RHODIUM INTERCALATORS AS NOVEL PEPTIDE DELIVERY SYSTEMS TO THE MAJOR GROOVE OF DNA: TOWARDS THE DESIGN OF ARTIFICIAL REPRESSORS

Thesis by

Niranjan Y. Sardesai

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1995

(Submitted February 13, 1995)

ACKNOWLEDGEMENTS

At the very outset I would like to express my sincerest regards for my advisor, Professor Jacqueline K. Barton, for all her encouragement, support and advise throughout my graduate career. She has been an ideal mentor and has influenced me in many positive ways. Jackie taught me to enjoy doing science by maintaining a positive attitude and trust my intuitions and perceptions. I feel privileged to have worked with Jackie and I hope I can take away some of her optimism and love for science with me.

I thank my thesis committee comprising of Prof. John Bercaw, Prof. Doug Rees and Prof. John Richards for their helpful comments and suggestions. I would like to acknowledge two post-doctoral research colleagues, Dr. Takashi Morii and Dr. Kaspar Zimmermann, very warmly. Takashi taught me the ropes as a first year graduate student and I thank him for setting me off on a good path. Kaspar played a steadying role in my graduate life both in and out of the laboratory as a colleague and a friend. The time we spent working together on adjacent lab benches was very productive. Kaspar played a key role in the development of a viable synthetic strategy for the metal-peptide complexes and first discovered the "glutamate switch" effect described in Chapter 3. I will also cherish my friendship with Vreni and Kaspar very much.

I sincerely thank Tara Chapman, who worked with me over the summer and fall of 1994, for her assistance and fondly wish her all the best in her graduate research. I am grateful to Jane Sanders, John Racs and Suzanna Horvath of the Biopolymer Synthesis and Analysis Resource Center at Caltech for their technical assistance with the mass spectrometric and amino acid analytical characterization of the metal-peptide complexes.

I thank all the members of the Barton group, past and present, for being a great set of colleagues. I enjoyed working with them immensely. In particular, I would like to thank Michelle Arkin, Dr. Richard Hartshorn, R. Erik Holmlin, Dr. Louis Kuo, Susanne Lin, Dr. Michael Pustilnik, Tom Shields and Dr. Ayesha Sitlani for their many scientific

discussions and insights. On a more personal level, I would like to show my appreciation for the enduring friendship I enjoyed with Ayesha, Dr. Inho Lee, Michelle and Dr. Christine Chow. I thank Maureen (Mo) Renta for all her help over the years. I would also like to acknowledge the many fun times I had with the chemistry soccer team and the "Hogs" softball team. On the whole, I leave with fond memories of Caltech and I thank everyone responsible for them.

Gargi Maheshwari deserves a lot of thanks and my deep appreciation for all her love and for bringing a sense of balance into my life particularly during the trying thesis writing days. Finally I thank my parents and my brother, Anand, for their constant support and faith in me. I draw strength and inspiration from my parents who raised me to be independent and free thinking. My mother showed me how to set goals and go about achieving them. My father taught me patience and understanding- two qualities that I suspect are a great asset in any profession.

ABSTRACT

Phenanthrenequinone diimine (phi) complexes of rhodium(III) bearing tethered peptides have been designed to serve as metallointercalating anchors to deliver peptide side chain functionalities for DNA recognition in the major groove. Metal-peptide complexes containing 11-15 amino acid residues were prepared using two complementary synthetic strategies: by direct coupling of a pendant carboxylate on the coordinatively saturated rhodium complex, $[Rh(phi)_2(phen')]^{3+}$ (phen'=5-amidoglutaryl-1, 10-phenanthroline), to the N-terminus of a resin-bound peptide in a manner analogous to the chain-elongation step in solid phase peptide synthesis; or by coupling phen' containing the pendant carboxylate to the resin-bound peptide, followed by coordination of $[Rh(phi)_2]^{3+}$ to the bidentate chelator attached to the peptide. With coordination complexes which are stable to peptide deprotection and cleavage conditions from the resin, the solid phase synthetic strategies prove convenient to apply. The metal-peptide complexes have been characterized by amino acid analysis, electronic spectroscopy, circular dichroism and mass spectrometry, where a novel pattern of peptide fragmentation facilitates the detailed sequence analysis of the appended peptide. All the metal-peptide complexes bind and, with photoactivation, cleave DNA with evidence of major groove chemistry. Significantly, the DNA site-specificity is seen to depend on the peptide side-chain functional groups. In one series, a single glutamate at position 10 is found to be essential in directing DNA site-recognition to the sequence 5'-CCA-3'. Methylation of the glutamate side chain or single E10Q, E10D, E10A mutations abolish this selectivity. The glutamate is essential to maintain α-helicity in the peptide and make base specific contacts, thereby providing a glutamate switch for site-specific DNA recognition. A second series, based on the recognition helix of the phage 434 repressor, reproduces operator binding. Photocleavage and MPE-Fe footprint analysis indicates that these metal-peptide complexes bind to the 5'-ACAA-3' operator sequences as monomers at

10 nM concentration and differentiate between operator site variants. These studies represent a new strategy to create an array of metal-peptide complexes with differing sequence specificity for DNA and suggest a route to the construction of small molecules that function as artificial repressors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

Chapter 1:
Introduction

1.1. Features of Protein-DNA Interactions

DNA-binding proteins play crucial roles in replication, transcription and repair of damaged DNA and as such there is a great deal of interest in determining the molecular parameters that govern the recognition of DNA by proteins. This basic knowledge is essential for the rational design of pharmaceuticals directed at DNA and to develop new tools to explore the genome. One of the most diverse classes of DNA-binding proteins are the transcription factors that regulate gene expression. Much of our understanding of the types of interactions that occur when proteins bind to specific DNA sites comes from studying structures of transcription factors bound to DNA. Many families of transcription factors have been identified, members of which share related structural motifs for recognition. These families include helix-turn-helix proteins, homeodomains, leucine zipper proteins, helix-loop-helix proteins, zinc finger proteins, steroid receptors and the β-sheet motif proteins; and they have been reviewed extensively.[1-9]

From the study of protein-DNA complexes, it is apparant that proteins use a combination of interactions to achieve specificity. Specific sites may be recognized by direct amino acid side chain-nucleic acid base contacts and by shape selection or "indirect readout." Site discrimination by shape selection is achieved mainly by steric constraints and the protein takes advantage of sequence dependent conformational differences in DNA to specifically bind to its binding site.[10] Frequently, protein recognition is more direct through a combination of base specific hydrogen bonds and van der Waals interactions. While the hydrogen bonds may be from side chain functionalities or from the peptide back bone amides, most base specific hydrogen bonds are made by the protein side chains. These contacts occur predominantly from the major groove, perhaps because here each base pair presents a different combination of functional groups for maximum

discrimination and the major groove is wide enough to accommodate elements of protein secondary structure for protein-DNA contacts. Even with a "direct readout," the shape complementarity of proteins with the DNA binding site is striking. Proteins may recognize 12-14 bp dimeric binding sites, but may protect over 20 bp of DNA from nucleases. Indeed, the DNA may bend or wrap around the protein in some instances to maximize the contacts.[11a] This extensive protein-DNA contact surface provides much of the non-specific binding affinity through interactions with the phosphate back bone. Usually proteins bind DNA with multiple DNA-binding domains to increase the specific affinity for the binding site. These domains may be identical but from different molecules, as in homodimers or heterodimers, or they may be placed in tandem on the same polypeptide chain.

While some broad generalizations can be made for the overall binding picture, the molecular details of the interaction are specific to the individual protein being studied. It appears that there is no universal recognition code for amino acid residue-nucleic acid base hydrogen bonds or van der Waals interactions. The same residue may recognize a number of DNA bases and, conversely, each DNA base may be recognized by a number of amino acid side chains. Indeed, as Figure 1.1 shows, several of the many possible hydrogen bonding interactions that may be envisioned between amino acid side chains and DNA bases are seen in nature. The particular set of interactions selected is highly context dependent. Similarly, a number of residues may recognize the 5-methyl group on thymine through a van der Waals contact. Hydrophobic contacts with Ala, Val, Leu, Ser, Thr and Gln are but a few structurally characterized examples.[11(d, h, j)] Another variable is the number and relative distribution of the two types of contacts. Frequently, the number of base-specific hydrogen bonds outnumber the van der Waals contacts. However, recent examples of the GATA-1 transcription factor and RNA-protein interactions in the HIV TAR-arginine or the glutamyl tRNA synthetase-$tRNA^{Glu}$

Table 1.1. Summary of Hydrogen Bonding Interactions Between Side Chains and Bases in the DNA Major Groove for Representative Protein-DNA Complexes.

	A	T	C	G
Arg		CAP		GR Trp Zif CAP Arc Gli
Asn	en GATA-1	Arc	Arc	lambda
Asp			Gli	
Gln	lambda 434 Arc	434		434
Glu			CAP	
His				Zif
Lys				GR lambda MetJ Gli
Ser	Gli			lambda Gli
Thr	MetJ			

The table has been adapted from reference 1 with additional data from the Arc, Gli and GATA-1 structures. Contacts are summarized for: *E. coli* CAP protein (CAP); *Drosophila* engrailed homeodomain (en); λ-repressor (lambda); 434 repressor (434); glucocorticoid repressor (GR); *E. coli* MetJ repressor (MetJ); *E. coli* Trp repressor (Trp); Zif268 zinc fingers (Zif); P22 Arc repressor (Arc); zinc fingers from the human *GLI* oncogene (Gli); erythroid specific transcription factor GATA-1 (GATA-1). The sources are cited in references 10, 11.

structure suggest that hydrophobic interactions may outnumber or outweigh hydrogen bonds in some instances.[11j, 12]

1.2. Designing Proteins with Altered Recognition or Function

Perhaps because there are several solutions to the recognition problem, making *apriori* predictions and designing specific DNA binding molecules *de novo* is a difficult task. However, numerous examples abound in the literature where natural systems have been modified in some way to alter the recognition or the function of the protein. These experiments take advantage of the existing protein scaffold to deliver specific functionalities to DNA. The zinc finger proteins have proved particularly convenient to manipulate and discover rudimentary rules for zinc finger-DNA recognition.[13] Phage selection techniques have been applied with success in the Zif268 zinc finger protein to discover mutant proteins with altered DNA specificity.[14] Naturally occurring proteins have been coupled with metal chelating groups to harness the reactivity of the metal ions. Dervan and co-workers have employed the affinity cleaving technique to map out the binding sites and binding orientations of proteins.[15] Sigman and co-workers have converted several DNA-binding proteins into site-specific redox activated nucleases by attachment of 1, 10-phenanthroline derivatives on to the proteins followed by addition of copper and reductants for oxidative DNA cleavage.[16] Sugiura and co-workers similarly converted the zinc finger transcription factor Sp1 into a DNA cleaving protein in the presence of Ni(II) and a peracid.[17]

Recently, Pabo and co-workers created a protein containing a fusion of fingers 1 and 2 from Zif268 and the homeodomain from Oct-1.[18] The fusion protein bound optimally to a sequence containing adjacent homeodomain and zinc finger subsites. The structure-based design of such hybrid transcription factors or repressors holds promise for the targeted regulation of specific cellular genes.

All the systems described above are macromolecular in design. The large size makes their manipulation difficult and understanding their specific interactions with DNA even harder. Attempts at minimalist design of DNA binding proteins likewise have their limitations. Simon and co-workers discovered that a 31-residue fragment from Hin recombinase containing the helix-turn-helix motif was not sufficient to impart sequence specific binding; instead a 52 residue piece was required.[19] Similar attempts have been made to isolate minimum DNA binding peptides in the leucine zipper proteins. The basic region α-helices of GCN4 have been linked together through a variety of chemical linkers doing away with the dimerizing leucine zipper unit. The minimum size of the covalent dimer, however, still utilized >40 residues (>20 amino acids per monomer) for site specific recognition.[20] The greatest handicap in designing shorter peptide based DNA binding molecules or artificial repressors is that the smaller peptides do not have sufficient binding affinity for specific binding to DNA.

1.3. Transition Metal Complex Directed Recognition of DNA

The Barton laboratory has been interested in developing octahedral metal complexes for DNA recognition.[21] The complexes of ruthenium, cobalt and rhodium developed in the group are all in a low spin d^6 electronic configuration and are substitutionally inert. Hence DNA recognition, as by proteins, is achieved through a set of non-covalent interactions rather than direct coordination. The complexes are small, adopt well defined three dimensional structures and are easily accessible synthetically in the laboratory. The rigid three dimensional structure implies that recognition elements may be built on a well defined scaffold and the ease of synthesis facilitates the isolation of the designed complexes. The use of transition metals also provides spectroscopic and photoreactive handles to study binding events. Summarized below are examples of ruthenium and rhodium complexes that mimic protein-nucleic acid interactions by taking

advantage of shape selection or hydrogen bonding and van der Waals interactions to discriminate between DNA sites.

(a) *Importance of Shape Selection:* The three dimensional structures of and the chirality associated with the metal complexes have been utilized to recognize global features of nucleic acids through shape selection. Perhaps the simplest form of recognition is illustrated by studies on $[Ru(phen)_3]^{2+}$ binding to DNA. The complementarity in the structures of B-form DNA and $[Ru(phen)_3]^{2+}$ discriminates between the two enantiomers of the complex such that surface binding in the minor groove is favoured by the Λ enantiomer, and intercalative binding in the major groove by the Δ enantiomer (Figure 1.1).[22] In another example, the construction of $[Ru(TMP)_3]^{2+}$ (TMP=3,4,7,8-tetramethyl-1,10-phenanthroline) led to the preferential targeting of A-form helices. The bulky and sterically hindered shape of $[Ru(TMP)_3]^{2+}$ prevents it from binding in the deep and narrow minor groove of B-form DNA, but allows binding through hydrophobic association with the wide and shallow minor groove of A-form helices (Figure 1.2).[23] Intercalative interactions with this complex are not feasible since the methyl groups preclude π-stacking.

The sensitivity of shape selection is illustrated by differences in DNA recognition between two related complexes, $[Rh(phen)_2(phi)]^{3+}$ and $[Rh(phi)_2(bpy)]^{3+}$.[24] Both complexes contain the phenanthrenequinone diimine (phi) ligand, an avid intercalator in the DNA base stack, which allows for the complexes to bind intimately and sense the local DNA structure at the intercalation site. Both complexes cleave DNA upon photoactivation, but the DNA sites recognized by the two complexes are markedly different. As shown in Figure 1.4, the 2, 9-H on the ancillary phen ligand in $[Rh(phen)_2(phi)]^{3+}$ clash with the DNA bases at most sites. Instead the complex binds at DNA sites that are open in the major groove to alleviate these steric clashes. Similar constraints are minimal for $[Rh(phi)_2(bpy)]^{3+}$ because the steric bulk associated with the ancillary ligands is pulled away from the axial direction (Figure 1.3). Indeed,

Figure 1.1. Enantioselective binding modes for Λ- and Δ- $[Ru(phen)_3]^{2+}$ to B form DNA. (Top) surface binding from the minor groove; (bottom) intercalative binding in the major groove. Symmetry considerations and steric factors favour intercalative binding by the Δ-enantiomer and groove binding by the Λ-enantiomer.

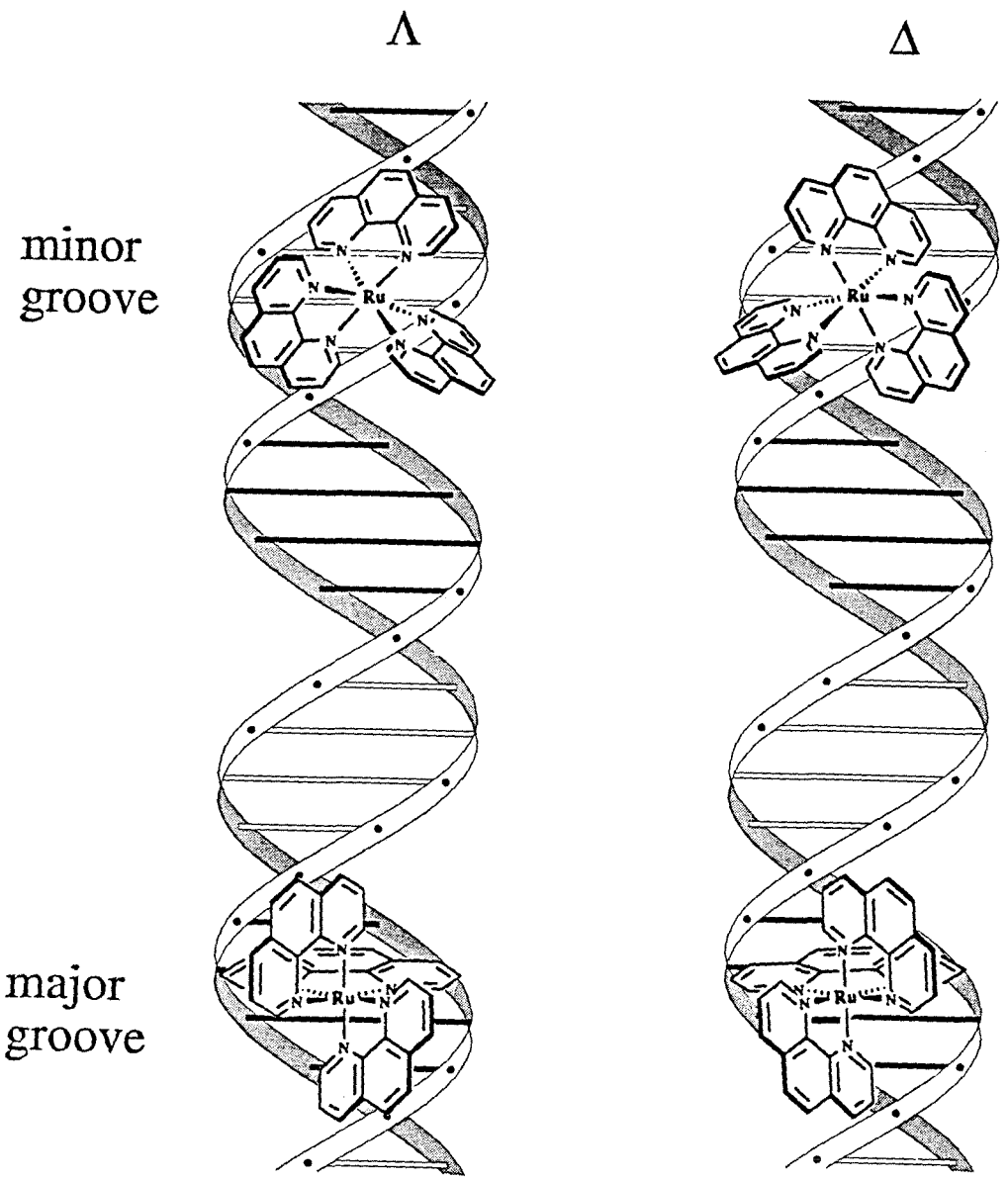
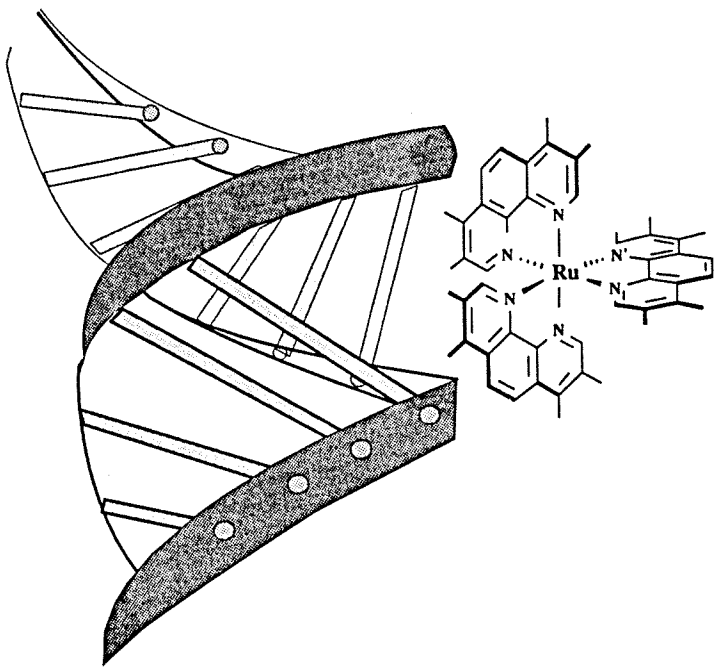

Λ
Δ
minor groove
major groove
Ru
N

Figure 1.2. Schematic representation of $[Ru(TMP)_3]^{2+}$ as a shape selective probe for A-form DNA. The surface of the minor groove of A-form DNA (wide and shallow) complements the shape of the tetramethylphenanthroline complex (top). The methyl groups preclude binding of the complex against the narrow minor groove of B-form DNA (bottom).


