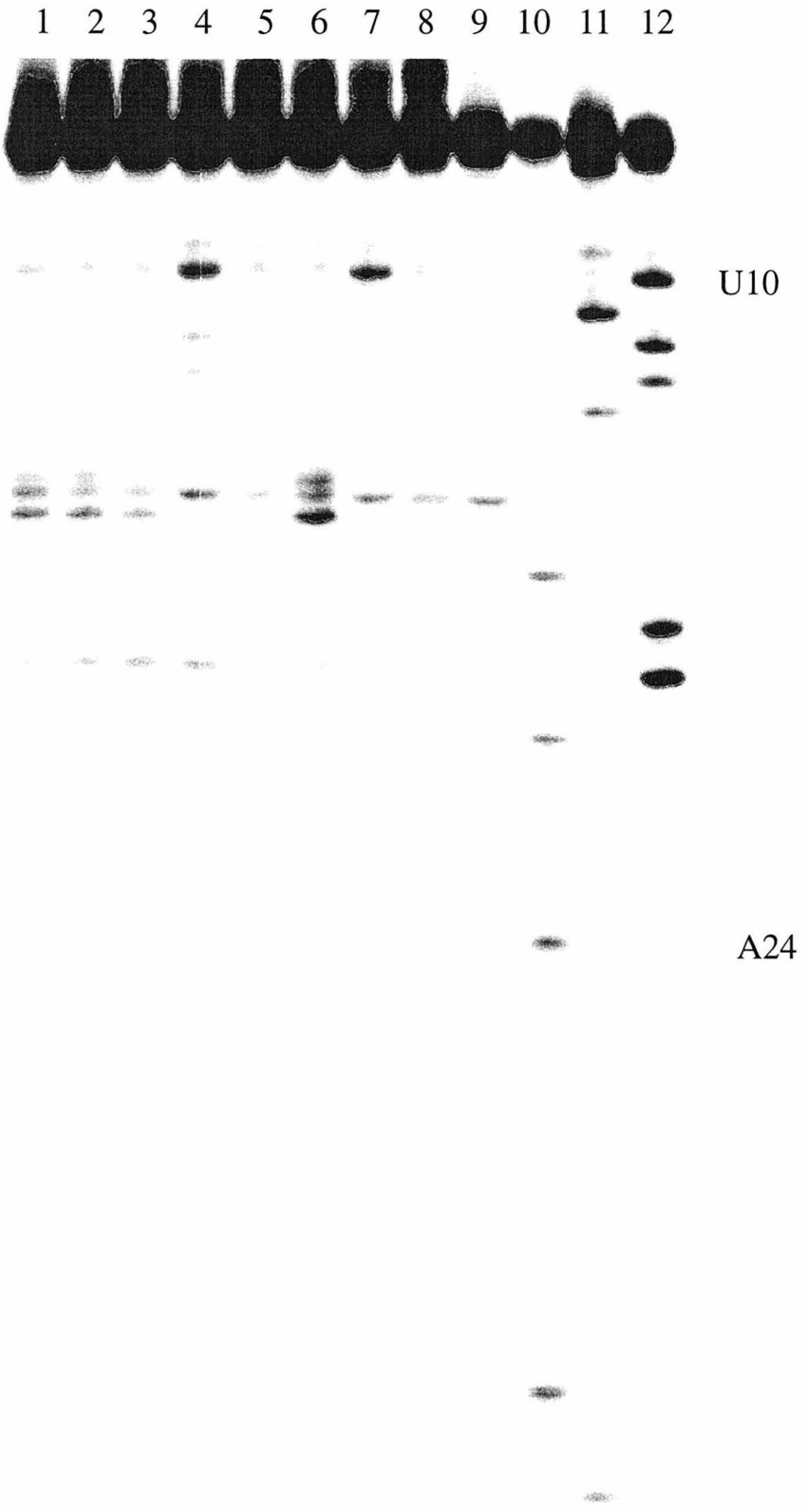

1 2 3 4 5 6 7 8 9 10 11 12
U10
A24

Table 5.2. Rhodium complex cleavage sites on BIV1 TAR RNA.