Ru
N
N'
N1

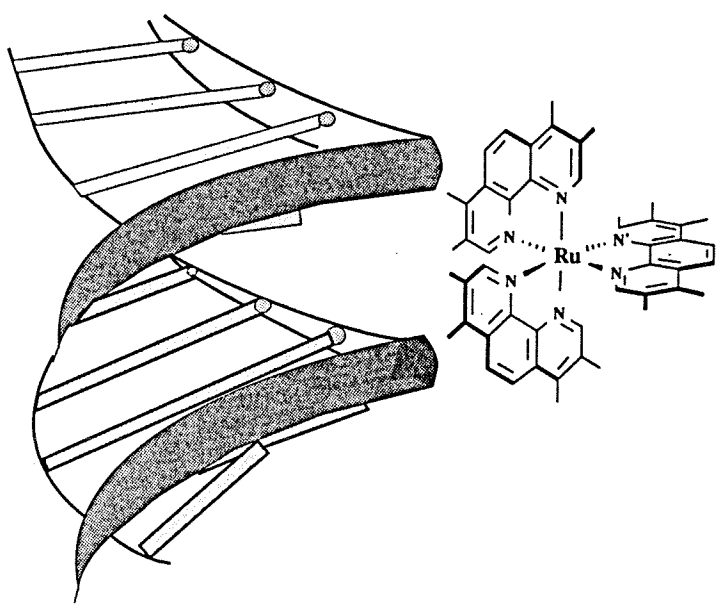
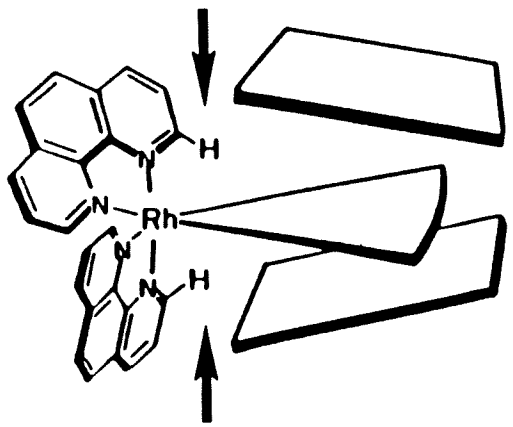
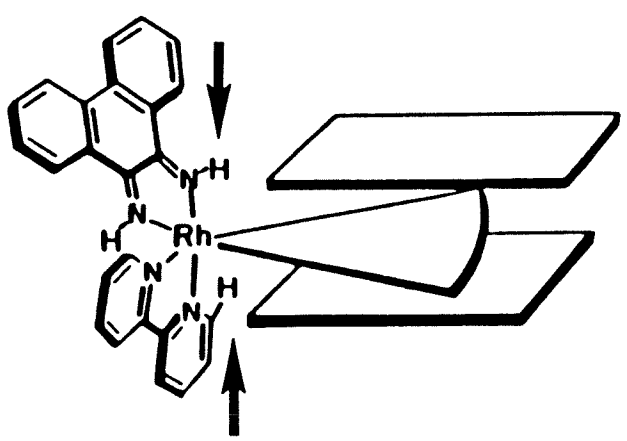

Ru
N
N'
N1

Figure 1.3. Shape-selective recognition of local DNA structure by $[Rh(phen)_2(phi)]^{3+}$ and $[Rh(phi)_2(bpy)]^{3+}$. (Top) Schematic representation of $[Rh(phen)_2(phi)]^{3+}$ intercalated into a site with an "open" major groove to avoid steric clashes with the 2, 9-H (arrows) on the ancillary phenanthroline ligands. (Bottom) Schematic representation of $[Rh(phi)_2(bpy)]^{3+}$ showing the analogous H atoms (arrows) pulled back in the axial direction. The reduced steric constraints allow the complex to intercalate at canonical DNA sites.



$[Rh(phen)_2(phi)]^{3+}$



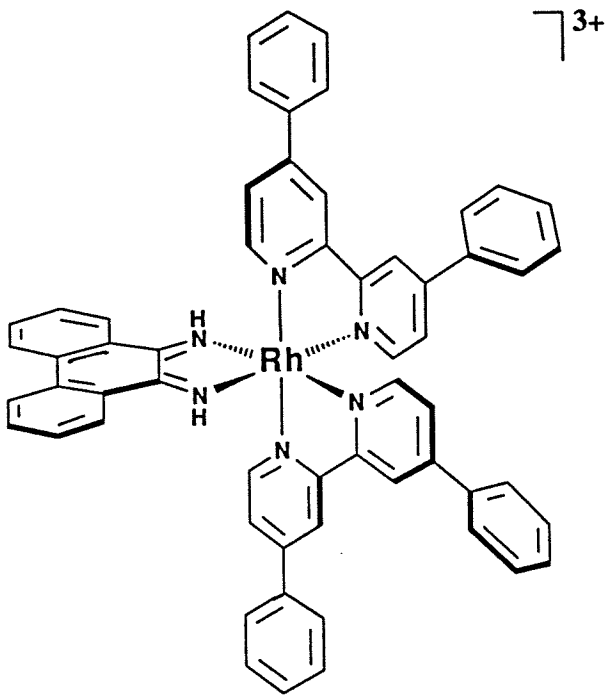
$[Rh(phi)_2(bpy)]^{3+}$

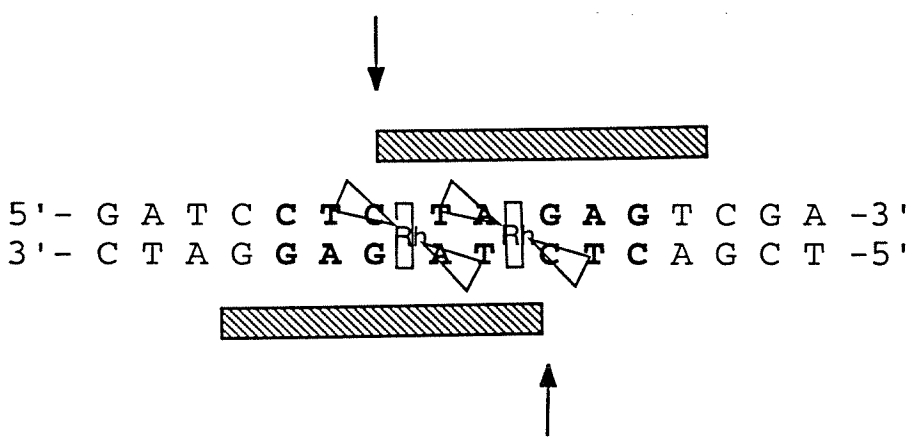
$[Rh(phi)_2(bpy)]^{3+}$ binds to most sites and is sequence-neutral in its DNA recognition. Thus, despite the possibility of hydrogen bonding by the ancillary phi-imine protons in $[Rh(phi)_2(bpy)]^{3+}$ with the DNA bases, it is $[Rh(phen)_2(phi)]^{3+}$ that shows the greater selectivity.

The opening of the major groove at a site is strongly governed by the differential propeller twist of the bases at that site. Binding and photocleavage by $[Rh(phen)_2(phi)]^{3+}$ may be used as a probe for DNA sites with an open major groove and in turn is a guage of the differential propeller twist associated with the site.[24b,c] Similarly, $[Rh(phen)_2(phi)]^{3+}$, when used as a probe for RNA tertiary structure, exclusively recognizes sites with base triplets, mismatches or bulges.[25] These sites are open for the complex to intercalate into as opposed to the double helical regions of RNA which are A-form and the major groove is too deep and narrow for effective intercalation by the complex.

All the metal complexes described above are devoid of hydrogen bonding functionalities and bind to DNA exclusively through hydrophobic interactions. Perhaps the most striking example of the extent to which we may exploit hydrophobic interactions is seen in $[Rh(DPB)_2(phi)]^{3+}$ (DPB=4,4'-diphenyl-2,2'-bipyridyl). The complex has a phi ligand for intercalative binding into DNA and large ancillary ligands for shape selection. Computer modeling indicates that the complex may span 6 bp when intercalated. Interestingly, photocleavage and MPE-Fe footprinting studies show that the Δ enantiomer of the complex binds specifically to an 8 bp palindromic site 5'-CTCTAGAG-3' (Figure 1.4).[26] The 8 bp site may be viewed as two overlapping 6 bp sites, 5'-CTCTAG-3', with the central CT base step being the intercalation and photocleavage site. A difference in binding affinity of 2.1 kcal/mol was observed for binding to the 8 bp site over an isolated 6 bp site. This remarkable 8 bp specificity and observed enhancement in affinity is indicative of cooperative binding of two molecules as a non-covalent dimer. $[Rh(DPB)_2(phi)]^{3+}$ effectively competes with enzymes for binding DNA sites. The

Figure 1.4. DNA recognition by Δ-[Rh(DPB)$_2$(phi)]$^{3+}$ as a non-covalent dimer. (Top): Structure of the complex. (Bottom) Schematic depicting the binding and photocleavage of the 8 base pair recognition sequence, 5'-CTCTAGAG-3', targeted by the complex. The recognition site is in bold letters, sites of phi intercalation are indicated as "Rh" and the photocleavage sites are marked by arrows. The shaded boxes represent DNA bases protected from MPE-Fe cleavage.

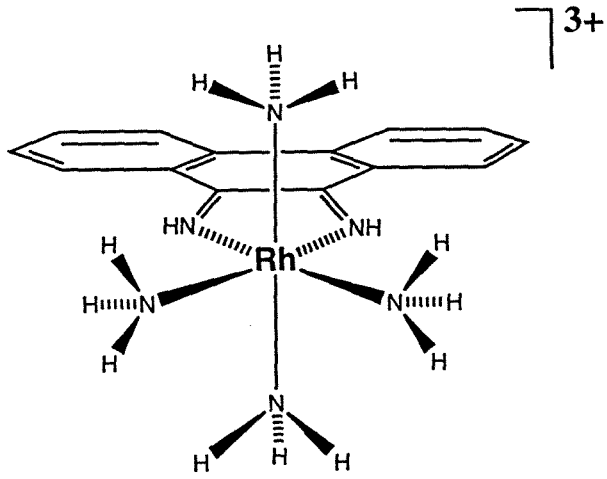

3+
Rh
N
H
N


5'- G A T C C T C T A G A G T C G A -3'
3'- C T A G G A G A T C T C A G C T -5'
Rh
Rh

dimerization interface may be modeled to consist of aromatic-aromatic interactions between the phenyl groups on one complex with the bpy ligand on the adjacent complex.

(b) *Importance of Hydrogen Bonding and van der Waals Contacts:* In parallel to the work on exploiting hydrophobic interactions for DNA recognition, the Barton group has also explored DNA recognition through hydrogen bonding. Towards this end, a series of phi complexes of rhodium have been prepared with aliphatic amines as the ancillary ligands. The simplest analog in this series, $[Rh(NH_3)_4(phi)]^{3+}$ preferentially recognizes 5'-GC-3' base steps.[27] This 5'-GC-3' preference was first identified by photocleavage studies and subsequently confirmed by ^{1}H-NMR spectroscopy.[27b] Molecular modeling based on NOE constraints attributes the 5'-GC-3' specificity to a hydrogen bonding contact between an axial ammine and the O6 of the 5'-G (Figure 1.5). Related work in the group demonstrates the enantioselective recognition of 5'-TA-3' base steps by Λ-$[Rh(en)_2(phi)]^{3+}$ (en=ethylene diamine) on account of a van der Waals contact between the 5-methyl group on thymine and the methylene groups on the en ligand in the Λ enantiomer (Figure1.6).[27] Both Δ and Λ enantiomers recognize 5'-GC-3' base pairs through hydrogen bonding contacts with the axial amines. This specificity may be built upon by introducing other functional groups. The complex, Δ–α-[Rh[(R,R)-Me_2trien](phi)]$^{3+}$, was designed to recognize 5'-TGCA-3' sequences (Figure 1.7) based on a combination of hydrogen bonds and van der Waals contacts. This rhodium complex also targets 5'-GC-3' steps through hydrogen bonding interactions similar to the parent complexes, $[Rh(NH_3)_4(phi)]^{3+}$ and $[Rh(en)_2(phi)]^{3+}$. However, at lower concentrations and with higher specificity, the complex cleaves at the predicted 5'-TGCA-3' sites.[28] The specificity is believed to be enhanced by the added methyl-methyl van der Waals contact. In fact, the site selectivity observed on DNA restriction fragments follows the trend: 5'-TGCA-3' > 5'-GCA-3' > 5'-GC-3'. This hierarchy of sites is consistent with the additive effects of a set of non-covalent interactions, each contributing approximately

Figure 1.5. Schematic illustration of the 5'-GC-3' preference observed for binding of $[Rh(NH_3)_4(phi)]^{3+}$ to DNA. The complex (top) is shown intercalated in a 5'-GC-3' base step (bottom). Note: the axial ammines of the complex are well positioned for hydrogen bonding with the guanine O6 atoms above and below the intercalation site. The complex can only span 2 base pairs.


3+
H
H
H
N
HN
NH
Rh
H
H
N
H
H
N
H
H
H
N
H
H
H

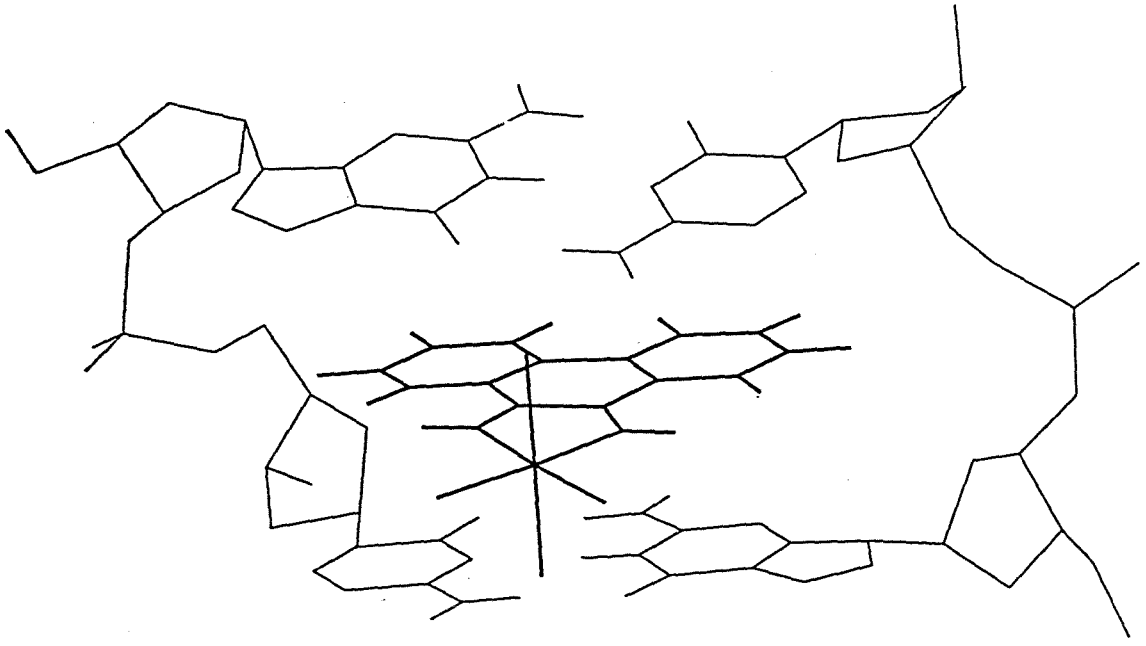
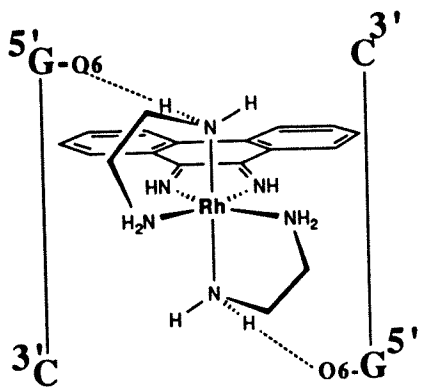
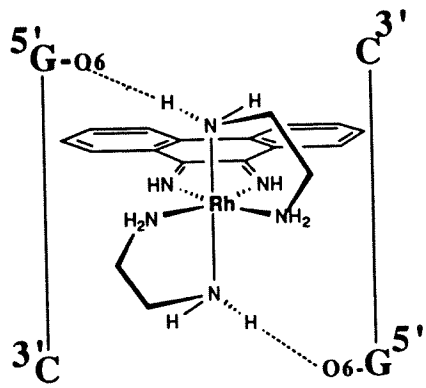


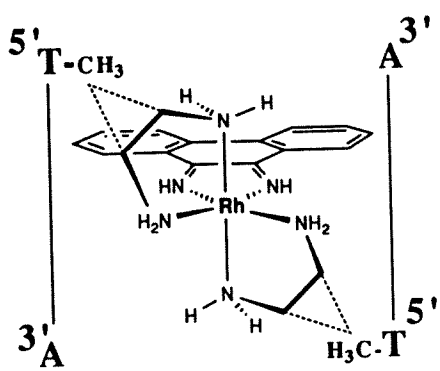
Figure 1.6. Schematic illustrating the basis for the recognition of 5'-GC-3' and 5'-TA-3' base steps by the enantiomers of $[Rh(en)_2(phi)]^{3+}$. Top: Both Δ- and Λ-$[Rh(en)_2(phi)]^{3+}$ contain axial amines for hydrogen bonding with the O6 of guanines at a 5'-GC-3' intercalation site. Therefore, both the enantiomers show 5'-GC-3' specificity. Bottom: The Λ-enantiomer is well positioned to make a van der Waals contact with the 5-methyl group on the 5'-T and is selective for 5'-TA-3' sites. The symmetry of the Δ-enantiomer does not afford similar interactions and, consequently, 5'-TA-3' sites show high Λ-enantioselectivity.

Λ-[Rh(en)$_2$ phi]$^{3+}$



Δ-[Rh(en)$_2$ phi]$^{3+}$





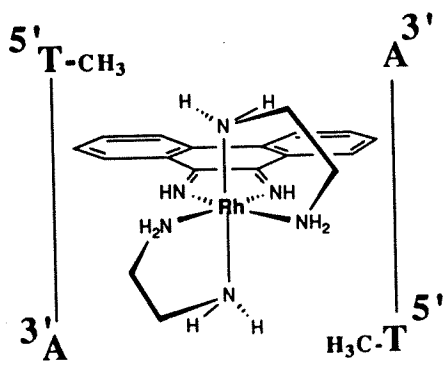
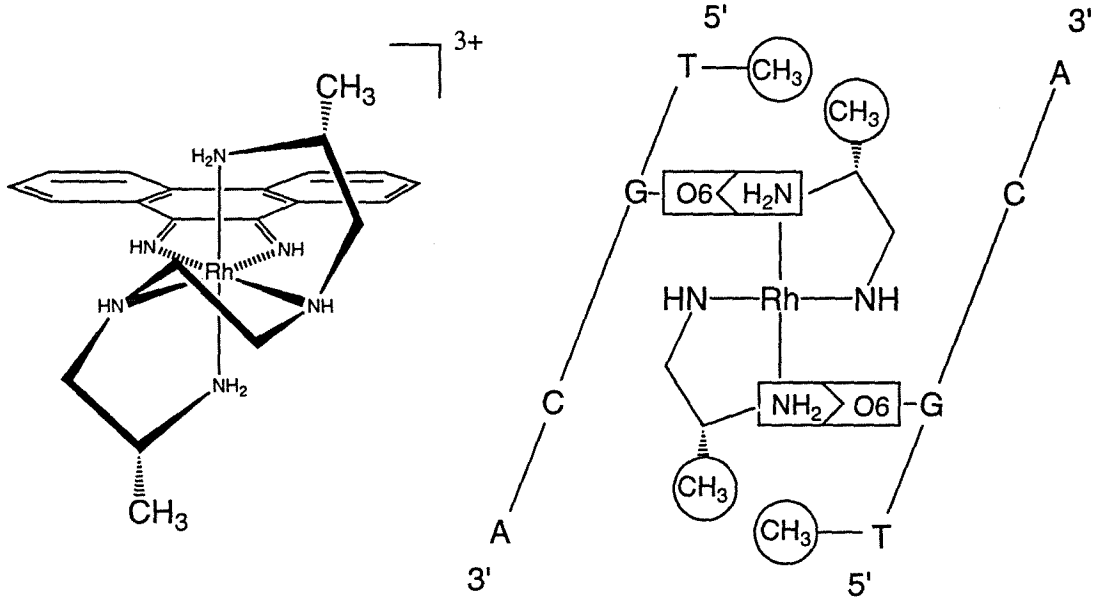


Figure 1.7. Schematic illustrating the assembly of non-covalent interactions to achieve 4 bp recognition by Δ–α-[Rh[(R,R)-Me_2trien](phi)]$^{3+}$. Left: Structure of the complex. Right: The axial amines hydrogen bond to the O6 of guanine at the central 5'-GC-3' intercalation site. The 5-methyl groups on the flanking 5'-T bases are in van der Waals contacts with the methyl groups on the rhodium complex. The two sets of interactions together account for the observed 5'-TGCA-3' recognition by Δ–α-[Rh[(R,R)-Me_2trien](phi)]$^{3+}$.


3+
CH3
H2N
HN
Rh
NH
HN
NH
NH2
CH3
5'
3'
T
CH3
CH3
A
G
O6
H2N
C
HN
Rh
NH
C
NH2
O6
G
A
CH3
CH3
T
3'
5'

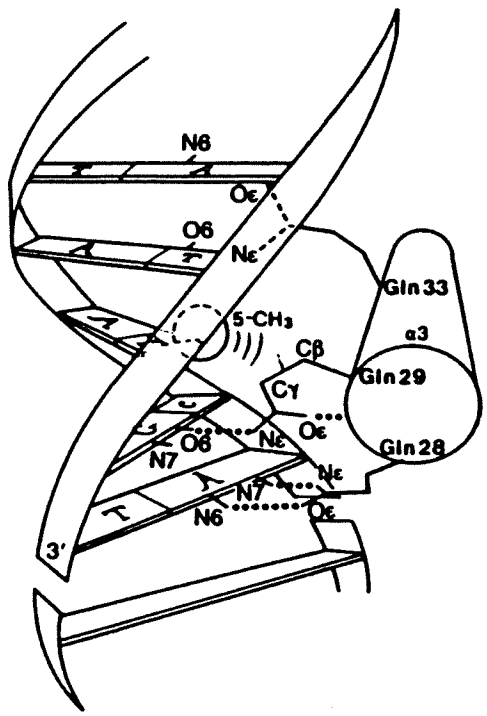
1 kcal/mole of binding stabilization. The binding of Δ–α-[Rh[(R,R)-Me_2trien](phi)]$^{3+}$ by intercalation between the 5'-GC-3' base step in the major groove is substantiated by 2D-NMR studies in the presence of the hexamer duplex d(GTGCAC)$_2$.[29] Efforts are currently underway in the group to increase the repertoire of DNA sites selected by appending different functional groups onto these polyamine scaffolds. It is hoped that larger DNA sites may be selected by building in additional DNA contacts piece by piece on these scaffolds.