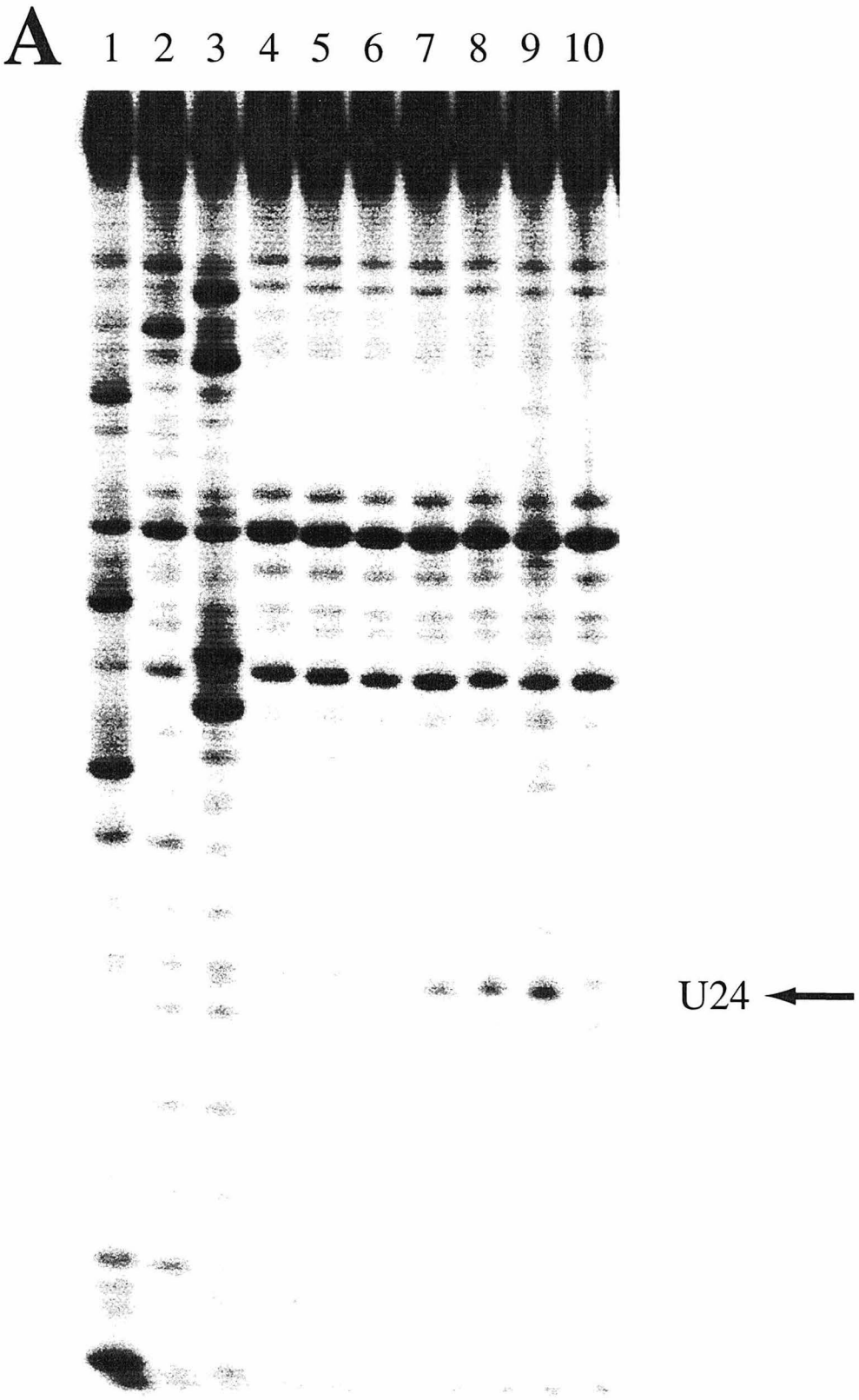
Rhodium species	Cleavage sites[a]
Δ-Rh(phen)$_2$phi^{3+}	G9, G10
Λ-Rh(phen)$_2$phi^{3+}	no cleavage
1Δ-Rh(MGP)$_2$phi^{5+}	U10, U20
racemic 2-Rh(MGP)$_2$phi^{5+}	no cleavage
3Δ-Rh(MGP)$_2$phi^{5+}	U10, U19, U20
3Λ-Rh(MGP)$_2$phi^{5+}	no cleavage
3Δ-Rh(GEB)$_2$phi^{5+}	U10
3Λ-Rh(GEB)$_2$phi^{5+}	no cleavage