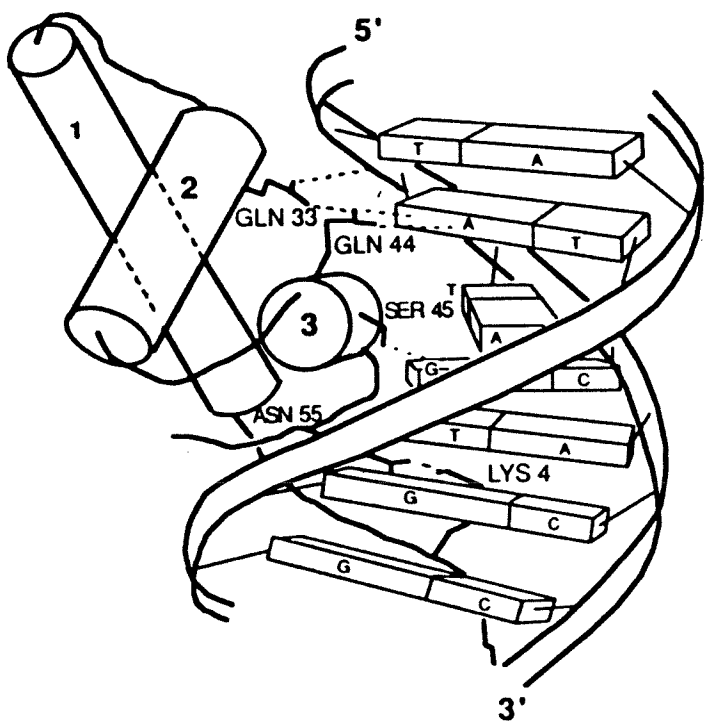
1.4. Design of Metallointercalator Based Repressors

Nature appears to have solved the DNA-recognition problem several times over by devising multiple families of DNA-binding proteins. A common feature in all these families is that the protein serves to deliver a DNA "recognition element," frequently an α-helix, to the major groove. The recognition element may also be an antiparallel β-sheet as in the β-ribbon motif DNA-binding proteins.[9] A significant number, or in several instances all, of the base specific contacts made by the proteins are from these recognition elements. Shown in Figure 1.8 are the recognition helices of the phage 434 repressor and the λ repressor proteins docked in the major groove. The positioning of the α-helix in the major groove establishes a sequence-specific ensemble of weak contacts between amino acid side chains and bases.

The work described in this thesis takes advantage of the ensemble of contacts presented by an α-helix by coupling it to the binding affinity provided by metallointercalators to design artificial repressors. We chose to develop [Rh(phi)$_2$(bpy)]$^{3+}$ or its close analogue, [[Rh(phi)$_2$(phen)]$^{3+}$, as the carrier for peptides because the complexes bind with high affinity ($K_d \leq 10^{-6}$ M) and are relatively sequence neutral in their DNA recognition.[24a] Indeed, [Rh(phi)$_2$(bpy)]$^{3+}$ serves as a useful photofootprinting reagent.[32] NMR and chemical reactivity studies of the binding of phi complexes of rhodium to DNA are consistent with intercalation from the major

Figure 1.8. Schematic illustration of the recognition helices of 434 repressor (top) and λ repressor (bottom) bound in the major groove of DNA. All the base specific contacts in the 434 repressor are made by residues in the recognition helix. The recognition helix in λ repressor makes significant base specific contacts but additional contacts are from other parts of the protein. Hydrogen bonds and van der Waals contacts are indicated.

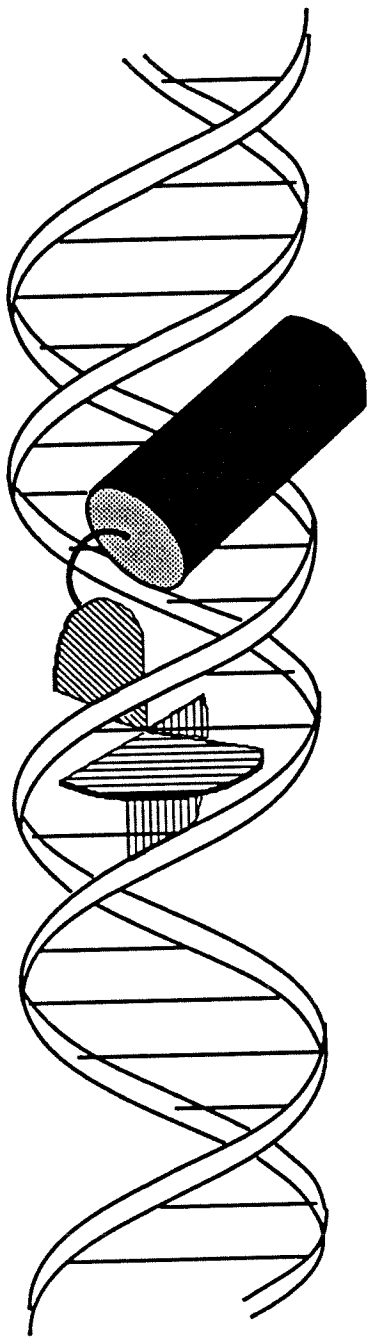

N6
Oε
O6
Nε
Gln 33
α3
5-CH3
Cβ
Gln 29
Cγ
Oε
Gln 28
Nε
O6
N7
N7
Nε
N6
Oε
3'


5'
1
2
GLN 33
GLN 44
3
SER 45
ASN 55
LYS 4
3'

groove,[24a, 27b, 30] a characteristic that makes these complexes particularly suitable as delivery systems for peptides (Figure 1.9). In contrast, most small molecules associate non-covalently with DNA in the minor groove.[31] The ability of the rhodium intercalator to cleave DNA upon photoactivation provides a convenient assay for determining the DNA binding site. In the simplest scenario, the phi center serves as the general binding module and the peptide as the site specific recognition module. However, more likely is the possibility that the metal-peptide chimera will serve as one discrete unit, much like DNA binding proteins, and recognition will be effected by interactions of both the modules with DNA. In either case, we hope to learn more about the principles of protein-DNA interactions.

The following chapters describe the synthesis (Chapter 2) and the DNA recognition properties of the designed metal-peptide complexes (Chapters 3 and 4). Two families of metal-peptide complexes based on the sequences of the recognition helices of phage P22 repressor (P22R) and phage 434 repressor (434R) have been investigated and they represent two aspects of DNA recognition by small peptides. Studies with the P22R based peptides led to the serendipitous discovery of a "glutamate switch" effect. Recognition in this family is exquisitely sensitive to the presence of a glutamate at position 10 in the sequence. An alternate model for DNA recognition by the peptides is proposed, taking into account the lack of orientational interactions with the rest of P22R. While these studies demonstrate the potential for using small metal-peptide complexes in site-specific DNA-recognition, they also underscore the importance of peptide orientation for reproducing protein-DNA recognition. The second family of metal-peptide complexes accurately reproduces features of operator recognition by the native 434R. The complexes recognize 5'-ACAA-3' operator sites, cleave specifically at 10 nM concentration, protect 7-10 bp of DNA around the operator sites from MPE-Fe cleavage and have low k_{off} rates, comparable to other DNA binding proteins, thus making them prototypes of artificial repressors. A comparison of the two families of metal-peptide

Figure 1.9. Schematic representation of a rhodium intercalator based peptide delivery system to the DNA major groove. The figure shows a helical peptide (cylinder) tethered via a short linker to an ancillary ligand on an octahedral rhodium complex. The rhodium complex binds to DNA by intercalation from the major groove and, as such, is disposed to deliver the tethered peptide to the DNA major groove for recognition.



complexes is made and considerations for future design of artificial repressors are discussed in Chapter 5.

The metal-peptide complexes incorporate elements of both shape selection as well as hydrogen bonding and van der Waals interactions to effect DNA recognition and compared to DNA binding proteins are 10-50 fold smaller in size. These studies illustrate the potential of metallointercalator based design of artificial repressors.

1.5. References.

1. Pabo, C. O.; Sauer, R. T. *Ann. Rev. Biochem.* **1992**, *61*, 1053.
2. Steitz, T. A. *Q. Rev. Biophys.* **1990**, *23*, 205.
3. (a) Pabo, C. O.; Sauer, R. T. *Ann. Rev. Biochem.* **1984**, *53*, 293. (b) Ptashne, M. *Trends Biochem. Sci.* **1984**, *9*, 142. (c) Johnson, P. F.; McKnight, S. L. *Ann. Rev. Biochem.* **1989**, *58*, 799. (d) Mitchell, P. J.; Tjian, R. *Science* **1989**, *245*, 371. (e) Frankel, A. D.; Kim, P. S. *Cell* **1991**, *65*, 717.
4. Helix-turn-helix motif. (a) Brennan, R. G.; Matthews, B. W. *J. Biol. Chem.* **1989**, *264*, 1903. (b) Harrison, S. C.; Aggarwal, A. K. *Ann. Rev. Biochem.* **1990**, *59*, 933.
5. Homeodomain. (a) Gehring, W.J.; Muller, M.; Affolter, M.; Percival-Smith, A.; Billeter, M.; Qian, Y. Q.; Otting, G.; Wuthrich, K. *Trends Genet.* **1990**, 6, 323. (b) Gehring, W. J.; Qian, Y. Q.; Billeter, M.; Furukubo-Tokunaga, K.; Schier, A. F.; Resendez-Perez, D.; Affolter, M.; Otting, G.; Wuthrich, K. *Cell* **1994**, *78*, 211.
6. Leucine Zipper motif. Kerppola, T. K.; Curran, T. *Curr. Opin. Str. Biol.* **1991**, *1*, 71.
7. Helix-loop-helix motif. (a) Jones, N. *Cell* **1990**, *61*, 9. (b) Davis, R. L.; Cheng, P.-F.; Lassar, A. B.; Weintraub, H. *Cell* **1990**, *60*, 733. (c) Anthony-Cahill, S. J.;

Benfield, P. A.; Fairman, R.; Wasserman, Z. R.; Brenner, S. L.; Stafford III, W. F.; Altenbach, C.; Hubbell, W. L.; DeGrado, W. F. *Science* **1992**, *255*, 979.

8. Zinc finger motif or steroid receptors. (a) Berg, J. M. *Science* **1986**, *232*, 485. (b) Klug, A.; Rhodes, D. *Trends Biochem. Sci.* **1987**, *12*, 464. (c) Berg, J. M. *Ann. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 405. (d) Vallee, B. L.; Coleman, J. E.; Auld, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, 88, 999. (e) Schwabe, J. W. R.; Klug, A. *Str. Biol.* **1994**, *1*, 345.

9. β-sheet motif. (a) Phillips, S. E. V. *Ann. Rev. Biophys. Biomol. Str.* **1994**, *23*, 671. (b) Raumann, B. E.; Brown, B. M.; Sauer, R. T. *Curr. Opin. Str. Biol.* **1994**, *4*, 36.

10. Otwinowski, Z.; Schevitz, R. W.; Zhang, R.-G.; Lawson, C. L.; Joachimiak, A.; Marmorstein, R. Q.; Luisi, B. F.; Siglar, P. B. *Nature* **1988**, *335*, 321.

11. (a) Schultz, S. C.; Sheilds, G. C.; Steitz, T. A. *Science* **1991**, *253*, 1001. (b) Kissinger, C. R.; Liu, B. S.; Martin-Blanco, E.; Kornberg, T. B.; Pabo, C. O. *Cell* **1990**, *63*, 579. (c) Jordan, S. R.; Pabo, C. O. *Science* **1988**, *242*, 893. (d) Aggarwal, A. K.; Rodgers, D. W.; Drottar, M.; Ptashne, M.; Harrison, S. C. *Science* **1988**, *242*, 899. (e) Luisi, B. F.; Xu, W. F.; Otwinowski, Z.; Freedman, L. P.; Yamamoto, K. R.; Sigler, P. B. *Nature* **1991**, *352*, 497. (f) Phillips, S. *Curr. Opin. Struc. Biol.* **1991**, *1*, 89. (g) Pavletich, N. P.; Pabo, C. O. Science 1991, 252, 809. (h) Pavletich, N. P.; Pabo, C. O. *Science* **1993**, *261*, 1701. (i) Raumann, B. E.; Rould, M. A.; Pabo, C. O.; Sauer, R. T. *Nature* **1994**, *367*, 754. (j) Omichinski, J. G.; Clore, G. M.; Schaad, O.; Felsenfeld, G.; Trainor, C.; Appella, E.; Stahl, S. J.; Gronenborn, A. M. *Science* **1993**, *261*, 438.

12. (a) Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3680. (b) Rould, M. A.; Perona, J. J.; Söll, D.; Steitz, T. A. *Science* **1989**, *246*, 1135.

13. (a) Desjarlais, J. R.; Berg, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 7345. (b) Desjarlais, J. R.; Berg, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 2256. (c) Nardelli, J.; Gibson, T.; Charnay, P. *Nucl. Acids Res.* **1992**, *20*, 4137.

14. (a) Rebar, E. J.; Pabo, C. O. *Science* **1994**, *263*, 671. (b) Jamieson, A. C.; Kim, S. H.; Wells, J. A. *Biochemistry* **1994**, *33*, 5689.

15. (a) Dervan, P. B. *Methods Enzymol.* **1991**, *208*, 497. (b) Mack, D. P.; Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7572. (c) Sluka, J. P.; Horvath, S. J.; Bruist, M. F.; Simon, M. I.; Dervan, P. B. *Science* **1987**, *238*, 1129. (e) Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847. (d) Mack, D. P.; Dervan, P. B. *Biochemistry* **1992**, *31*, 9399.

16. Chen, C. B.; Sigman, D. S. *Science* **1987**, *237*, 1197. (b) Pan, C. Q.; Landgraf, R.; Sigman, D. S. *Mol. Microb.* **1994**, *12*, 335. (c) Sigman, D. S.; Kuwabara, M. D.; Chen, C. B.; Bruice, T. W. *Methods Enzymol.* **1991**, *208*, 414.

17. Nagaoka, M.; Hagihara, M.; Kuwahara, J.; Sugiura, Y. *J. Am. Chem. Soc.* **1994**, *116*, 4085.

18. Pomerantz, J. L.; Sharp, P. A.; Pabo, C. O. *Science* **1995**, *267*, 93.

19. Bruist, M. F.; Horvath, S. J.; Hood, L. E.; Steitz, T. A.; Simon, M. I. *Science* **1987**, *235*, 777.

20. (a) Talanian, R. V.; McKnight, C. J.; Kim, P. S. *Science* **1990**, *249*, 769. (b) Talanian, R. V.; McKnight, C. J.; Rutkowski, R.; Kim, P. S. *Biochemistry.* **1992**, *31*, 6871 and ibid. **1993**, *32*, 1688. (c) Cuenoud, B.; Schepartz, A. *Science*, **1993**, *259*, 510. (d) Morii, T.; Simomura, M.; Morimoto, S.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 1150. (e) Ueno, M.; Sawada, M.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1994**, *116*, 11137.

21. (a) Pyle, A. M.; Barton, J. K. *Prog. Inorg. Chem.* **1990**, *38*, 413. (b) Chow, C. S.; Barton, J. K. *Methods Enz.* **1992**, *212*, 219. (c) Sitlani, A.; Barton, J. K. *Handbook of Metal-Ligand Interactions of Biological Fluids*, Marcel-Dekker,

Inc., New York, **1994** in press. (d) Dupureur, C. M; Barton, J. K. *Comprehensive Supramolecular Chemistry*, **1995**, Pergamon Press, in press.

22. (a) Barton, J. K.; Danishefsky, A. T.; Goldberg, J. M. *J. Am. Chem. Soc.* **1984**, *106*, 2171. (b) Barton, J. K.; Goldberg, J. M.; Kumar, C. V.; Turro, N. J. *J. Am. Chem. Soc.* **1986**, *108*, 2081. (c) Rehman, J. P.; Barton, J. K. *Biochemistry* **1990**, *29*, 1701.

23. (a) Mei, H. Y.; Barton, J. K. *J. Am. Chem. Soc.* **1986**, *108*, 7414. (b) Mei, H. Y.; Barton, J. K. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 1339.

24. (a) Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520. (b) Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432. (c) Campisi, D.; Morii, T.; Barton, J. K. *Biochemistry* **1994**, *33*, 4130.

25. (a) Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 2839. (b) Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972.

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

27. (a) Krotz, A.; Kuo, L. Y.; Sheilds, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 3877. (b) Collins, J. G.; Shields, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 9840.

28. Krotz, A.; Hudson, B. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12577.

29. Hudson, B. P.; Collins, G. J.; Barton, J. K. *unpublished results.*

30. (a) Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303. (b) David, S. S.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 2984.

31. Searle, M. S. *Prog. NMR Spec.* **1993**, *25*, 403.

32. Uchida, K.; Pyle, A. M.; Morii, T.; Barton, J. K. *Nucl. Acids Res.* **1989**, *17*, 259.

Chapter 2:

Coordinatively Saturated Rhodium Complexes Containing Appended Peptides: Synthesis and Characterization of their Physical Properties#

2.1. Introduction

There has been increased attention focused on the assembly of peptides containing coordinated transition metal complexes.[1-7] The coordination geometry about a metal ion provides a rigid, well defined center from which to append peptides for a specific function. As with larger metalloproteins, the metal ion may serve a structural role in bringing together discrete elements of peptide secondary structure or the metal ion may serve a functional role in catalysis.[4-7] Furthermore, as is also found with larger protein systems, the presence of the transition metal center provides a convenient spectroscopic handle to assay function. Examples of the incorporation of coordination chemistry in peptide design include the application of metal ion coordination to stabilize peptide α-helices and β-turns,[1, 2] the *de novo* design of three and four helix bundle proteins through crosslinking peptide helices by metal ions,[1, 5] and the construction of donor-acceptor assemblies in studies of photoinduced electron transfer across peptides.[3]

The Barton laboratory has focused on the design of transition metal complexes to explore site-specific recognition of nucleic acids.[8] As a part of this effort to construct smaller functional mimics of DNA-binding proteins, we became interested in incorporating appended peptides in our design. The goal is to create metal-peptide constructs, as shown in Figure 2.1, for site-selective recognition of double helical DNA in which the recognition characteristics of the complex are governed by the appended peptide. Towards this end, peptides (11-15 residues) were appended to a sequence-

---

Adapted from Sardesai, N. Y.; Lin, S., C.; Zimmermann, K.; Barton, J. K. *Bioconjugate Chem.* **1995**, accepted for publication.

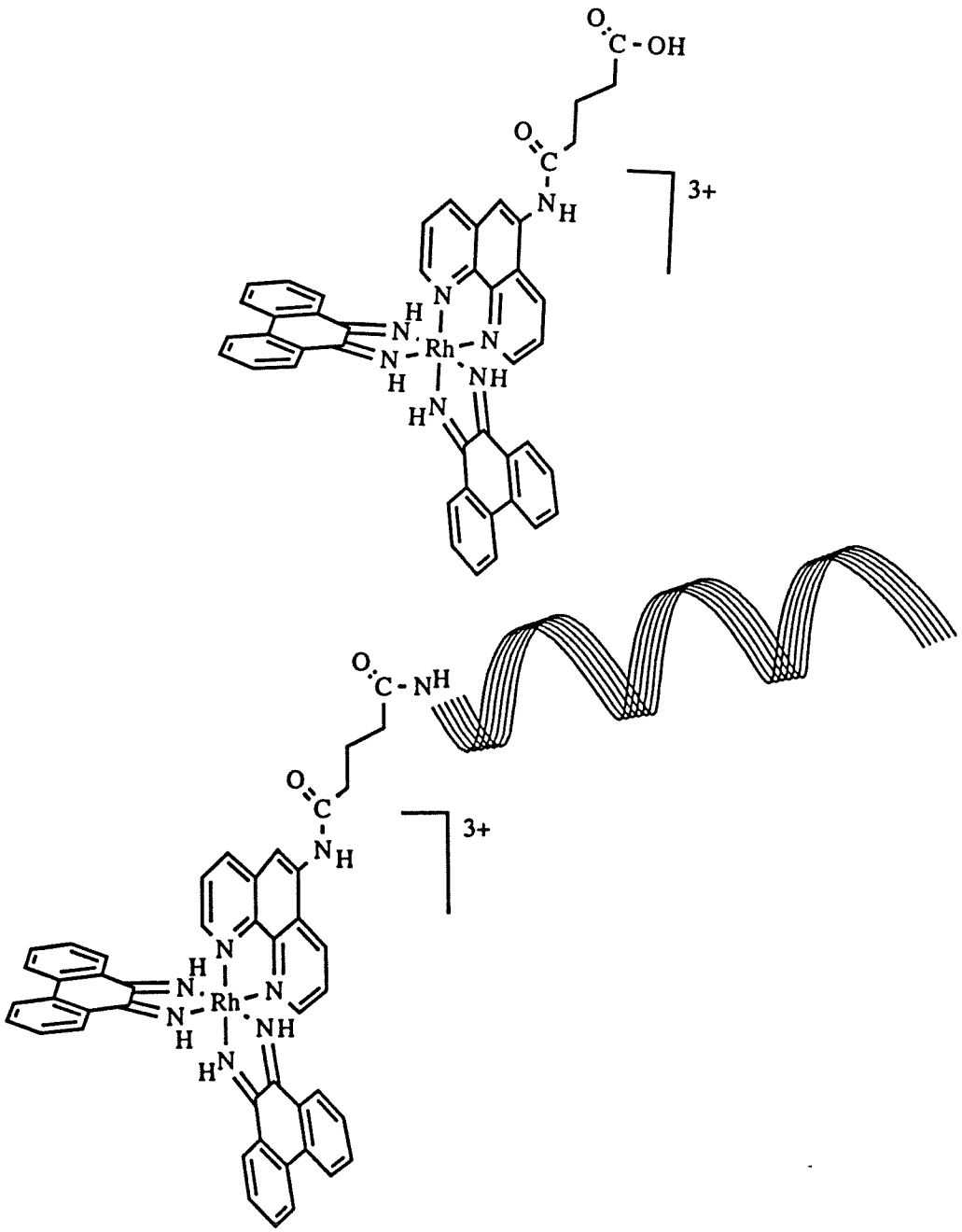


Figure 2.1. Structures of $[Rh(phi)_2(phen')]^{3+}$ (top) and a generic metal-peptide complex, $[Rh(phi)_2(phen')]^{3+}$-*peptide* (bottom). The N-terminus of the peptide is linked to the parent rhodium complex via a glutaryl linker. The peptide is shown schematically as an α-helix.

neutral[9] DNA-metallointercalator, $[Rh(phi)_2(phen)]^{3+}$ (phi=9,10-phenanthrenequinone diimine; phen=1,10-phenanthroline). This chapter describes the synthesis and characterization of metal complexes containing different tethered peptides, and the advantages and limitations of the different methodologies adopted.

Initial synthetic strategies involved solution phase coupling or coordination reactions. These were later abandoned in favor of assembling the metal-peptide complexes on a solid support. Peptides of length 11-15 amino acids have been synthesized by standard solid phase synthesis.[10, 11] Thereafter, metal complexes are appended onto the resin through two distinct coupling methodologies: (i) by direct coupling of $[Rh(phi)_2L]^{3+}$ (L= a bidentate chelator containing a pendant carboxylate) to the peptide on the resin; or (ii) by first coupling L to the resin-bound peptide, followed by coordination of $[Rh(phi)_2]^{3+}$ to L. Meyer and co-workers have used a similar strategy of direct coupling to synthesize a tripeptide containing $[Ru(bpy)_3]^{2+}$.[3] These syntheses complement solution phase coordination methodologies (vide infra).[1] The solution method relies, however, on the selective coordination or coupling of the metal center to the desired side chain functionalities on the peptide. Regioselective ligation in solution may be limited if the desired peptide sequence contains, for example, several cysteine or histidine residues. Similarly, regioselective coupling through amide bond formation may be limited if the peptide contains more than one free amine including the amino-terminus, for example, owing to the presence of lysine residues. Instead, coordination directly on the resin offers control through the selective deprotection of side-chain functionalities for ligation. More recently, solid phase synthetic methodology has proved convenient in the synthesis of combinatorial libraries of compounds and, in particular, of peptide based ligands and receptors.[12] Developing a viable solid phase synthetic protocol for the metal-peptide complexes is therefore essential for any such future applications.

The characterization of the metal-peptide constructs using mass spectrometric analysis in addition to the more conventional spectrophotometric methods is also

described in this chapter. Here, too, the metal ion yields distinct advantages. In the presence of the appended coordination complex, a novel pattern of peptide fragmentation is found which is not observed in the absence of the pendant metal complex, and this fragmentation greatly facilitates the sequence analysis of the appended peptide. The sequence information thus obtained includes location of a deletion of specific residues from amongst several alternate positions in the peptide sequence and presence of protecting groups on select residues following incomplete deprotection. The unique fragmentation pattern and sensitivity of detection has important implications towards developing a generalized mass-spec based sequencing strategy for peptides.

2.2. Experimental

2.2.1. Materials. $RhCl_3.6H_2O$ was purchased from Aesar Johnson-Matthey and 5-amino-1,10-phenanthroline was purchased from Polysciences Inc. All other chemicals were purchased from Aldrich. Anhydrous solvents were purchased from Fluka. $[Rh(phi)_2Cl_2]Cl$ was synthesized (83% yield) following published protocols.[13] MBHA[14] and PAM resins for tBoc synthesis were purchased from Applied Biosystems Inc. (ABI) and the amino acids from ABI or Bachem. *N*-α-tBoc-L-amino acids were used with the following side chain protecting groups: Arg(Mts), Asp(OBzl), Glu(OBzl), Lys(Cl-Z), Ser(Bzl), Thr(Bzl), Trp(CHO). FMOC peptides in the N-deprotected form on PAL resin were provided by Lederle Laboratories and had the following side chain protecting groups: Arg (PMC), Glu (tBu), Lys (tBoc), Ser (tBu).

2.2.2. Instrumentation. ^{1}H-NMR spectra were recorded on a 300 MHz GE QE Plus spectrometer or a Brucker AM500 spectrometer. Ultraviolet-visible spectra were recorded on a Hewlett-Packard 8452A diode array or Cary 219 spectrophotometer. Circular dichroism studies were performed on a Jasco J-500 or J-600 spectrometer using 1 cm path length cells. The peptide concentrations were determined by quantitative

amino acid analysis as the average of three runs on an ABI420 amino acid analyzer. The concentrations of metal-peptide complexes were determined by UV-visible spectroscopy using ε_{350} (isosbestic)=23,600 $M^{-1}cm^{-1}$. ^{252}Cf plasma desorption mass spectrometry (PDMS) was recorded on a time-of-flight spectrometer (Bio-Ion/Applied Biosystems 20 K, Uppsala, Sweden). Automated tBoc syntheses were done on an ABI430 peptide synthesizer. High performance liquid chromatography (HPLC) was carried out on a Waters 600E system equipped with a Waters 484 tunable detector. A Pharmacia FPLC equipped with a LKB 2141 variable wavelength monitor and a FPLC-Manager work station was also utilized for chromatographic separations of products.

2.2.3. Synthesis of Ligands and Metal Complexes.

5-(amidoglutaryl)-1, 10-phenanthroline, phen', 2. A suspension of 5-amino-1,10-phenanthroline, **1**, (1.0 g, 5.1 mmol) in 50 mL anhydrous pyridine was heated to 70°C. Then, glutaric anhydride (1.17 g, 10.2 mmol) was added. The solution was stirred at 100°C and another 585 mg (5.1 mmol) and 1.17 g (10.2 mmol) of anhydride were added after 1 h and 2 h respectively. After 3 h, the pyridine was reduced in vacuo to 5 mL and 250 mL of acetonitrile was added. The solution was stirred at room temperature for 1 h to precipitate the desired product. The precipitate was collected by filtration and washed with acetonitrile (2 x 20 mL) to yield phen' as an off-white powder (785 mg, 49.6%). TLC [silica gel, without fluorescent indicator; CH_2Cl_2-MeOH (1:1); stained with $(NH_4)_2Fe(SO_4)_2$] : $R_f = 0.41$.

^1H NMR (d_6-DMSO, 300 MHz):[15] 1.88 (*quint.*, $J = 7.3$, 2H, CH_2-$\underline{CH_2}$-CH_2); 2.33 (*t*, $J = 7.3$, 2H, CH_2-COO); 2.56 (*t*, $J = 7.3$, CH_2-CON); 7.70 (*dd*, $J = 8.1$, 4.3, 1H, H-C(8)); 7.78 (*dd*, $J = 8.4, 4.3$, 1H, H-C(3)); 8.15 (*s*, 1H , H-C(6)); 8.41 (*dd*, $J = 8.1, 1.7$, 1H, H-C(7)); 8.60 (*dd*, $J = 8.4, 1.6$, 1H, H-C(4)); 8.99 (*dd*, $J = 4.3, 1.7$, 1H, H-C(9)); 9.09 (*dd*, $J = 4.3, 1.6$, 1H, H-C(2)); 10.11 (*s br*, 1H, NH); 12.12 (*s br*, 1H, OH).

A substantial by-product observed in the synthesis is the corresponding glutarimide that results from cyclization of phen'. The filtrate from above was stripped of acetonitrile to yield a viscous oil. The N-5-(1, 10-phenanthroline)-glutarimide, **3**, was precipitated as a white powder by addition of ca. 100 mL of water. The precipitate was collected on a frit, washed with water, and dried under vacuum. The yield of this cyclized product accounts for the remainder of the starting material, 5-amino-1,10-phenanthroline.

TLC (same conditions as above): R_f=0.35.

^{1}H NMR (d_6-DMSO, 500 MHz): The O<u>H</u> and N<u>H</u> peaks at 12.12 and 10.11 ppm respectively (in the spectrum for phen' above) are absent. The other signals are slightly shifted to, respectively: 2.15 (*m*, 2H); 2.80 (*m*, 2H); 2.96 (*m*, 2H); 7.75 (*dd*, 1H); 7.80 (*dd*, 1H); 7.91 (*s*, 1H); 8.28 (*dd*, 1H); 8.48 (*dd*, 1H); 9.10 (*dd*, 1H); 9.14 (*dd*, 1H). The three aliphatic signals are more complicated because the protons become diastereotopic upon cyclization.

[Rh(phi)$_2$(DMF)$_2$](OTf)$_3$, 5. [Rh(phi)$_2$Cl$_2$]Cl, **4** (124 mg, 200 μmol), prepared as described earlier,[13] and silver triflate (154.2 mg, 600 μmol) were suspended in 10 mL dry DMF and heated at 65° C for 24 h in the dark. The solution becomes clear (dark red) with a lot of AgCl precipitate (white) and some free silver (black). The solution was filtered through a medium fritted funnel and used immediately in coordination reactions.

Bis(phenanthrenequinone diimine)((5-amidoglutaryl)-1,10-phenanthroline) rhodium(III) trichloride, [Rh(phi)$_2$(phen')]Cl$_3$, 6. **5** (200 μmol) was filtered directly into a flask containing **2** (100 mg, 320 μmol) and heated at 65° C for 18 h. At this stage a simple purification step may be attempted. The reaction mixture was rotovapped to dryness and taken up in ca. 10 mL 1:1 H_2O/MeOH. The solution was filtered to remove any insoluble particulates (most likely, some unreacted **4**) and passed through an anion

exchange column (Sephadex QAE-A25) in the Cl^- form to exchange the triflate counter ions for chlorides. The eluent was stripped of solvent and taken up in ca. 20 mL H_2O. Only the desired product, **6**, is water soluble. The unreacted **4** and **2** being sparingly water soluble can be removed by filtering through a medium frit. The aqueous solution was lyophilized to dryness to yield an orange fluffy powder.

Alternatively, the reaction mixture was diluted with 50 mL 1:1 H_2O/CH_3CN and adsorbed on a Sephadex SP C-50, 40-120μ (H^+-form) cation exchange column. The column was washed with 300 mL H_2O/CH_3CN. The product mixture was then chromatographed with an HCl-gradient (100 mL 0.1 N and 100 mL 0.2 N HCl in H_2O/CH_3CN). The main fraction was eluted at ca. 0.15 N HCl. The solvents are removed in vacuo and the residue dissolved in 30 mL H_2O. After lyophilization, **6** (185 mg, 89.8%) was obtained as an orange fluffy powder.

^{1}H-NMR (d_6-DMSO, 300 MHz):[15] 1.86-1.96 (*m*, 2H, $CH_2C\underline{H}_2CH_2$); 2.36 (*t, J = 7.3*, 2H, CH_2COO); 2.71 (*t, J = 7.3*, 2H, CH_2CON); 7.52 (*t, J = 7.6*, 1H); 7.63 (*t, J = 7.6*, 1H); 7.78 (*t, J = 7.7*, 1H); 7.84 (*t, J = 7.8*, 1H); 8.09 (*dd, J = 8.3, 5.4*, 1H); 8.18 (*dd, J = 8.6, 5.3*, 1H); 8.43-8.54 (*m*, 8H); 8.61-8.64 (*m*, 1H); 8.66 (*s*, 1H); 8.75 (*d, J = 7.8*, 1H); 8.86 (*d(br), J = 7.9*, 1H); 8.90 (*d, J = 5.4*, 1H); 8.984 (*d, J = 7.7*, 1H); 8.985 (*d, J = 8.3*, 1H); 9.03 (*d, J = 5.3*, 1H); 9.28 (*d, J = 8.6*, 1H); 10.94 (*s*, 1H, H-N (amide)); 14.09, 14.12 (*2s*, 1H ea.); 14.26, 14.29 (*2s*, 1H ea.).

(D_2O, 500 MHz):1.98 (*m*, 2H); 2.5 (*t*, 2H); 2.8 (*t*, 2H); 7.34 (*t*, 2H); 7.47 (*t*, 2H); 7.64 (*t*, 2H); 7.72 (*t*, 2H); 7.9-8.83 (15H).

UV-Vis (H_2O, pH 5) λ_{max} (ε $M^{-1}cm^{-1}$): 251 (69,800); 270 (83,300); 380 (32,900).

PDMS obsd. (calc.): $[Rh(phi)_2(phen')]^{3+}$: 823.9 (824.7); $[Rh(phi)(phen')]^{3+}$: 618.9 (618.5); $[Rh(phi)_2]^+$: 515.4 (515.4); $[Rh(phen')]^+$ or $[Rh(phi)_2(phen')]^{2+}$: 412.3 (412.2); $[Rh(phi)]^+$: 309.6 (309.3); $[phi]^+$: 206.5 (206.3).

2.2.4. Measurement of pH-Dependent Optical Changes in the Spectrum of $[Rh(phi)_2(phen')]^{3+}$. A 19.5 μM solution of **6** in 10 mL of 0.001N HCl (pH 3.03) was used in the experiment. The solution was neutralized by addition of small amounts of NaOH (10, 5, 4 or 2 μL of 0.5, 0.1, 0.05 or 0.01 N NaOH depending on the pH) with vigorous stirring such that the total volume increase over the entire titration was at most 100 μL (corresponding to a concentration change of < 1%). Upon addition of NaOH, the pH was recorded and 1 mL solution was taken out to record the absorption spectra at each pH value. After scanning, the solution was returned to the bulk solution before further addition of NaOH. The titration was carried out over a pH range of 3.03 - 9.0 and back to 3.06 by acidifying with HCl in order to ensure that the changes were reversible.

2.2.5. Resolution of Enantiomers of $[Rh(phi)_2(phen')]Cl_3$. The eluent potassium (+) *tris*[*l*-cysteinesulphinato(2-)-S,N]cobaltate(III), $K_3[Co(l\text{-cysu})_3]$, was synthesized according to literature protocols.[16] A 115 x 2.5 cm column was filled with Sephadex-SP C-25 cation exchange resin that had been swelled in water. The resin was washed with 0.1 M KCl and then with copious amounts of water. **6** (35 mg) was dissolved in water and loaded on a minimum of resin. A 0.1 M $[Co(l\text{-cysu})_3]^{3-}$ solution was recirculated at a flow rate of about 1 mL/min. After 12 h, two distinct orange bands could be seen. Each band was separated and eluted off the resin using 0.2 M HCl in 1:1 H_2O/CH_3CN. The bands were dried in vacuo, dissolved in water and lyophilized to yield Δ- and Λ-$[Rh(phi)_2(phen')]Cl_3$ (14 mg each). For the Δ- isomer, $\Delta\varepsilon_{280}$= -26 $M^{-1}cm^{-1}$; $\Delta\varepsilon_{450}$= -10 $M^{-1}cm^{-1}$.[17]

2.2.6. Peptide Synthesis. Manual tBoc peptide synthesis was performed according to standard procedures.[10] The amino acids were coupled by using tBoc amino acid (4 eq.) and DCC (4 eq.) except for Arg and Glu which were HOBt esters (4 eq. amino acid, 4 eq. DCC and 4 eq. and HOBt) and Gln and Asn which were symmetric anhydrides (4 eq.

amino acid and 2 eq. DCC). All couplings were monitored by ninhydrin[18] and the cycle was repeated until >99% coupling efficiency was achieved.

A portion of each resin (50-200 mg) was cleaved and deprotected for characterization by HPLC, amino acid analysis and PDMS. The peptide resins were stored dry at -20°C.

2.2.7. Solution Phase Synthesis of Metal-Peptide Chimeras. Two solution phase coupling strategies were attempted.

Direct Coupling to Free Peptide: The desired peptide was deprotected, cleaved from the resin, and HPLC purified (see below). Reactions were carried out on a 2-40 μmole scale with 1:1, 1:2 or 4:1 ratios of **6**/free peptide. A typical coupling reaction involved activation of the carboxylate group on **6** through the formation of an HOBt ester. **6** (10 μmole), HOBt, DMAP, and DCC (in CH_2Cl_2) were reacted in a 1:1:0.25:1 eq. ratio respectively in DMF for 30 min. A solution of the free peptide (20 μmole in 1 mL DMF) was added dropwise to the solution of the activated ester and stirred overnight under an atmosphere of nitrogen. The reaction was stripped off solvents and redissolved in 30% CH_3CN/H_2O (0.1% TFA). An insoluble white precipitate of urea was filtered and the products isolated from the dark red solution by reverse phase FPLC or HPLC.

Other coupling reagents and non-nucleophilic bases were also tried. These include: (i) DIPC/DSC/pyridine; (ii) DCC/DSC/pyridine; (iii) DCC/DSC/DMAP; (iv) DIPC/HOBt/DMAP; (v) $SOCl_2$; (vi) TBTU/DIEA; (vii) TSTU/DIEA.

Coordination to Free Peptide: Phen' was first coupled to the N-terminus of the desired peptide on the resin using standard peptide coupling methodology with DCC as the coupling reagent. Preformed symmetric anhydride of phen' was formed using 2 eq. of phen'. This was added to 1 eq. of N-terminal deprotected peptide (on resin) in DMF and stirred at room temperature for 48-72 hours. Ninhydrin tests were carried out periodically

and additional DCC was added to ensure quantitative coupling.[18] The phen'-peptide was then cleaved from the resin as described below and purified on a Pharmacia FPLC reverse phase C18 PEP-RPC 10/10 column using a CH_3CN (0.1% TFA)/H_2O (0.1% TFA) gradient. The free peptides usually elute at 15% CH_3CN while the phen'-peptides elute at 25-40% CH_3CN. An optimized protocol is described in detail below.

Synthesis of [Rh(phi)$_2$(phen'-peptide)]: To a suspension of [Rh(phi)$_2$Cl$_2$]Cl (31 mg, 0.05 mmol) in 4 ml DMF (covered with Al foil and in dark) was added 500µl of a 0.3 M soln. of AgOTf in DMF. The solution was stirred at room temperature for 5 h. The precipitated AgCl was spun down to a pellet in a centrifuge (1.5min, 14000g). The clear dark red supernatant was transferred to another flask and used as such without characterization. This was presumed to be **5**. Phen'-peptide (0.05 mmol in 40% CH_3CN (0.1% TFA)/H_2O ; 3 ml total volume) was added to the solution of **5** and the mixture stirred at 60° C for 60h. The reaction was stripped of the solvent (to remove DMF) and the solid redissolved in 20% CH_3CN/H_2O (0.1% TFA). The solution was spun down to remove insoluble material and the orange-red supernatant was injected into an FPLC for purification by reverse phase chromatography.

2.2.8. Solid Phase Synthesis of Metal-Peptide Chimeras. Two strategies for coupling [Rh(phi)$_2$(phen')]$^{3+}$ have been successfully employed with both tBoc and FMOC peptides and are described below. The rhodium complex is stable to HF, TFMSA and TFA peptide cleavage and deprotection conditions, allowing it to be coupled to the peptide on the resin.

Removal of Amino-terminal Protecting Group. Resin containing the required peptide (40 µmol) was washed with CH_2Cl_2 (3 x 10 mL). The resin was treated with the TFA solution (25% (v/v) and 0.1% anisole in CH_2Cl_2) for 1.5 min and 30 min. After washing the resin with CH_2Cl_2 (6 x 10 mL), it was treated with 4 mL 10% (v/v) DIEA in

CH_2Cl_2 for 1.5 min (2x) and washed again with CH_2Cl_2(6 x 10 mL). The resin was dried on the aspirator and then under vacuum.

Removal of Base Labile Protecting Groups. Since phi complexes of rhodium are unstable to aqueous bases,* the base labile protecting groups like formyl-Trp in tBoc synthesis must be removed prior to coupling the rhodium complex. To deprotect the Trp side chain, the resin was treated with 10 mL 2-aminoethanol (6% in 5:95 water/DMF) for 30 min. The solution was drained and the resin washed with DMF (3 x 4 mL). The 2-aminoethanol treatment and washing were repeated. Finally, the resin was washed with CH_2Cl_2, 1:1 CH_2Cl_2/MeOH and absolute EtOH and dried under vacuum.

Coordination Strategy. *(a) Activation of Phen':* Phen' (24.6 mg, 80 μmol), TBTU (32 mg, 100 μmol), DMAP (ca. 1 mg, 4 μmol) and NMM (22 μL, 200 μmol) were taken in a 25 mL round bottom flask with 2 mL NMP and stirred at room temperature for 15 min.

(b) Coupling of Phen' to the N-terminus of Peptide: The N-terminus deprotected peptide-resin (40 μmol) was transferred to a 25 mL round-bottom flask and stirred with 5 mL CH_2Cl_2 for 15 min. The phen' solution was added to the slurry, and the reaction was stirred at ambient temperature for 36 h. The reaction was followed by ninhydrin assay.[18] Double coupling of phen' may be required to ensure > 95 % coupling efficiency. The resin was filtered to remove excess reagents and washed with CH_2Cl_2, 1:1 CH_2Cl_2/MeOH and absolute EtOH and dried under vacuum.

(c) Coordination of Resin-Peptide-Phen' to $[Rh(phi)_2(DMF)_2](OTf)_3$: The phen'-peptide resin was placed in a 25 mL round bottom flask in 5 mL CH_2Cl_2 and stirred for

---

* After treatment of **6** with 2 -aminoethanol at 0°C at pH 10 for 5 minutes, about 10% of the metal complex is degraded.