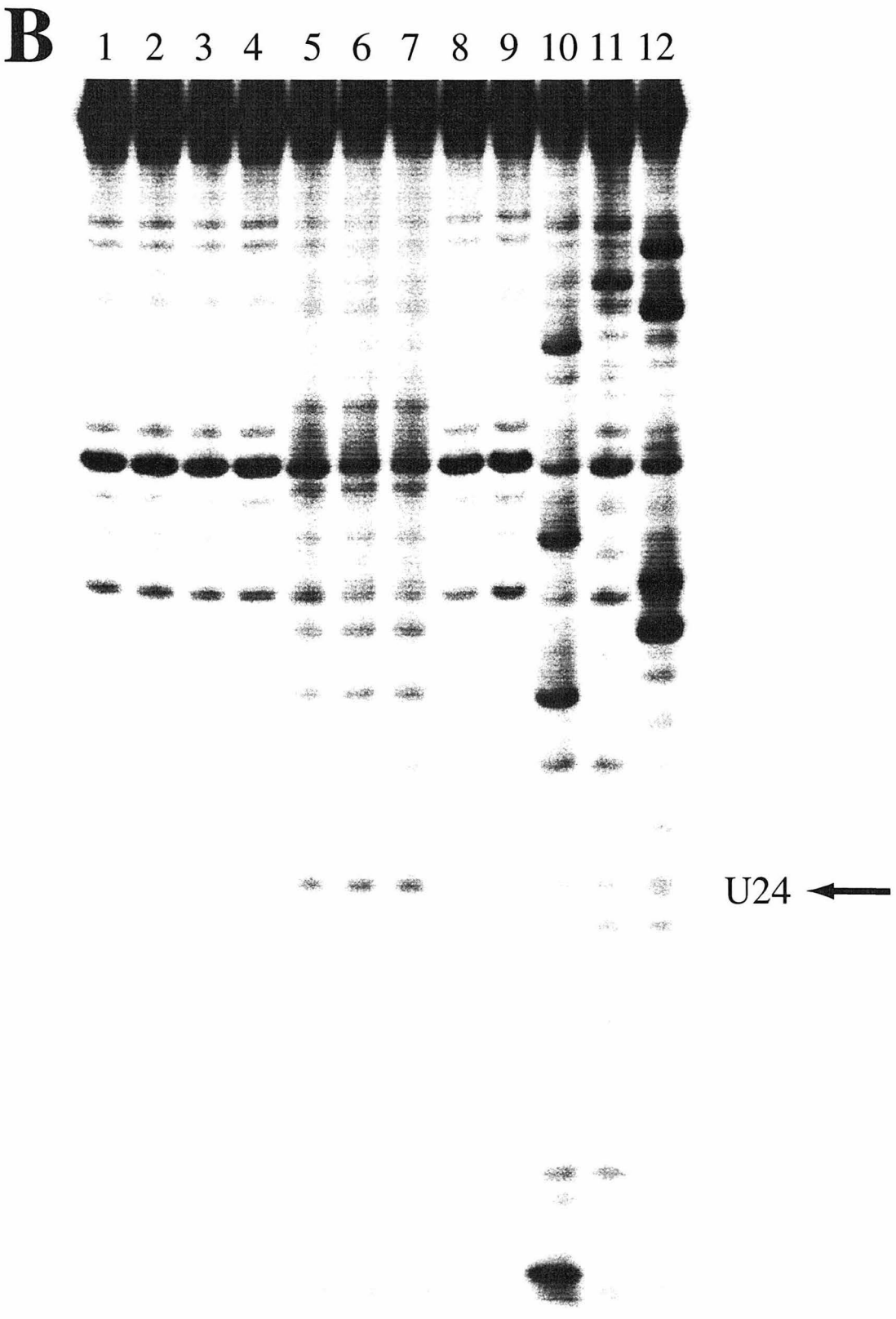
[a] Strong sites are in bold, weak site are italicized.