15 min to swell the resin. The $[Rh(phi)_2(DMF)_2](OTf)_3$ solution (40 μmol, in 5 mL DMF) was added and the resulting mixture stirred in the dark at 70°C for 36 h. The solution was removed by filtration and the dark red resin was washed several times with DMF and 1:1 CH_2Cl_2/DMF over 30 min with agitation. Finally, the resin was washed with CH_2Cl_2, 1:1 CH_2Cl_2/MeOH and absolute EtOH and dried under vacuum.

Direct Coupling Strategy. *(a) Activation of $[Rh(phi)_2(phen')]Cl_3$:* $[Rh(phi)_2(phen')]Cl_3$ (13 mg, 14 μmol), HOBt (4.4 mg, 14 μmol), DCC (3.1 mg, 14 μmol) and DMAP (ca. 1 mg, 4 μmol) were taken in a 5 mL round-bottom flask. Anhydrous NMP (0.5 mL) was added and the solution was stirred at room temperature for 15 min. The activation can also be done using DSC (1 eq.) and DMAP (ca. 0.3 eq.).

(b) Coupling of $[Rh(phi)_2(phen')]Cl_3$ to the N-terminus of the Peptide: The N-terminus deprotected peptide-resin (7 μmol) was placed into a 5 mL round-bottom flask. The rhodium solution was added to the resin. The whole mixture was stirred under argon in the dark for 24 h. The resin was washed with NMP until the filtrate was clear and further washed with CH_2Cl_2, 1:1 CH_2Cl_2/MeOH and absolute EtOH. The resin was dried in vacuo for several hours and the coupling was repeated. The resin turns from orange to dark red in conjunction with the coupling of the metal complex. Ninhydrin analysis indicated *ca.* 70% coupling efficiency.[18]

Deprotection and Cleavage of Peptides and Metal-peptide Complexes.

tBoc: Cleavage and deprotection of the tBoc peptides was accomplished according to either the two step TFMSA[19] or HF[10a] cleavage procedures. The HF cleavage was done in the presence of *p*-cresol and *p*-thiocresol for 60 min at 0° C by the Biopolymer Synthesis and Analysis Resource Center, Caltech. The two step TFMSA procedure is as follows:

Step 1: The resin (ca. 120 mg) following coupling of **6** to the peptide was taken in a 10 mL r.b. flask. m-Cresol (190 μL) and dimethyl sulfide (500 μL) were added to the flask and the mixture was cooled to 0° C under argon. This was followed by addition of 500 μL TFA and 200 μL TFMSA and stirred for 1h at 0° C. The reaction mixture was filtered through a sintered glass funnel and the resin washed extensively with CH_2Cl_2 and 1:1 CH_2Cl_2/EtOH. The resin was dried on an aspirator and then under high vacuum for 1h.

Step 2: The dry resin was transferred to a 10 mL r.b. flask with a stirrer and placed in an ice bath. Thioanisole (160 μL), ethane dithiol (80 μL) and TFA (1 mL) were added to the flask and allowed to cool to 0° C. TFMSA (200 μL) was added slowly and the reaction stirred for 25 min at 0° C and then for 35 min at room temperature following removal of the ice bath.

FMOC: The cleavage and deprotection of the FMOC peptides was done using 82.5% TFA, 5% H_2O, 5% phenol, 5% thioanisole and 2.5% ethanedithiol for 2-5 h.[20] The resin (ca. 100 mg) was transferred to a 10 mL r.b. flask with a stirrer and allowed to cool to 0° C in an ice bath. 1.5 mL of reagent B (250 mg phenol, 90 μL ethane dithiol, 180 μL thioanisole, 180 μL water and 3.3 mL TFA) was added to the flask and stirred. After 10 min the ice bath was removed and the reaction mixture was allowed to warm up to room temperature with stirring over 1.5-2 h.

Work-up of Cleaved Peptides and Metal-peptide Complexes. The cleavage solutions were filtered into chilled (-20° C) *tert*-butyl methyl ether (50 mL). After storing the solution at -70° C for at least 30 minutes, the solution was centrifuged for 10 minutes at 5,000 rpm. The ether was decanted off and the orange precipitate was washed with cold ether twice. Alternatively, the peptide was isolated by filtration through a fine sintered glass funnel. The precipitate was then taken up in 5% acetic acid and lyophilized to dryness or directly purified by HPLC. The residual resin should be pale orange, but

some of the metal-peptide chimera may remain associated with the resin giving it a darker color. Additional cleaved chimera may be recovered by soaking the resin with DMF/CH_2Cl_2 for a longer period of time.

2.2.9. Chromatographic Purification of Metal-peptide Complexes. The metal-peptide complexes were purified by HPLC on a Vydac semi-preparative C_{18} reverse phase column using a water (0.1% TFA)/acetonitrile (0.1% TFA) gradient (15%-40% acetonitrile over 19 min) with a flow rate of 4.5 mL/min. Alternatively, the complexes were purified by FPLC on a PEP-RPC 10/10 C_{18} reverse phase column using appropriate gradients as indicated. The free peptides eluted first followed by phen'-peptides and finally the metal-peptide complexes (35-40% acetonitrile). The elution profiles were monitored at 220 nm and 370 nm with peaks containing the intact metal-peptide complexes showing the strongest absorbance at 370 nm. The chromatographed fractions corresponding to the metal-peptide complexes were lyophilized and stored dry at -20° C.

2.2.10. Mass Spectrometric Characterization. Molecular weight determinations of the peptides, phen'-peptides and the tethered peptide complexes were carried out using ^{252}Cf PDMS at an accelerating voltage of 15 kV. The samples were adsorped on nitrocellulose surfaces by applying solutions (200 pmol) of the samples in 25% acetonitrile (0.1%TFA) - 75% water and allowed to dry. Excess salt, if present, was removed from the adsorbed samples by rinsing with 1:1 EtOH/water prior to data accumulation.

2.2.11. Representative Metal-Peptide Complexes. (The complete characterization data for all the metal-peptide complexes synthesized has been included in section 2.6 as an appendix.)

[Rh(phi)$_2$(phen')]$^{3+}$-LQQAIEQLQNAAAA-COOH. The peptide was synthesized on the PAM resin using automated tBoc techniques and the rhodium complex coupled using both the direct coupling and coordination method. After cleavage and deprotection by TFMSA, the chimera was isolated as an orange powder. Amino acid analysis observed (calculated) ratio: (By direct coupling) Asx 0.9 (1), Glx 4.8 (5), Gly 1.2 (0), Ala 5.1 (5), Val 0.6 (0), Ile 1.0 (1), Leu 2.0 (2). Small amounts of Gly and Val are seen in the hydrolysis blank of the amino acid analyzer and in the hydrolysis of the skeletal complex which become significant if the amount of sample analyzed is small. Nonetheless, the other amino acids are in the correct ratio. (By coordination) Asx 1.1 (1), Glx 5.0 (5), Ala 5.0 (5), Ile 1.2 (1), Leu 2.2 (2). Molecular ion peak observed (calculated): 2274.7 (2275.3). ^{1}H NMR (D_2O, 500 MHz): Spectrum shows the characteristic signals for the parent complex, **6**, in the aromatic region (7-12 ppm) and, the linker and peptide signals in the aliphatic region (0-5 ppm). There are small changes in the positions of the various aromatic signals upon attachment of the peptide, particularly the singlet corresponding to C6-H on the phen.

[Rh(phi)$_2$(phen')]$^{3+}$-TQQSKKQLQNKAAA-CONH$_2$. The peptide was synthesized on the MBHA resin using automated tBoc techniques and the rhodium complex coupled using the coordination method. After cleavage and deprotection by TFMSA, the chimera was isolated as an orange powder. Amino acid analysis observed (calculated) ratio: Asx 0.7 (1), Glx 3.2 (4), Ser 1.0 (1), Thr 1.0 (1), Ala 3.0 (3), Leu 1.2 (1), Lys 2.5 (3). Molecular ion peak observed (calculated): 2348.2 (2349.5).

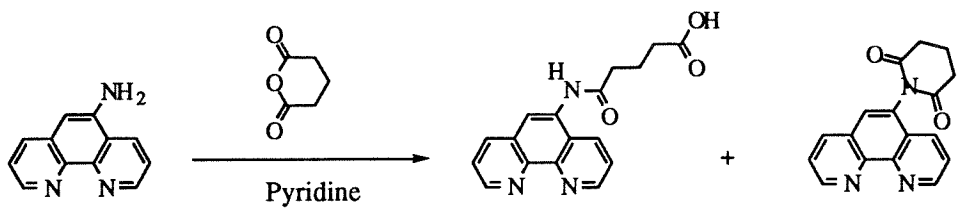
[Rh(phi)$_2$(phen')]$^{3+}$-KAEGAKAFFKRSA-CONH$_2$. The peptide was obtained from American Cyanamid Lederle Labs following automated FMOC synthesis and the rhodium complex coupled using the coordination method. After cleavage for 2h, the metal-peptide complex was isolated as an orange powder. Amino acid analysis observed (calculated) ratio: Glx 1.0 (1), Ser 0.9 (1), Gly 1.1 (1), Arg 1.1 (1), Ala 4.1 (4), Phe 2.0 (2), Lys 2.8 (3). Molecular ion peak observed (calculated): 2216.0 (2217.6).

2.3. Results

2.3.1. Synthesis of Phen' and $[Rh(phi)_2(phen')]^{3+}$. The synthesis of phen' is achieved in one step by reaction of 5-amino-1, 10-phenanthroline with glutaric anhydride in dry pyridine in 50% yield (Scheme 2.1). A substantial biproduct of the reaction is the corresponding glutarimide that results from intramolecular cyclization of phen' to form a six membered cyclic imide and accounts for the remainder of the starting material, 5-amino-1, 10-phenanthroline. The NMR spectra of **2** and **3** are shown in Figures 2.2 and 2.3 respectively. The aliphatic signals of **3** appear as two sets of multiplets indicating that the protons are diasterotopic, possibly because of hindered rotation about the C5-N bond.

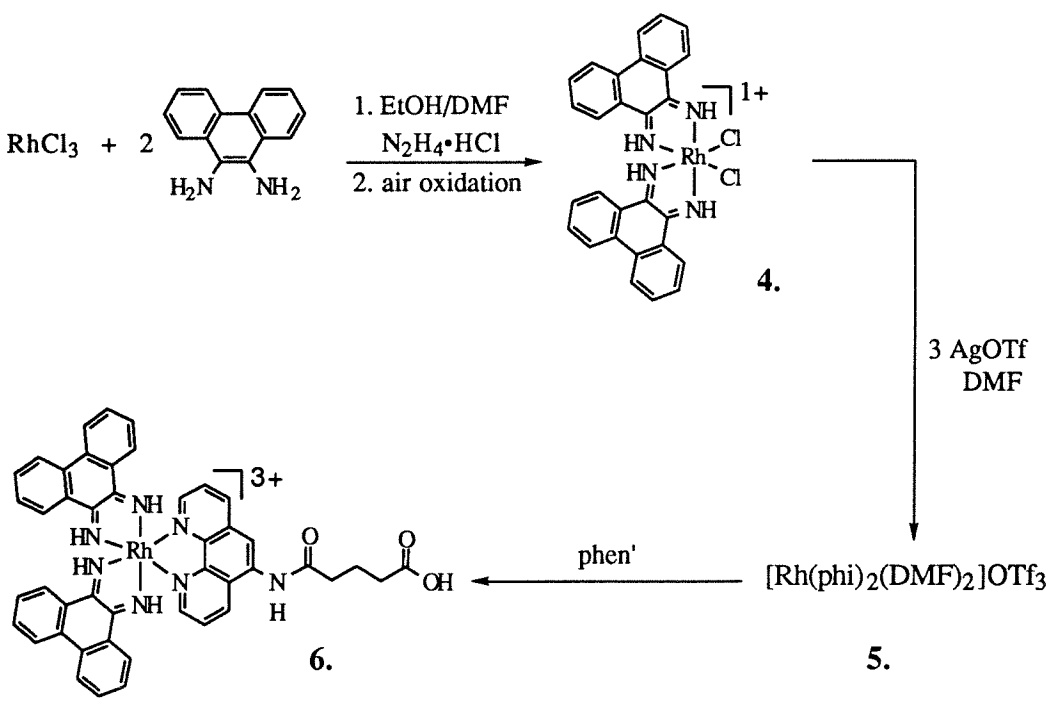
Scheme 2.2 illustrates the synthesis of **6** containing a pendent carboxylate for attachment of a peptide. **4** is synthesized by the previously reported literature method[13] and then reacted with silver triflate to exchange the chloride ligands. After complete removal of silver ions, the substitutionally facile **5** is then heated with phen' to promote coordination of the third chelating ligand. No coordination to the rhodium center by the ancillary carboxylate has been observed. The NMR spectrum of $[Rh(phi)_2(phen')]^{3+}$ is shown in Figure 2.4 and the ^{252}Cf PDMS spectrum in Figure 2.5.

2.3.2. pH Dependent Optical Changes in the Spectrum of $[Rh(phi)_2(phen')]^{3+}$. The pH dependent UV-Visible spectrum of **6** is shown in Figure 2.6. Isosbestic points were observed at 220, 280, 312, 349, and 422 nm. The absorbance at 350 nm (ε_{350}=23600 $M^{-1}cm^{-1}$ for $[Rh(phi)_2(bpy)]Cl_3$)[21] has been used to estimate the concentration of $Rh(phi)_2$ complexes in this report. As with $[Rh(phi)_2(bpy)]^{3+}$ the pH-dependent spectral shifts for $[Rh(phi)_2(phen')]^{3+}$ are dramatic and completely reversible. The phi-centered transition shifts 29 nm to the blue when titrated from pH 3.0 to 9.0 (Figure 2.6). A plot of the phi absorbance maximum vs. pH yields a classic pH titration curve with a point of inflection (pK_a) at approximately 6.4 pH units (Figure 2.7). The intensity of the transition decreases by approximately 30% on going from pH 3.0 to pH 9.0. For comparison, the



1. 2. 3.

Scheme 2.1. Synthesis of phen'.



Scheme 2.2. Synthesis of $[Rh(phi)_2(phen')]^{3+}$.

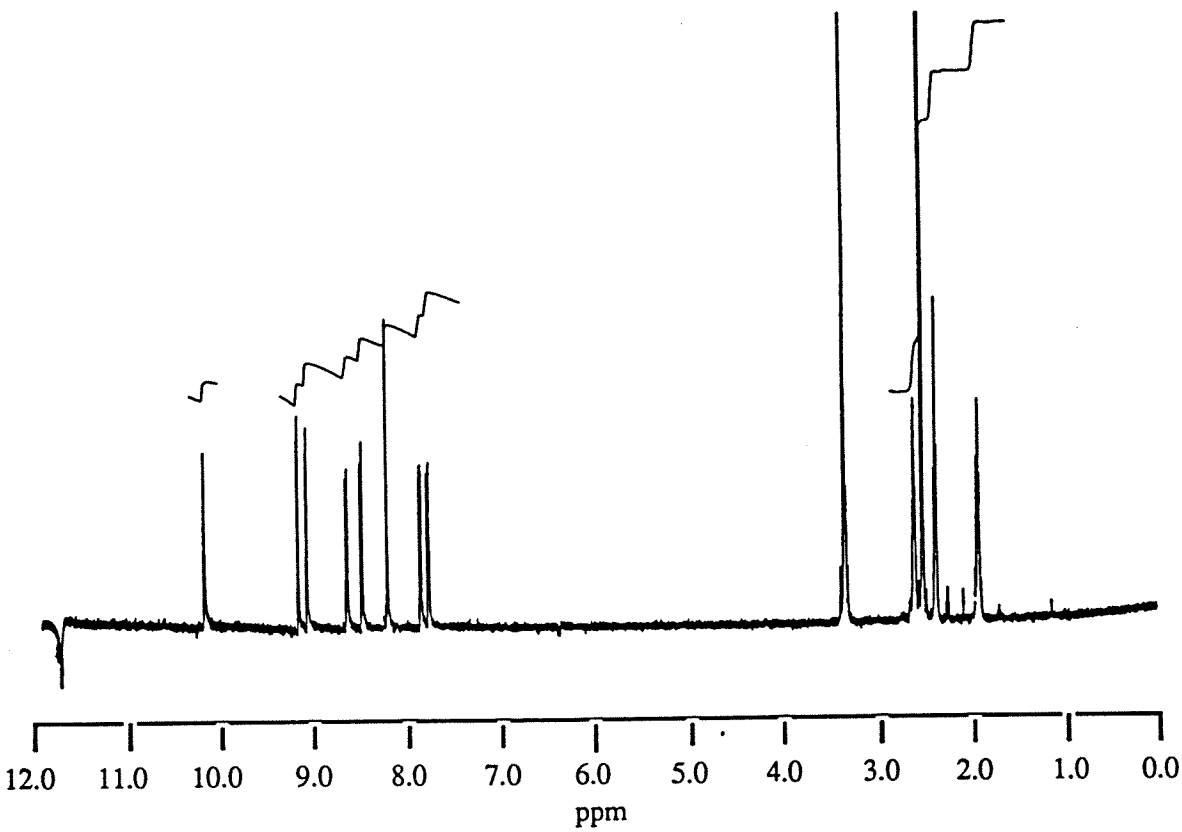
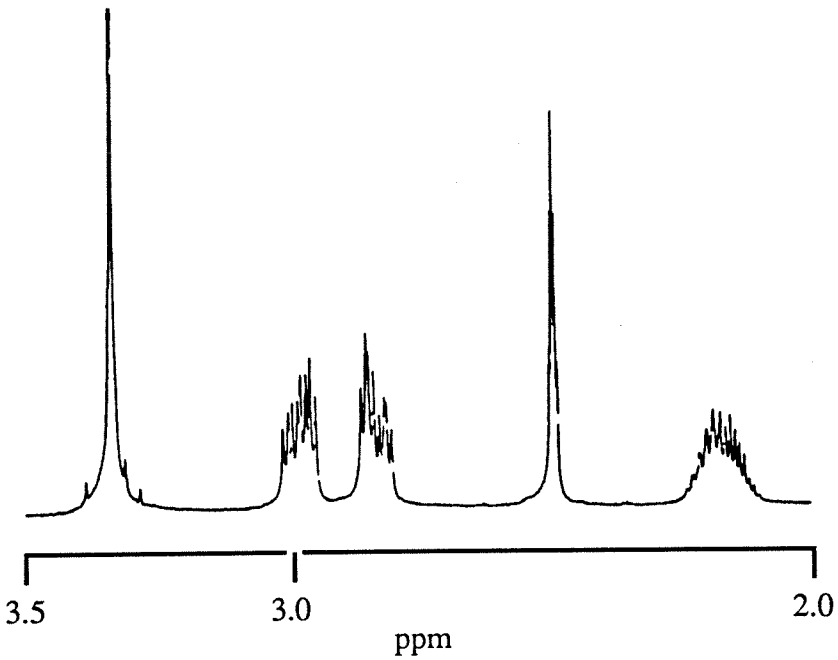


Figure 2.2. 500 MHz ^{1}H-NMR spectrum of phen' in DMSO.

Figure 2.3. 500 MHz ^{1}H-NMR spectrum of cyclized phen' in DMSO (bottom) and an expansion of the aliphatic signals (top).


3.5
3.0
2.0
ppm

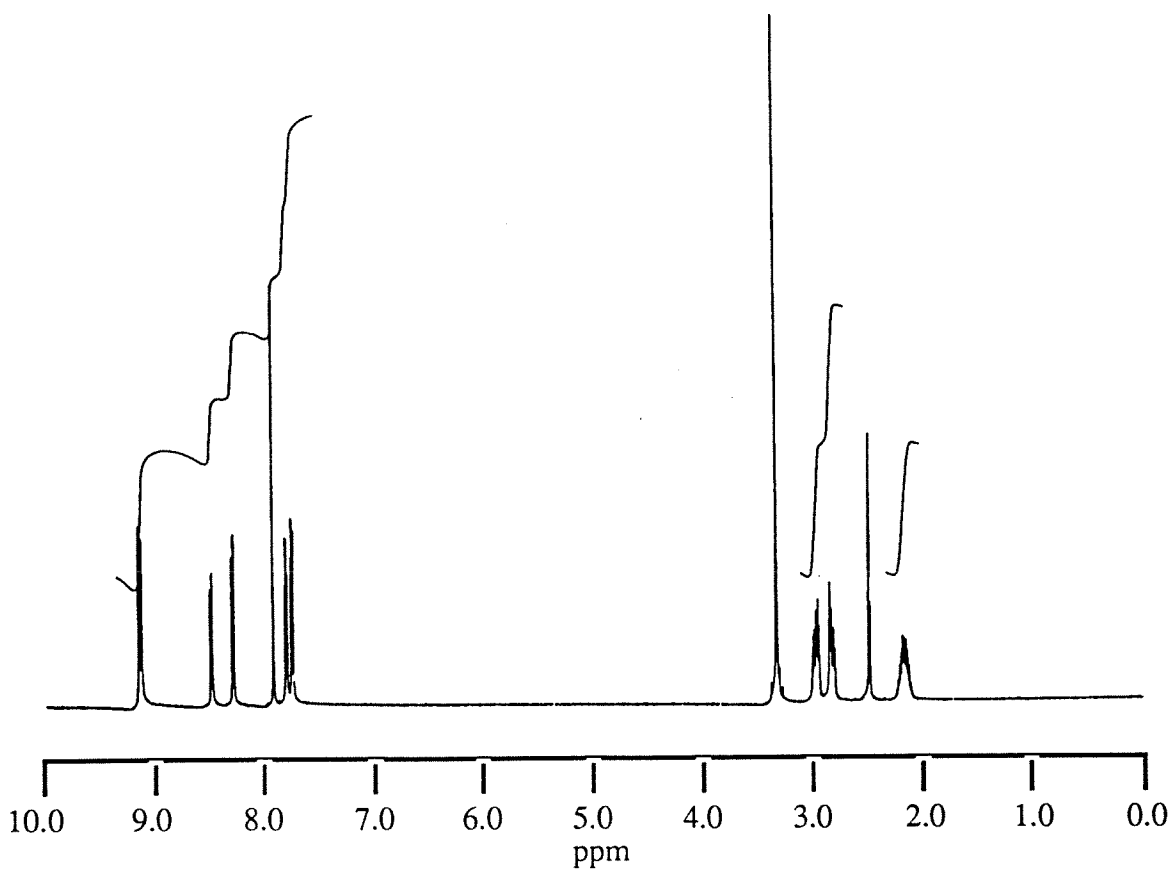
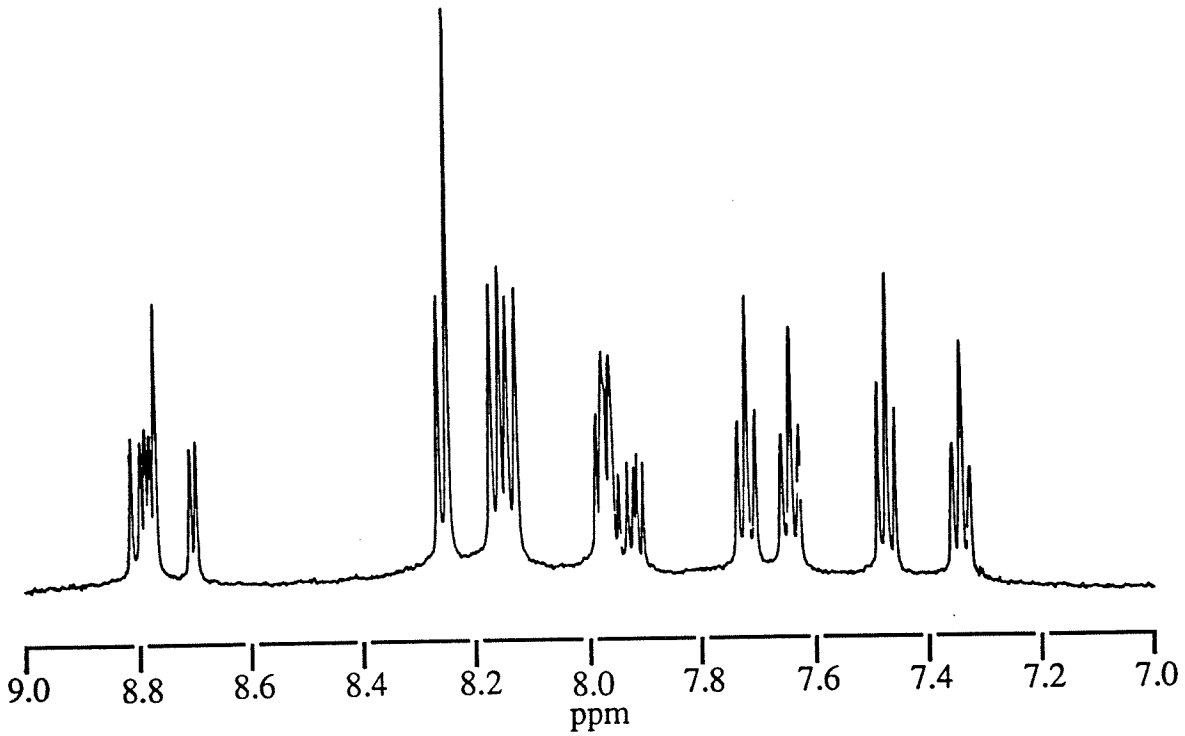
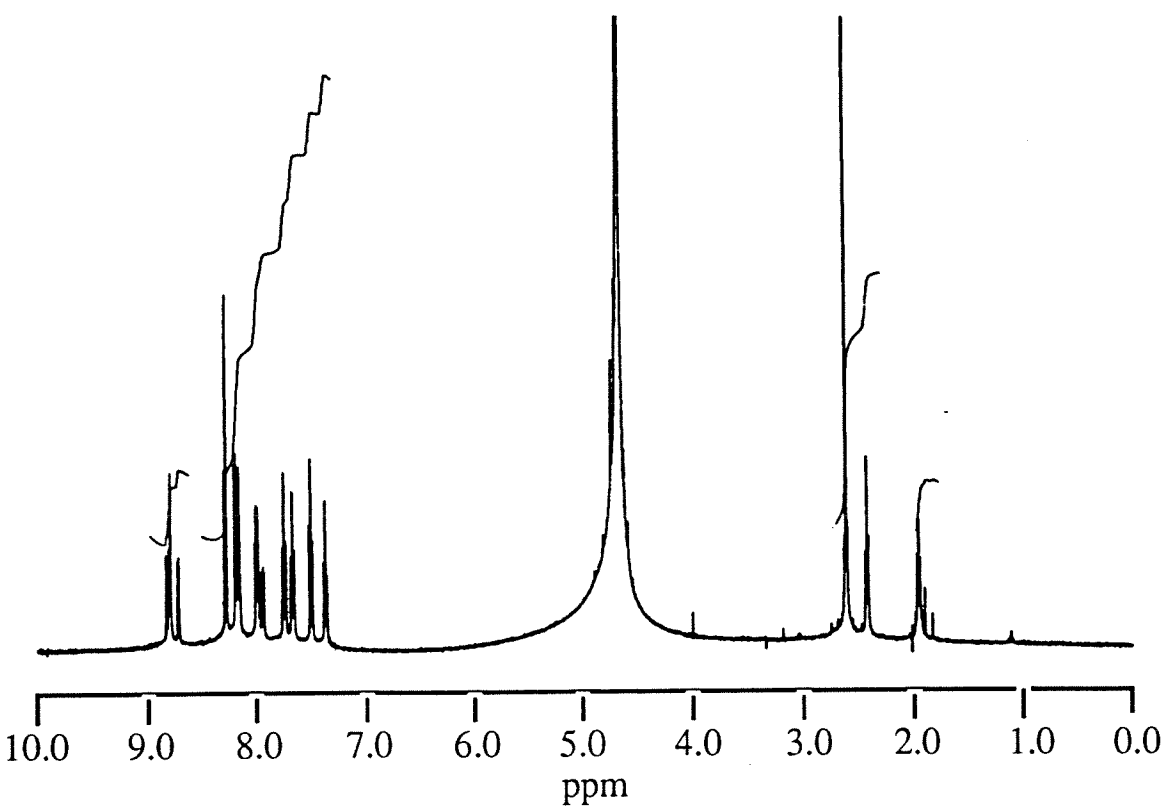

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm

Figure 2.4. 500 MHz ^{1}H-NMR spectrum of [Rh(phi)$_2$(phen')]Cl$_3$ in D_2O (bottom) and an expansion of the aromatic signals (top).


9.0
8.8
8.6
8.4
8.2
8.0
7.8
7.6
7.4
7.2
7.0
ppm


10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm

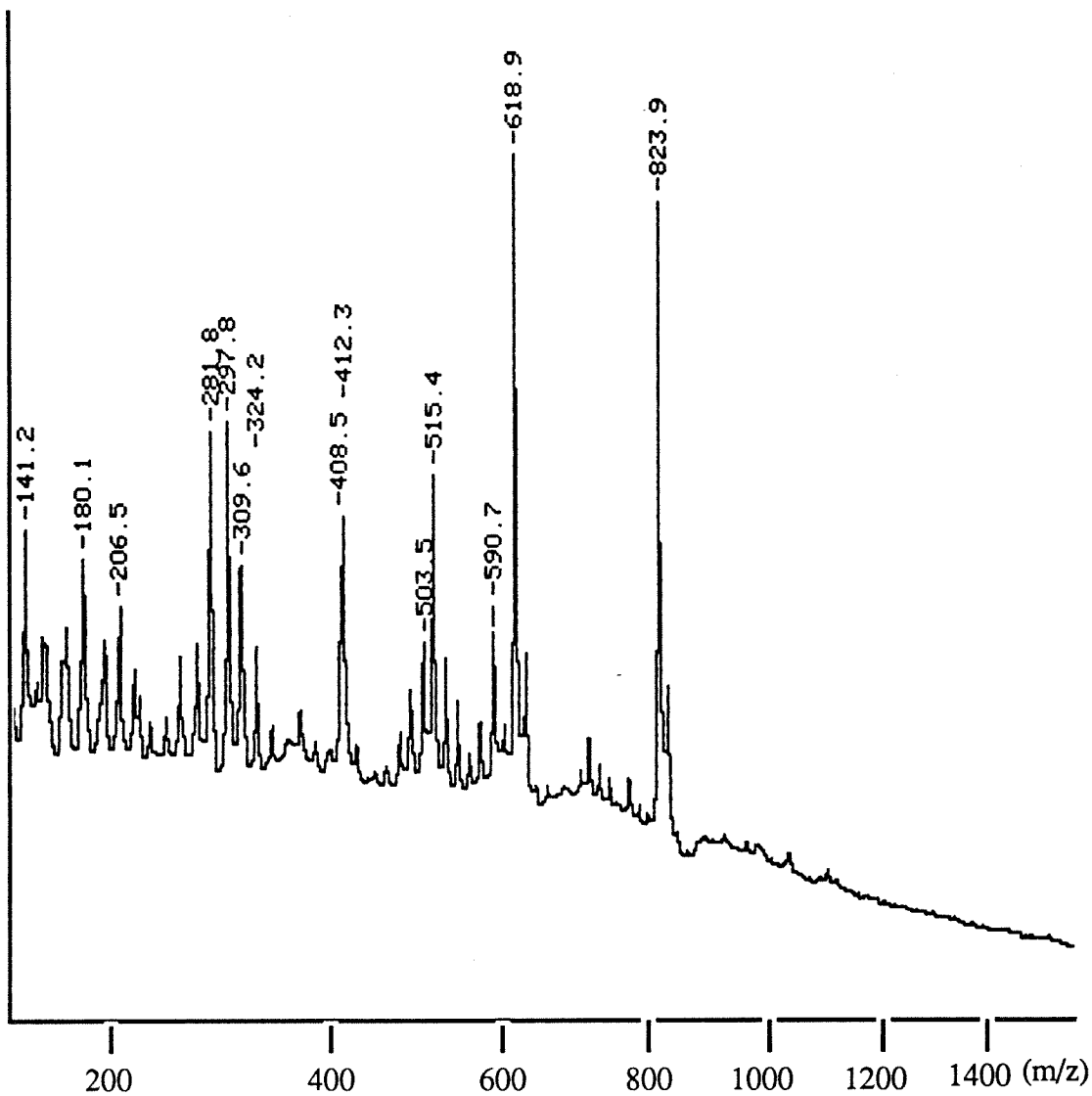


Figure 2.5. ^{252}Cf PDMS spectrum of $[Rh(phi)_2(phen')]^{3+}$.

Figure 2.6. UV-Visible spectra of $[Rh(phi)_2(phen')]^{3+}$ in water as a function of pH. The spectra were measured over a range of 3.0-9.0 pH units.

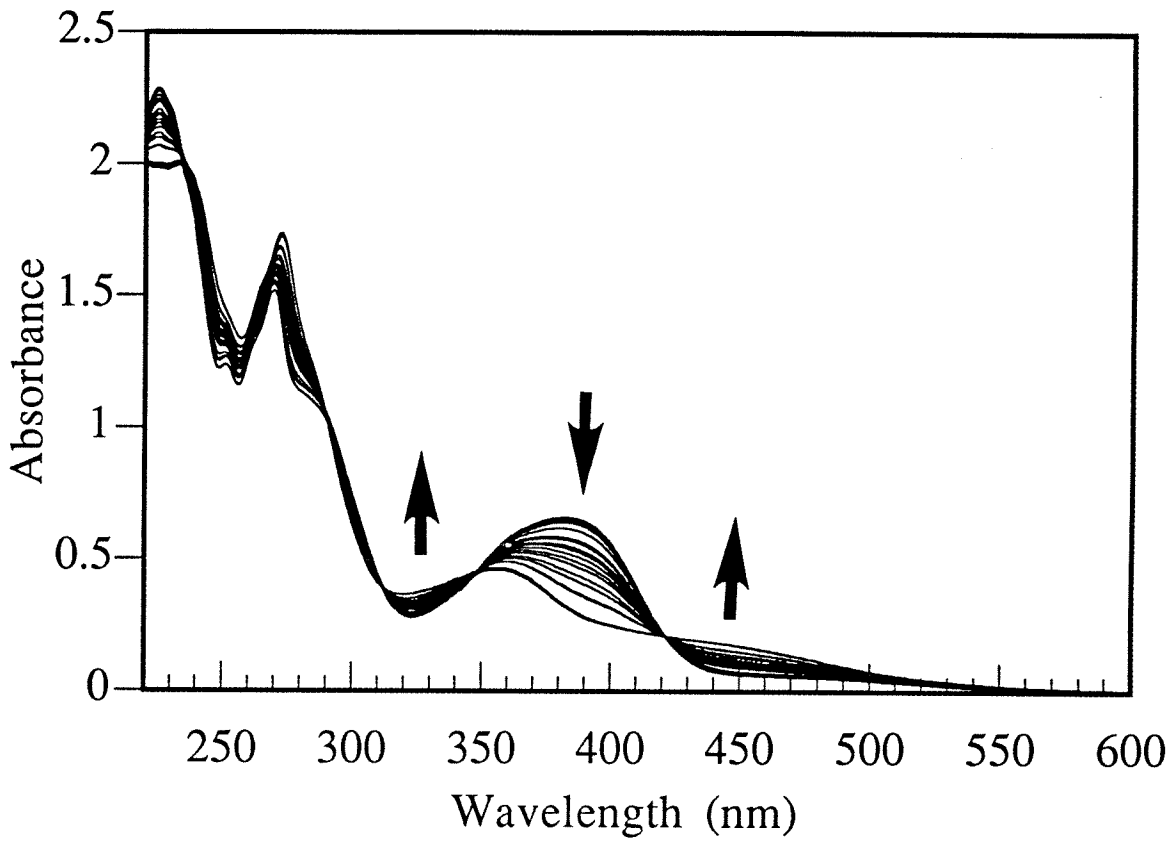
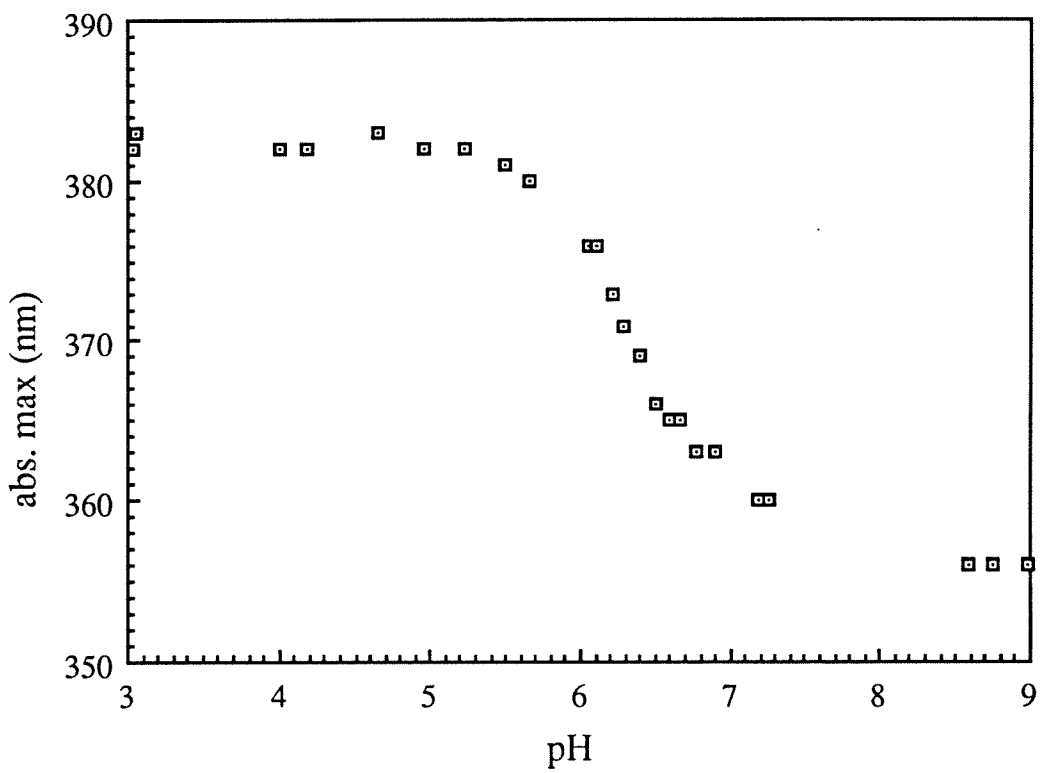

2.5
2
1.5
1
0.5
0
Absorbance
250
300
350
400
450
500
550
600
Wavelength (nm)

Figure 2.7. Plot of λ_{max} (nm) of the phi-transition from figure 2.6 as a function of pH.

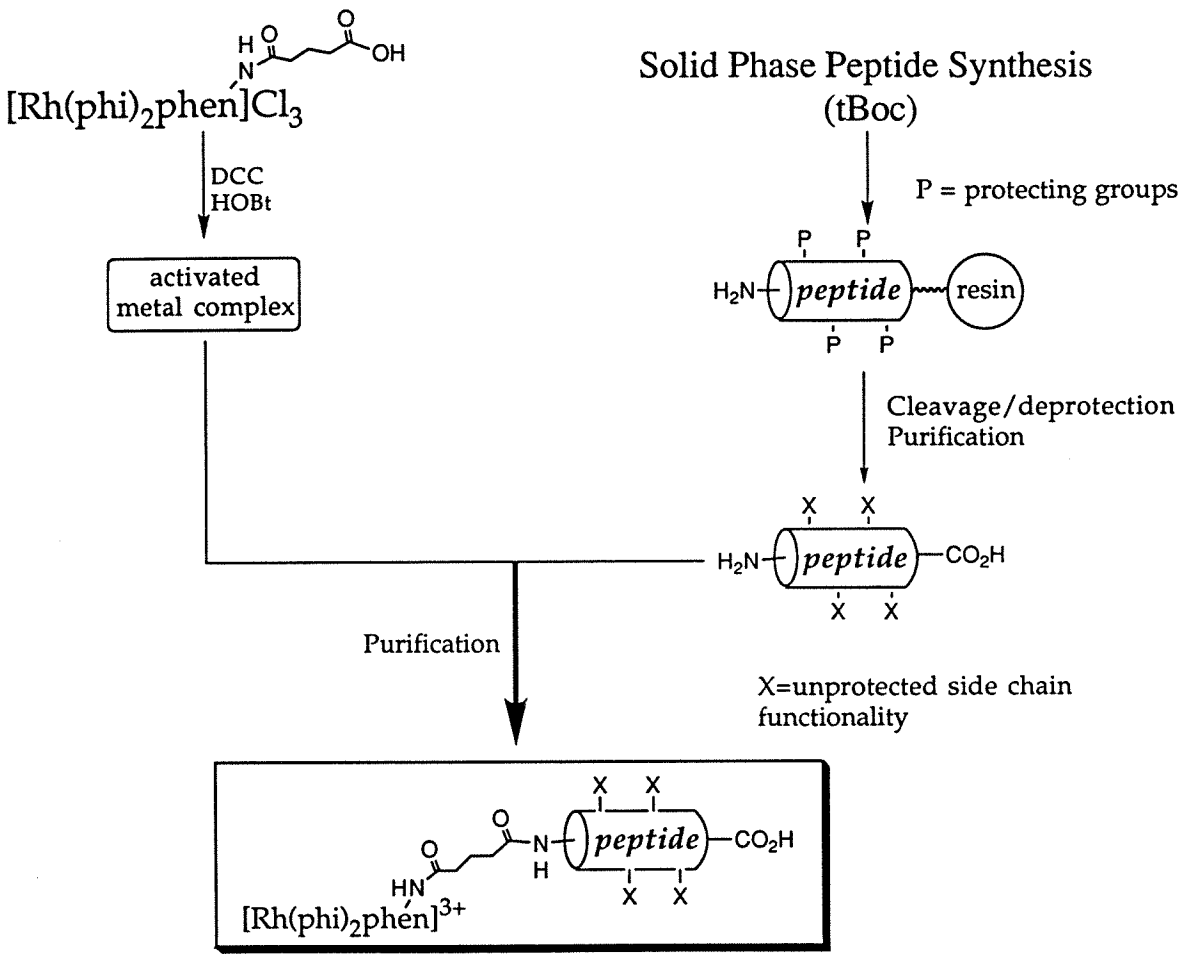

390
380
370
360
350
abs. max (nm)
3
4
5
6
7
8
9
pH

pK_a of $[Rh(phi)_2(bpy)]^{3+}$ is 6.8-6.9, while the loss in intensity is approximately 34% and the λ_{max} for the phi-centered transition is blue shifted by 29.7 nm over a similar pH range.[21] The free carboxylate likely contributes to the lowering of the observed pK_a of $[Rh(phi)_2(phen')]^{3+}$ compared to $[Rh(phi)_2(bpy)]^{3+}$.