Figure 5.6. Competition for BIV TAR RNA between $Rh(phen)_2phi^{3+}$ and HIV (**A**) Tat peptide or BIV (**B**) Tat peptide.

A. Gray-scale representation of a phosphor autoradiogram of a 20% denaturing polyacrylamide gel showing $Rh(phen)_2phi^{3+}$ cleavage on 3'-^{32}P-labeled BIV RNA in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. Lanes 1, 2 & 3: A-, G- and U-specific reactions on BIV TAR RNA respectively. Lanes 4, 5, 6, 7, 8 & 9: labeled BIV after incubation with 2 μM HIV Tat and 0.1, 0.5, 1, 2, 5 & 10 μM $Rh(phen)_2phi^{3+}$, respectively, and irradiation for 15 minutes at 365 nm. Lane 10: labeled BIV with 2 μM HIV Tat and 5 μM $Rh(phen)_2phi^{3+}$.

B. Gray-scale representation of a phosphor autoradiogram of a 20% denaturing polyacrylamide gel showing $Rh(phen)_2phi^{3+}$ cleavage on 3'-^{32}P-labeled BIV in 50 mM Tris-HCl, 18 mM NaCl, 10 mM NaOAc, pH 7.0. Lanes 1, 2, 3, 4, 5, 6 & 7: labeled BIV after incubation with 2 μM BIV Tat and 1, 2, 5, 10, 25, 50 & 100 μM $Rh(phen)_2phi^{3+}$, respectively, and irradiation for 15 minutes at 365 nm. Lane 8: labeled BIV with 2 μM BIV Tat and 5 μM $Rh(phen)_2phi^{3+}$. Lane 9: labeled BIV with 2 μM BIV Tat upon irradiation for 15 minutes at 365 nm without metal. Lanes 10, 11 & 12: A-, G- and U-specific reactions on BIV respectively.