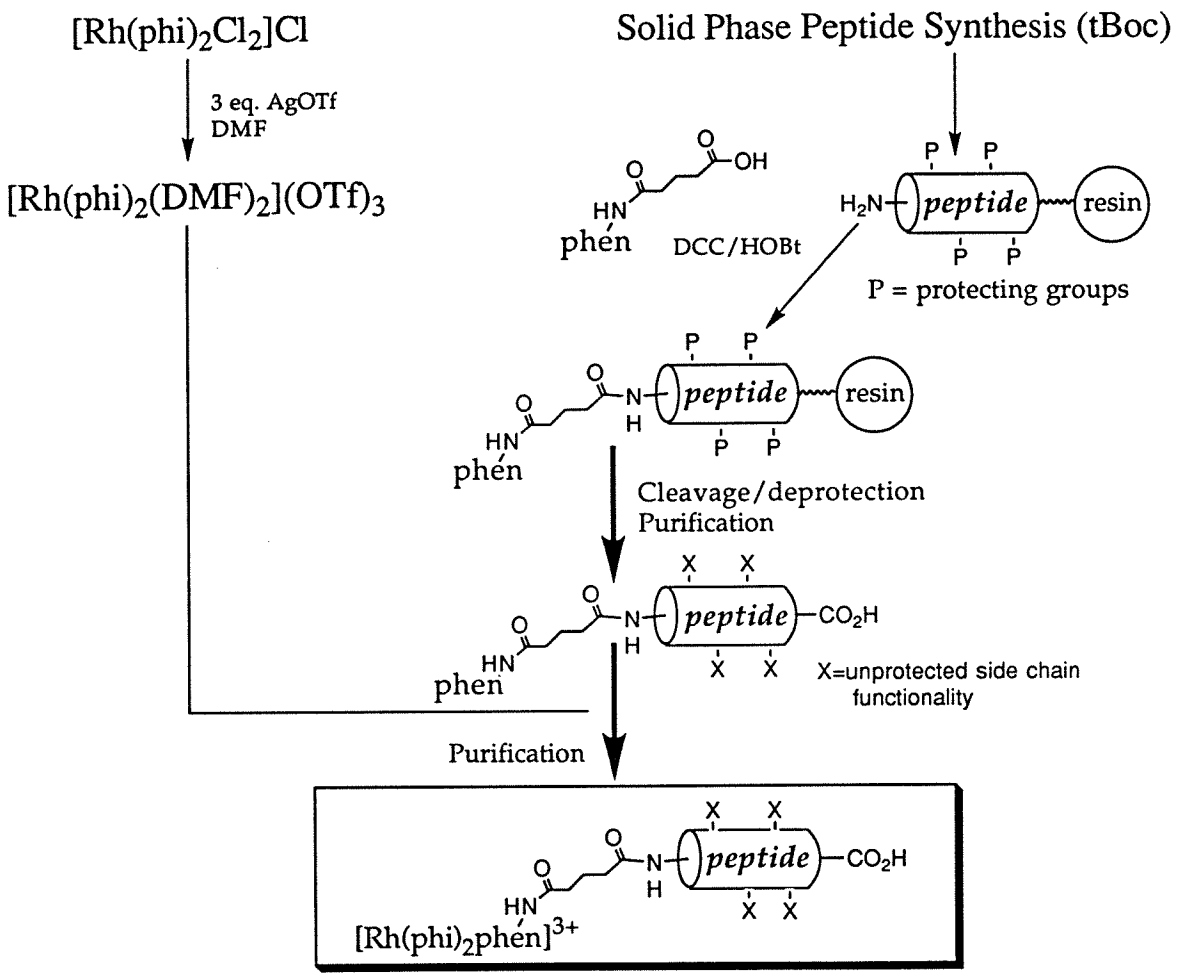
2.3.3. Coupling of the Rhodium Complex to the Peptide. Two general strategies viz., direct coupling and coordination, have been explored in the solution phase (Schemes 2.3 and 2.4) as well as on a solid support (Schemes 2.5 and 2.6).

Solution Phase Methodologies. The earliest attempts at synthesizing metal-peptide complexes involved reaction of the free amino terminus of a purified peptide to the carboxylate on **6**. A range of coupling conditions and reagents including DCC, DIPC, $SOCl_2$, TBTU, TSTU, EDC (water soluble carbodiimide), HOBt, HOSu, and DSC have been explored with limited success. The procedure is of limited utility and works only if the peptide sequence has a unique nucleophilic group such as a single free amine.

A second solution phase method involving coordination of **4** onto a phen'-peptide held some promise because the analogous coordination step in the synthesis of **6** is a high yield reaction (>89%). The coupling of **2** onto the peptide is also a nearly quantitative reaction and is similar to capping the N-terminus during standard solid phase peptide synthesis. The phen'-peptides are deprotected and cleaved from the resin and purified by HPLC. Mass spec and amino acid analysis characterization of the phen'-peptides confirms their identity. The mass spectrum also shows the presence of trace metals including iron associated with the phen'-peptides, possibly by chelating to the attached phen. However, the coordination reaction proves to be disastrous in solution. A large number of peaks corresponding to products containing a rhodium complex and an attached peptide are seen (based on NMR and UV-Vis spectra of collected peaks). These peaks may be attributed to phen' coordinated complexes of peptides of different length, i.e., arising from failure sequences; or to coordination by the peptide. Preliminary



Scheme 2.3. Solution phase strategy for direct coupling.



Scheme 2.4. Solution phase strategy for coordination.

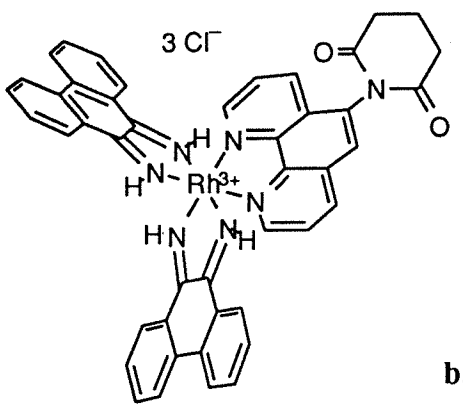
attempts to differentiate between phen' coordination versus peptide coordination by monitoring the formation of a Fe(II)-phen complex in the latter case were inconclusive. The differences in the UV-Vis spectra were too small to draw any conclusions. Furthermore, the characteristic MLCT band for the $[Fe(phen)_3]^{2+}$ (λ_{max} at 510 nm) may be swamped by the absorbance of the rhodium complex. The large number of peaks observed during HPLC purification makes this approach impractical to pursue.

Two side products that repeatedly show up from reaction of **6** were isolated in the direct coupling methods (both solution phase and solid phase) and are shown in Figure 2.8. The first is the glutarimide of **6** arising from the intramolecular cyclization of phen' on the rhodium complex. The second is the N-acetamide of urea observed in reactions with DCC as the coupling reagent. This product was characterized by PDMS and showed the characteristic M^+ signal at 1030 mass units. The yields of these products can be substantial depending on the coupling reagent used.

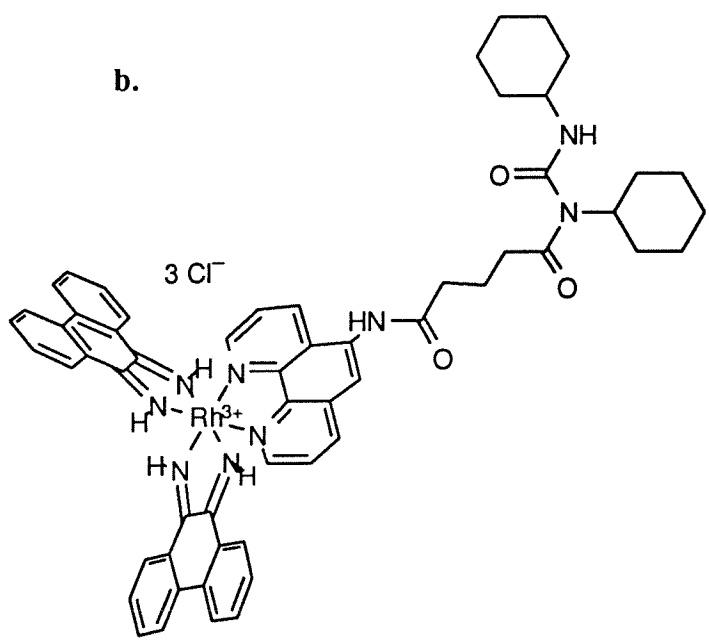
Solid Phase Methodologies. The metal-peptide chimera is synthesized using one of two strategies: (i) the coordination method or (ii) direct coupling. In the coordination strategy (Scheme 2.5), the chelating ligand containing a pendant carboxylate, phen', is first coupled onto the amino terminus of the peptide on the resin. Then, the resin-bound peptide containing the chelating ligand is reacted with $[Rh(phi)_2(DMF)_2](OTf)_3$, in a manner similar to the synthesis of the parent rhodium complex. In the direct coupling strategy (Scheme 2.6), the coordinatively saturated metal complex containing the pendant carboxylate is first assembled. Then the functionalized metal complex and the terminal amine of the peptide bound to the resin are condensed in one step, in a manner that is analogous to the addition of another residue onto the growing peptide chain. For both strategies, following coupling of the metal complex onto the resin, the metal-peptide complex is deprotected and cleaved from the resin in the same manner as for the free peptide.

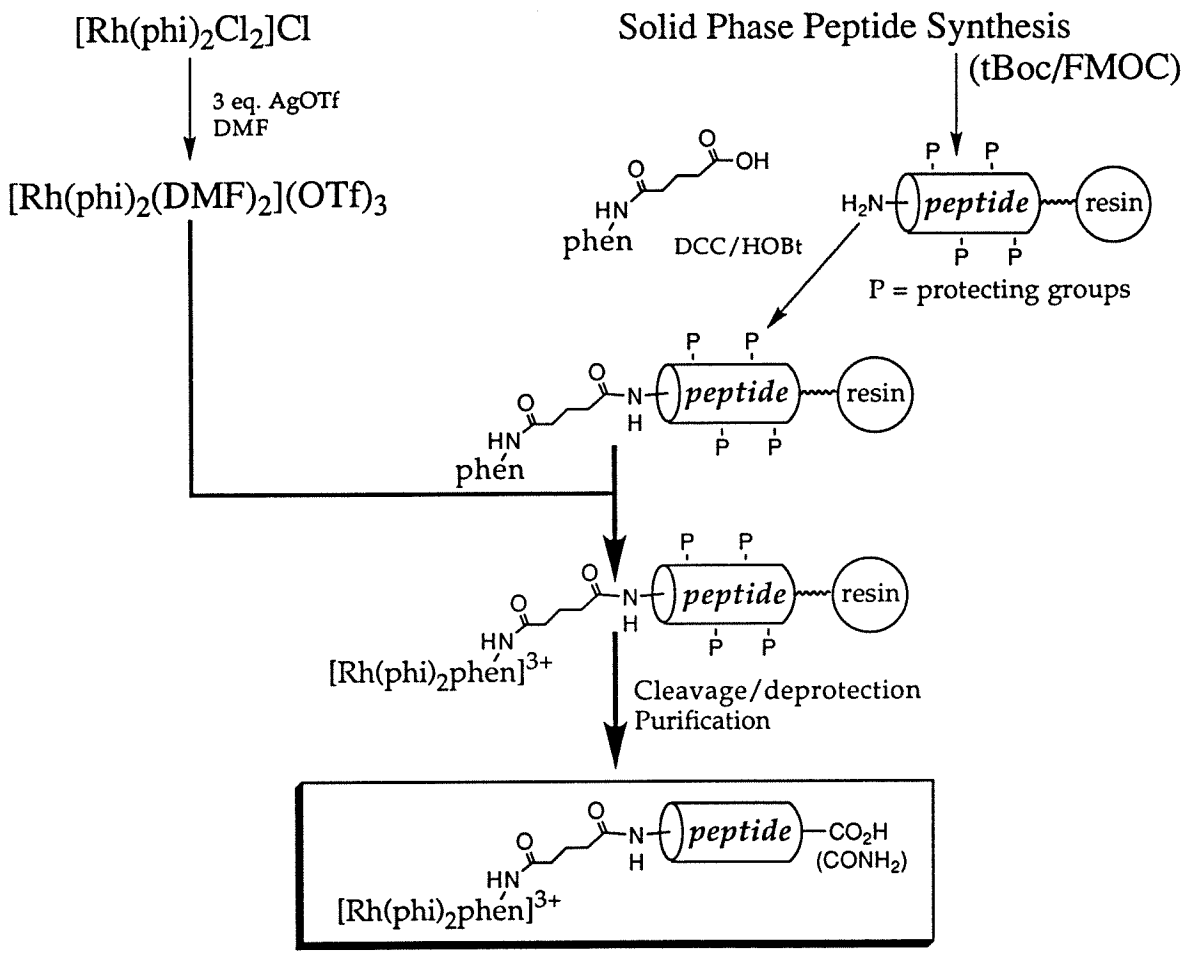
Figure 2.8. Side products observed during the direct coupling of $[Rh(phi)_2(phen')]^{3+}$ to a peptide. **a.** Rhodium complex containing the cyclized form of phen'; **b.** N-acetamide of urea formed when DCC is used as the coupling reagent.

a.

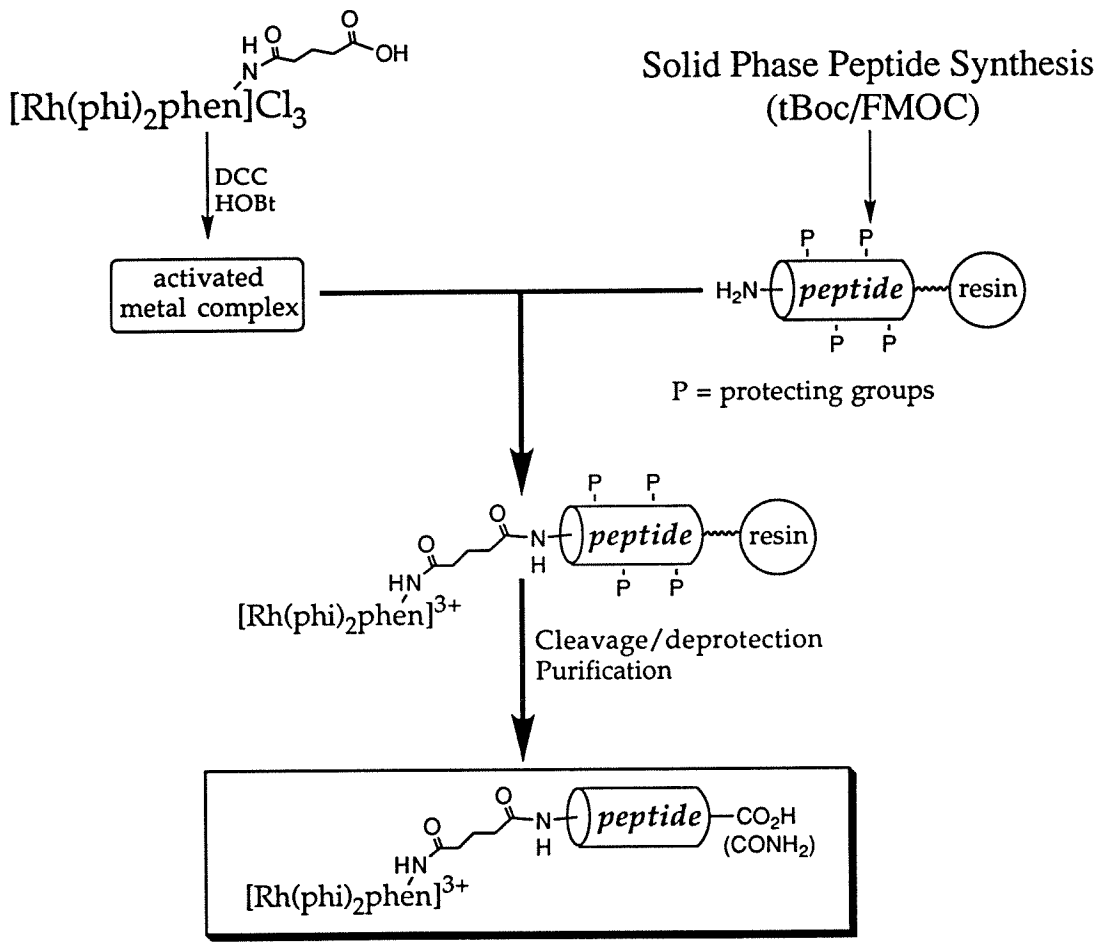


b.





Scheme 2.5. Solid phase strategy for coordination.



Scheme 2.6. Solid phase strategy for direct coupling.

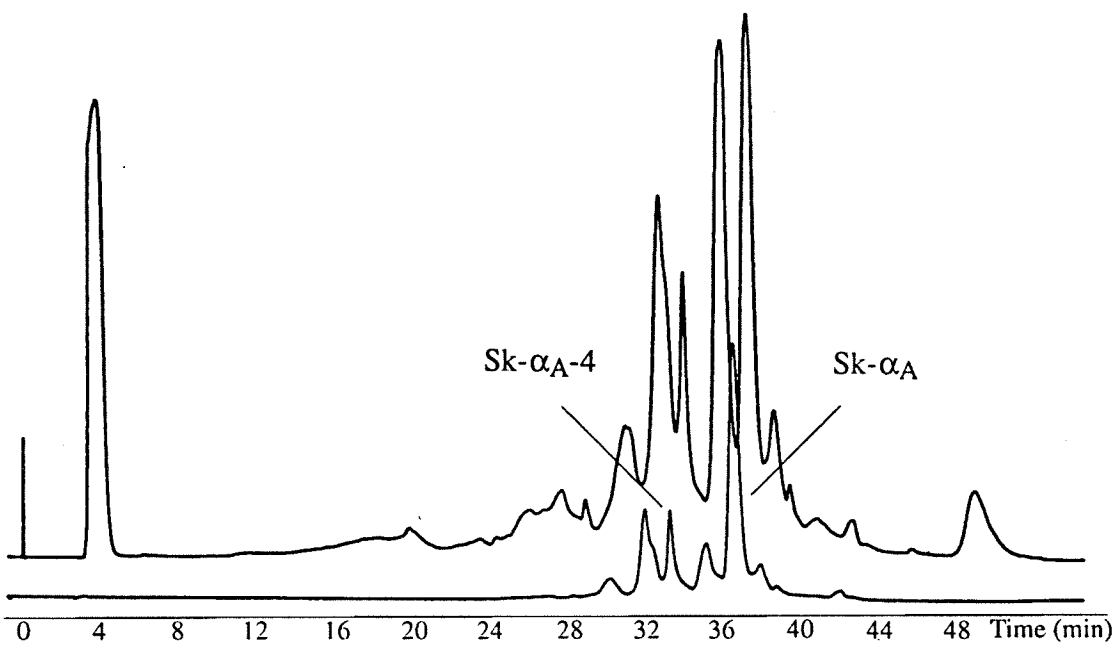
A variety of conditions for synthesis have been examined. Peptides have been constructed using both FMOC and tBoc methodologies and using manual as well as automated solid phase techniques.[10, 11] Prior to coupling of the ligand or metal complex onto the resin, a small portion of the resin is cleaved and the free peptide analyzed by HPLC to check the fidelity and efficiency of the synthesis. Yields of the final metal-peptide chimera (*vide infra)* are limited by the efficiency of the peptide synthesis. Indeed, as shown in Figure 2.9, the FPLC (or HPLC) chromatogram for the purification of the rhodium-peptide complex reflects the homogeniety of the initial peptide synthesis.

A range of coupling agents have been examined using both strategies. In the case of the coordination method, several different coupling reagents have been used with similar success. These reagents include DCC/HOBt, DSC, TBTU, and TSTU. With the direct coupling method, we observe that the presence of the metal center makes the coupling reaction less efficient. The explanation for this poorer reactivity may be a function of either electronic deactivation or steric accessibility on the resin, or likely both. Because of the lower reactivity, more potent coupling agents should be favored. However, we also observe in the case of phen' that intramolecular cyclization can be competitive with very potent activating agents (Figure 2.8). Given these factors, only DCC/HOBt and DSC have been found to promote reaction by the direct coupling strategy.

The metal-peptide complexes are observed to be more difficult to cleave off the resin than the peptide alone. Longer reaction times and harsher conditions (HF versus TFMSA for tBoc peptides) may be needed. The two step TFMSA cleavage procedure works better than the one step procedure used for peptides.[19] Neither the peptide nor the metal complex are stable to extended exposure to the cleavage and deprotection conditions thereby limiting the length of the cleavage reaction. Several linkages to the resins such as MBHA and PAM were also examined, but the variation in linker does not appear to affect the yield of cleaved product.

Figure 2.9. Chromatogram of the semi-preparative scale FPLC purification of $[Rh(phi)_2(phen')]^{3+}$-LQQAIEQLQNAAAA-COOH (Sk-α_A). Upper trace: 220 nm (5.0 AUFS); lower trace: 360 nm (3.0 AUFS). An approximate time scale is shown below the chromatogram. The peaks corresponding to Sk-α_A and a second product, Sk-α_A-4, are indicated. These were collected and re-purified by HPLC to homogeniety. The other peaks correspond to coupling to failed sequences (evident in the 370 nm trace) and analogous uncoupled peptides or uncoordinated phen'-peptides (evident in the 220 nm trace). Note the clear resolution of the desired product peak from the peaks corresponding to the failed peptide sequences. FPLC gradient used is as follows:

Buffer A:	0.1% TFA/water	Buffer B:	0.1% TFA/CH_3CN
Flow rate:	2.5 mL/min	Column:	PEP-RPC 10/10 (C_{18})
0-5 min:	0% B		
25 min:	20% B		
40 min:	40% B		
46 min:	40% B		
51 min:	0% B		
51-60 min:	0% B		


Sk-αA-4
Sk-αA
0
4
8
12
16
20
24
28
32
36
40
44
48
Time (min)

The metal-peptide complexes are purified by HPLC. In all cases observed so far, the retention time of the metal-peptide complexes is found to be longer than that of the corresponding free peptide. The collected peaks are repurified by HPLC to ensure that the metal-peptide complexes have been purified to homogeniety and that the complexes are stable under the HPLC conditions. A list of the metal-peptide chimeras synthesized is shown in Table 2.1 (appendix to Chapter 2).