A
1 2 3 4 5 6 7 8 9 10
U24


B
1 2 3 4 5 6 7 8 9 10 11 12
U24

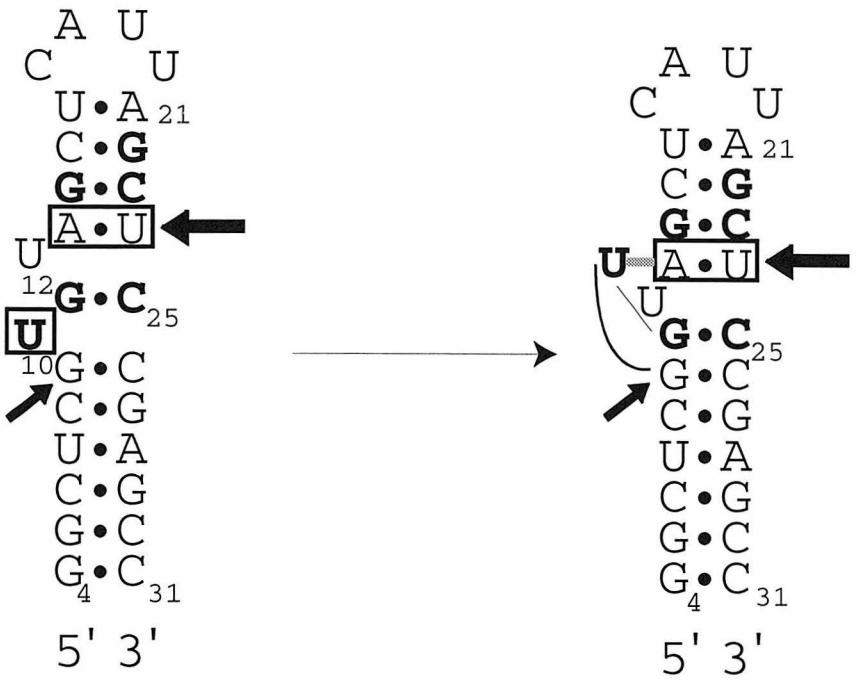
Tighter specific binding of BIV Tat peptide is evident to its target BIV RNA than is the HIV counterpart.[1] Affinities of the BIV Tat peptide for BIV TAR RNA have been previously reported to be 2 x 10^8 M^{-1}.[1] At stoichiometric rhodium and BIV Tat peptide concentrations, no specific cleavage by the rhodium complex is therefore evident. Given its lower affinity for the RNA site, higher concentrations of rhodium complex are required instead for effective competition. Instead, as also demonstrated in Figure 5.6, competitive titrations reveal extensive photocleavage at U24 by the rhodium complex at 20 μM concentration. Therefore, the rhodium complex does effectively compete with binding of the Tat peptide to its target.

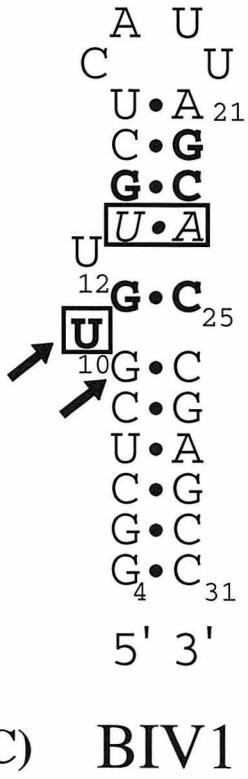
5.4. Discussion.

The sites targeted by $Rh(phen)_2phi^{3+}$ on BIV TAR RNA show a correspondence to those which are important in peptide binding and function. This is represented schematically in Figure 5.7. The strongest site is U24, a residue shown to be involved in triple base formation in the presence of BIV Tat peptide. Here we see that specific binding of the metal complex also depends upon base triple formation. Cleavage at U24 by the rhodium complex is abolished when the A13-U24 base pair is mutated to U13-A24 in the BIV1 RNA. This conservative mutation necessarily disrupts the base triple. This sequence is no longer recognized by the rhodium complex, probably because the major groove is not widened by the triple any longer. Thus, the metal complex is sensitive to structural perturbations associated with this simple change in sequence. Remarkably, the sensitivity of metal complex recognition to the base triple even exceeds that of the Tat peptide, where the U24A-A13U mutation produces 25% binding relative to wild type.[1]

Based upon the photocleavage data for the complex and the differential recognition of BIV TAR RNA versus BIV1, we would propose that the metal complex intercalates at the triple base site. Given that strand scission involves direct reaction of the activated phi

Figure 5.7. (**A**) Schematic of the BIV TAR RNA showing the correspondence between nucleotides targeted by $Rh(phen)_2phi^{3+}$ cleavage (arrows with size corresponding to relative intensity) and those important for protein binding (bold) and tertiary folding (boxed). (**B**) Folded BIV TAR RNA. (**C**) BIV1 TAR RNA showing sites of $Rh(phen)_2phi^{3+}$ cleavage.


A U
C U
U•A 21
C•G
G•C
A•U
U
12G•C 25
U
10G•C
C•G
U•A
C•G
G•C
G4•C 31
5' 3'
(A)
BIVWT
(B)


A U
C U
U•A 21
C•G
G•C
U•A
U
12G•C 25
U
10G•C
C•G
U•A
C•G
G•C
G4•C 31
5' 3'
(C) BIV1

ligand and the RNA, the phi ligand is likely stacked between the U10-A13•U24 triple and the G11•C25 base pair, in the widened major groove. Although the bulged site lacking the triple could provide some opening for the metal complex, given weak cleavage at G9 on BIV1, stacking against the triple base array provides critical stabilization of the intercalator in the fully folded site of BIV TAR RNA. A comparison of the cleavage characteristics of the rhodium complex on the folded WT BIV TAR RNA and the BIV1 RNA is also shown schematically in Figure 5.7. The site is also likely a somewhat restricted one given the high enantioselectivity we observe. In the NMR structure of the TAR RNA bound by Tat peptide,[3] the width of the major groove at the G11•C25 base pair is thought to be 17Å. These helical dimensions would likely allow access of the right-handed Δ-$Rh(phen)_2phi^{3+}$ for intercalation but not the Λ-isomer.[16] We cannot establish from these data whether the metal complex induces the conformational change of the RNA upon binding or whether, in the absence of metal complex as well as without peptide, the RNA adopts a folded form with triple base interaction. It is clear, however, that the same folded conformation required for peptide binding is required for recognition by the metal complex.

Perhaps most remarkable is the simple observation of high site-selectivity for the octahedral coordination complex, which lacks the array of functionalities present in the peptide and proposed to be critical for the peptide in achieving recognition of its nucleic acid site. As described in this and the previous chapter, experiments in our laboratory have shown that the addition of guanidinium moieties to the ancillary ligands of the rhodium complex,[11] mimicking the arginine side chains on the native peptide, may enhance to the same extent both the affinity and selectivity of the metal complex for HIV and BIV RNA sequences. However, as shown here, even without these functionalities, *shape-selection*, matching the shape of the small metal complex to its nucleic acid target, provides sufficient selective stabilization for RNA site discrimination. Indeed, we find that $Rh(phen)_2phi^{3+}$ competes effectively with the Tat peptide for its binding site on the TAR RNA. Our results therefore underscore once again the value of shape-selective recognition. Shape-selection

can be powerfully applied not only in probing structural variations along the nucleic acid polymer but also in the design of novel small molecules to target nucleic acid sites with high site-selectivity, in the development of molecules to inhibit protein recognition, and, potentially, in the design of new chemotherapeutics.

References

1. Chen, L.; Frankel, A. D. *Biochemistry* **1994**, *33*, 2708-2715.
2. Puglisi, J. D.; Chen, L.; Blanchard, S.; Frankel, A. D. *Science* **1995**, *270*, 1200-1203.
3. Ye, X.; Kumar, R. A.; Patel, D. J. *Chemistry & Biology* **1995**, *2*, 827-840.
4. Puglisi, J. D.; Tan, R.; J., C. B.; Frankel, A. D.; Williamson, J. R. *Science* **1992**, *257*, 76-80.
5. Johann, T. W.; Barton, J. K. *Phil. Trans. Royal Soc. A.* **1996**, *1996*, 299-324.
6. Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972-982.
7. Neenhold, H. R.; Rana, T. M. *Biochemistry* **1995**, *34*, 6303-6309.
8. Miligan, J. F.; Groebe, D. R.; Witherall, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783-8798.
9. England, T. E.; Uhlenbeck, O. C. *Nature* **1978**, *275*, 560-561.
10. Chow, C. S.; Barton, J. K. *Meth. Enzym.* **1992**, *212*, 219-242.
11. Terbrueggen, R. H.; Barton, J. K. *Biochemistry* **1995**, *34*, 8227-8234.
12. Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839-2841.
13. Peattie, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1760.
14. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303-2312.
15. Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995-11003.

16. Barton, J. K. *Science* **1986**, *233*, 727-734.