2.3.4. Overall Yields. Theoretical yields of metal-peptide complex may be determined based upon the initial substitution of the resin. Actual recovered yields of pure chimera (after two rounds of HPLC purification) are found to be in the range of 2 -18%. Not surprisingly, the major determinant of the yield for the reaction is the initial synthesis of the peptide. Yields of metal-peptide chimeras correlate closely with the recovered yields for the individual peptides before metal-complex attachment. The presence of the metal complex, does, however, significantly decrease the overall yield. A broad correlation between the hydrophobicity of a peptide and the yield of the metal-peptide complex may be drawn. Generally, the more hydrophilic peptides tend to give higher yields of the corresponding metal-peptide complex.

Yields do not differ substantially between the coordination and direct coupling strategies. However, the isolated yields are higher for the solid phase methods compared to solution phase synthesis because of lesser undesired products. The presence of the metal center is observed to inhibit the coupling reaction both in solution and on the resin, and the presence of the solid support may somewhat interfere with coordination. The coordination of $[Rh(phi)_2(DMF)_2](OTf)_3$ and phen' in solution is accomplished in 90% yield. The competing intramolecular cyclization reaction does not appear to limit the yield because similar yields are also obtained by using a modified 2,2'-bipyridyl based ligand (bpy') lacking the ability to cyclize.[15] Furthermore, in the solid phase method, the problem of competing reactions can be, in principle, alleviated by multiple coupling

cycles and using excess **6**. A chief difficulty in the synthesis is recovery of metal-peptide chimera from the resin. The presence of the covalently bound metal complex certainly inhibits cleavage of the chimera from the resin; whether electronic or steric factors are dominating is not known.

2.3.5. Characterization of the Chimeras.

Electronic Spectroscopy. As shown in Figure 2.10, the UV-visible spectra of the metal-peptide complexes are observed to be similar to the spectrum of the parent rhodium complex. Since the peptides do not absorb significant light at wavelengths $\geq$ 300 nm, the concentration of the chimera may be quantitated based upon the extinction coefficient of the rhodium complex at 350 nm. As with the parent metal complexes, there are pH dependent changes observed in the spectrum which depend upon the protonation state of the coordinated phi ligands.[13, 21, 22]

NMR Spectroscopy. NMR spectra of products isolated during the HPLC purification of reaction mixtures were helpful in determining which fractions had the intact rhodium complex as well as the peptide. These were then characterized definitively by amino acid analysis and mass spectrometry. The NMR spectrum of a representative metal-peptide complex, Sk-α_A, is shown in Figure 2.11. The aromatic region is sharp and diagnostic of **6**, while the aliphatic region is characteristic of the linker and the attached peptide. The 8 methyl groups in the peptide sequence are present as a broad cluster around 0.8 ppm, α-protons around 4.0 ppm, and the β-methylenes of the peptide side chains as well as the linker protons of phen' between 1.3-2.8 ppm. The broad average signals observed for the various peptide resonances indicates that this metal-peptide complex does not adopt a unique preferred conformation in solution under the NMR conditions. NMR titration of Sk-α_F with increasing amounts of TFE are reported in Chapter 4 along with CD studies of the metal-peptide complexes to ascertain the conformation of the attached peptides.

Figure 2.10. Electronic spectra of a representative metal-peptide complex, $[Rh(phi)_2(phen')]^{3+}$-TQQSKKQLQNKAAA-$CONH_2$ (Sk-α_F), at pH 7.0 (-------) and at pH<5.0 (———).

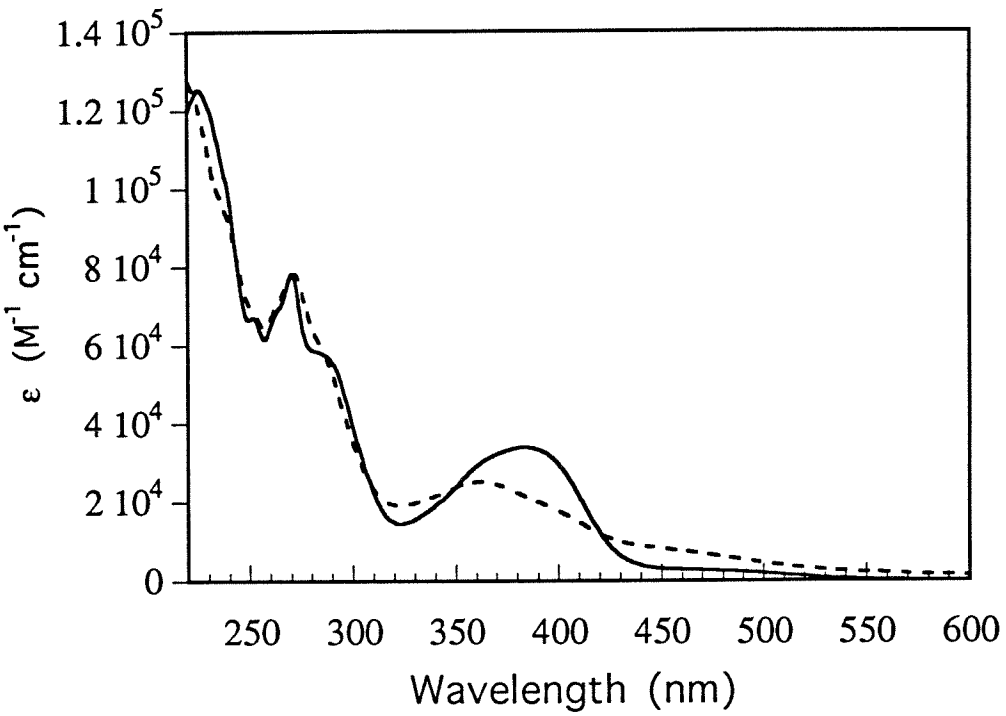
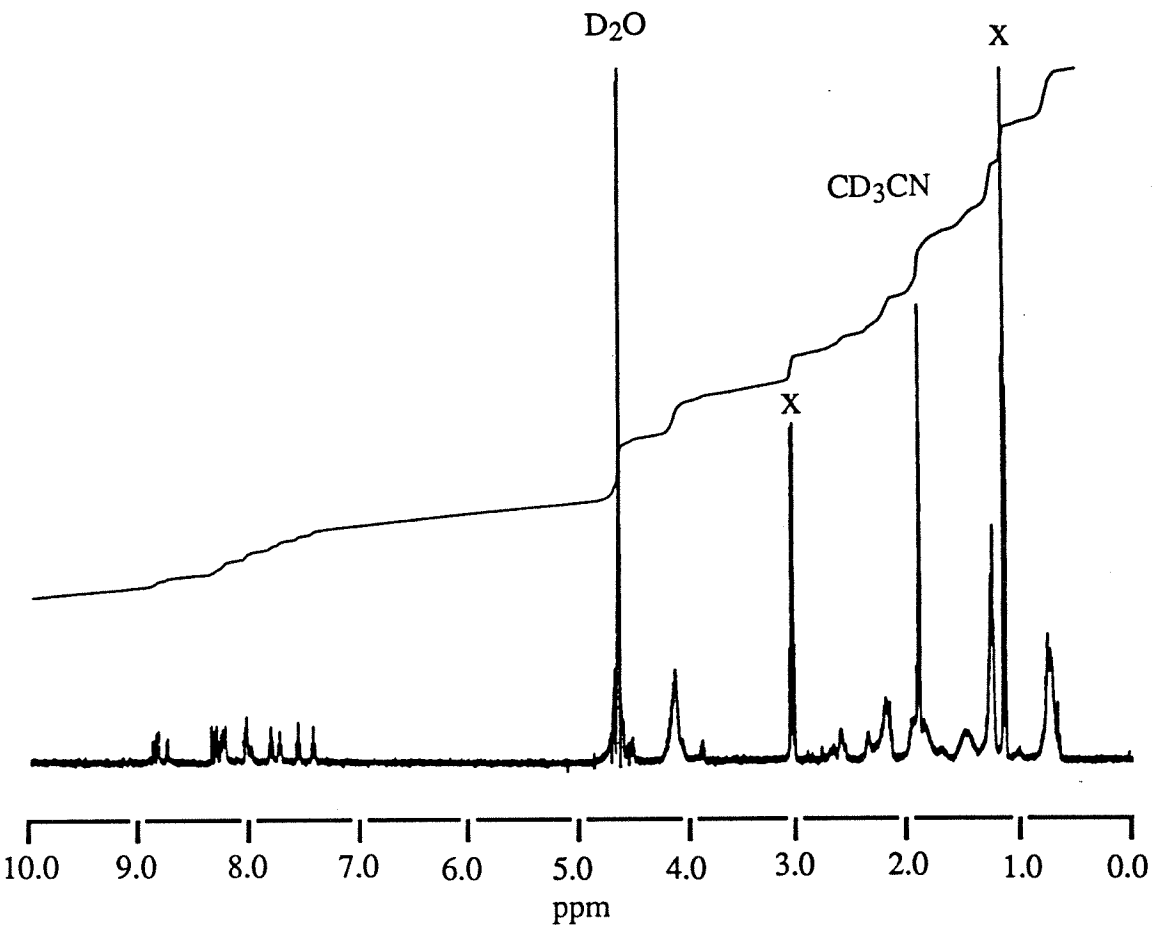

1.4 10^5
1.2 10^5
1 10^5
8 10^4
6 10^4
4 10^4
2 10^4
0
ε (M^-1 cm^-1)
250
300
350
400
450
500
550
600
Wavelength (nm)

Figure 2.11. 500 MHz ^{1}H NMR spectrum of a representative metal-peptide complex, $[Rh(phi)_2(phen')]^{3+}$-LQQAIEQLQNAAAA-COOH (Sk-α_A), in D_2O (0.1% TFA, 20% CD_3CN).


D2O
X
CD3CN
X
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm

Amino Acid Analysis. The peptide content of the chimera is ascertained by quantitative amino acid analysis. Products from the hydrolysis of **6** (as a control) were determined by comparison between the following solutions:

i. **6** alone

ii. **6** + amino acids standard

iii. amino acids standard

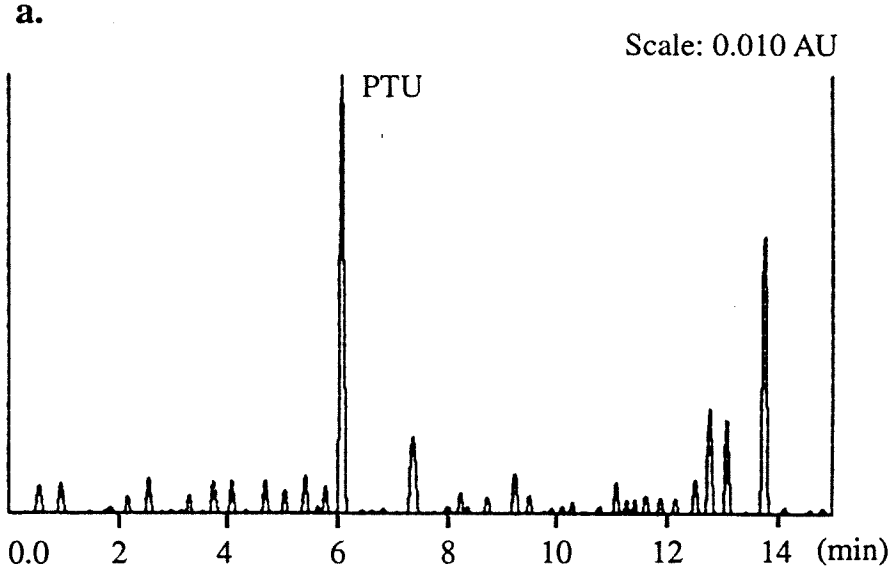
iv. hydrolysis blank of the amino acid analyzer

Hydrolysis of **6** and, subsequent separation of the products by analytical HPLC in the amino acids analyzer, shows two large peaks at 11.26 and 13.77 min (Figure 2.12) and a peak with a retention time of 9.23 min which is close to that for valine (9.25 min). The peak at 13.77 min is also observed in the hydrolysis blank. However, the peak at 11.26 min is neither seen in the hydrolysis blank nor in the amino acids standard indicating that this is a product originating from **6**. The peak at 11.26 min is observed even when **6** is hydrolyzed in the presence of an amino acids standard. This further substantiates the idea that the peak at 11.26 min is not a trace amino acid with an anomalous retention time or an artifact of the hydrolysis procedure. Similarly, quantitative amino acid analysis of the amino acids standard (0.25 nmole) in the presence of **6** yielded the respective amino acids in the correct ratios (compared to nor-leucine as an internal standard) indicating that **6** does not interfere with the amino acid analysis either.

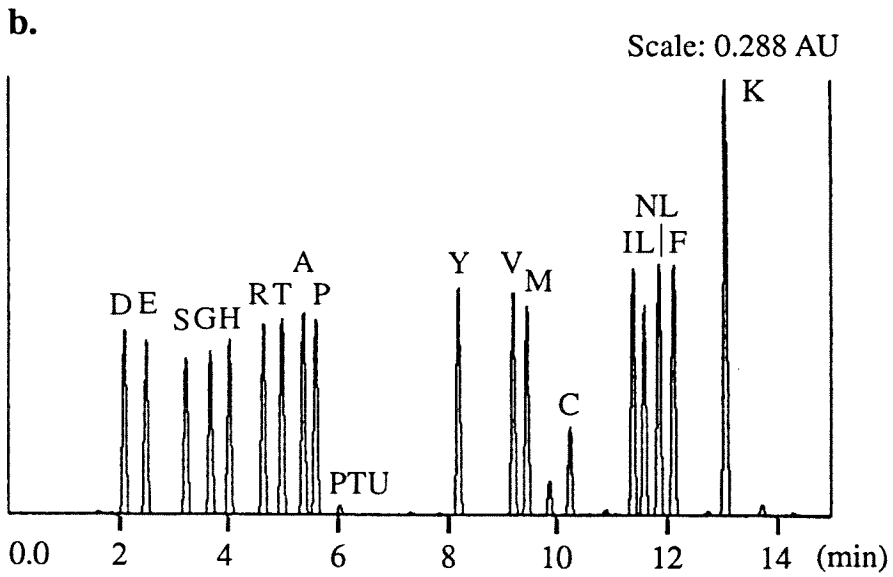
Since hydrolysis products of **6** have an elution profile distinct from normal amino acids (with the possible exception of Val), quantitative amino acid analysis may be used as an independent measure for the concentration of the metal-peptide complexes. The ratios obtained for amino acids match up well with those expected for the attached peptide. Comparison of the quantitation by UV-visible spectroscopy and amino acid analysis reveals a near 1:1 ratio of rhodium complex to peptide in all cases.

Figure 2.12. Amino acid analysis control for $[Rh(phi)_2(phen')]^{3+}$. **(a)** hydrolysis blank for the amino acid analyzer (no amino acids loaded); **(b)** hydrolysis of amino acids standard; **(c)** hydrolysis of $[Rh(phi)_2(phen')]^{3+}$; and **(d)** hydrolysis of a mixture of amino acids standard and $[Rh(phi)_2(phen')]^{3+}$. The peak P1 (at 11.26 min) corresponds to a product originating from $[Rh(phi)_2(phen')]^{3+}$. Peak P2 (at 13.77 min) is present in all samples and is a product of the hydrolysis blank. Nor-leucine (NL) is used as an internal standard and phenylthio urea (PTU) is a product of the derivatization procedure for the amino acids prior to detection.

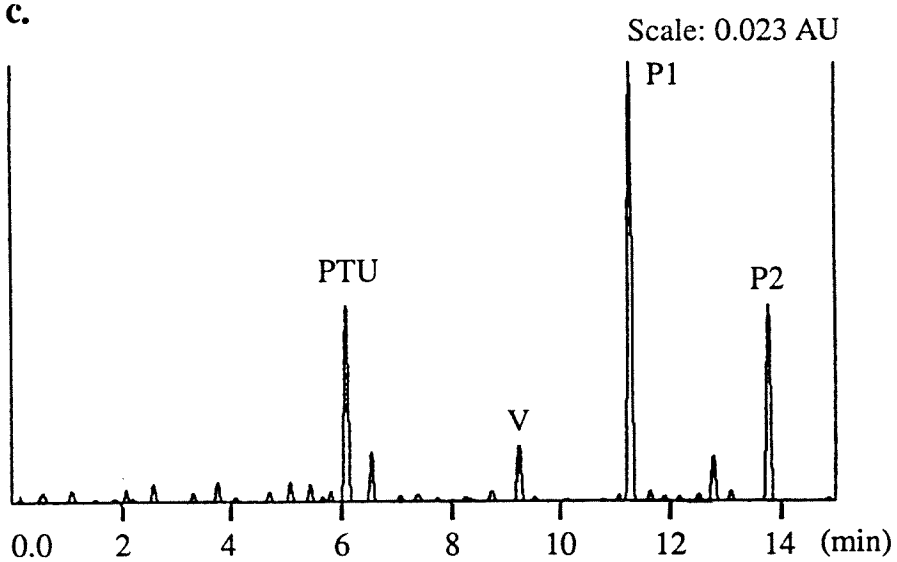
a.


Scale: 0.010 AU
PTU
0.0
2
4
6
8
10
12
14
(min)

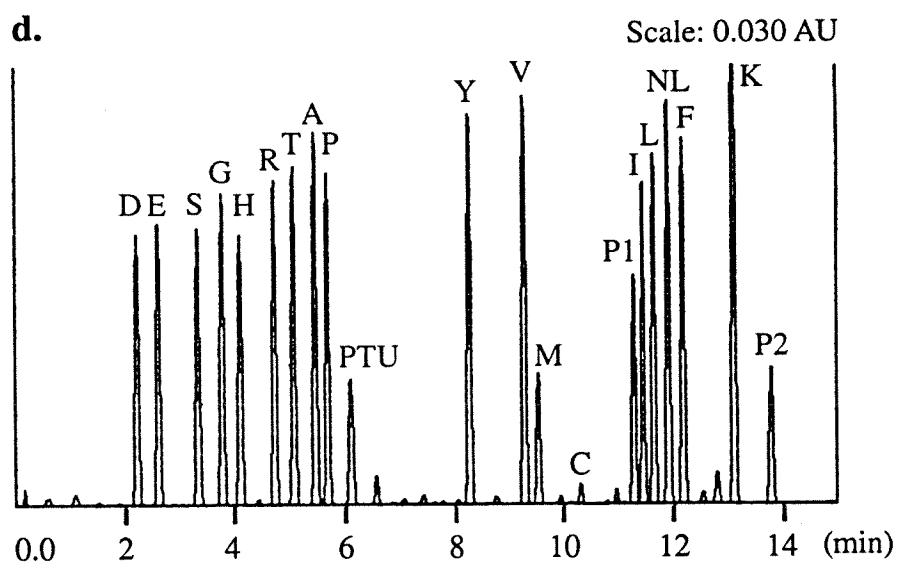
b.


Scale: 0.288 AU
K
NL
IL|F
Y
V
M
A
D E
S G H
R T
P
C
PTU
0.0
2
4
6
8
10
12
14
(min)

c.


Scale: 0.023 AU
P1
PTU
P2
V
0.0
2
4
6
8
10
12
14
(min)

d.


Scale: 0.030 AU
D E S G H R T A P
PTU
Y
V
M
C
P1
I L NL F K
P2
0.0
2
4
6
8
10
12
14
(min)

Considerations of Chirality. The enantiomers of $[Rh(phi)_2(phen')]^{3+}$ are resolved on a cation exchange column using a chiral eluent (+) *tris*[*l*-cysteinesulphinato(2-)-S,N]cobaltate(III).[16] As may be seen in Figure 2.13, the isomers separate into two bands which show characteristic CD spectra. The assignment of enantiomeric configuration is made based upon comparison to spectra of $[Rh(phen)_2(phi)]^{3+}$ and $[Rh(en)_2(phi)]^{3+}$ (en=ethylenediamine).[23] Upon coupling a peptide to **6**, two diastereomers arise. These would correspond to coupling Λ- or Δ- enantiomers of **6** to a peptide composed entirely of l-amino acids. Pure diastereomers may then be obtained by direct coupling of the enantiomers of the rhodium complex to the peptide. Conditions for coupling and deprotection/cleavage from the resin do not lead to racemization at the metal center.[17]

Alternatively, optically pure metal-peptide diastereomers are isolated from chimeras which are racemic about the metal center using a protocol of chiral elution on Sephadex CM-C25 analogous to that used for the parent metal complex.[16] The CD spectrum of the diastereomerically pure metal-peptide complex (Figure 2.14) isolated by chiral elution is observed to be the composite of the spectrum of the corresponding enantiomer of the rhodium complex and that of the peptide.

The possibility of chiral induction about the metal center either in the direct coupling or coordination of metal complexes to the diastereomerically pure peptides on the resin was also considered. However, the synthesis of the metal-peptide complexes is found to proceed without diastereoselectivity using either strategy. Figure 2.15 shows the CD spectrum of a metal-peptide chimera produced by coordination of racemic $[Rh(phi)_2]^{3+}$ onto a 14-residue peptide containing coupled phen'. The circular dichroism shows negative ellipticity below 240 nm which is characteristic of the peptide, but no features above 250 nm are evident, which is indicative of a racemate at the rhodium center. This result is to be contrasted to the results of crosslinking peptide α-helical bundles in solution by $[Fe(bpy)_3]^{2+}$ or $[Ni(bpy)_3]^{2+}$ where isomeric induction was observed.[1]

Figure 2.13. Circular dichroism of the enantiomers of $[Rh(phi)_2(phen')]Cl_3$ in 10 mM Tris·HCl, pH 7.0 : (-------) Λ-enantiomer and (———) Δ-enantiomer (Spectra adapted from reference 17).

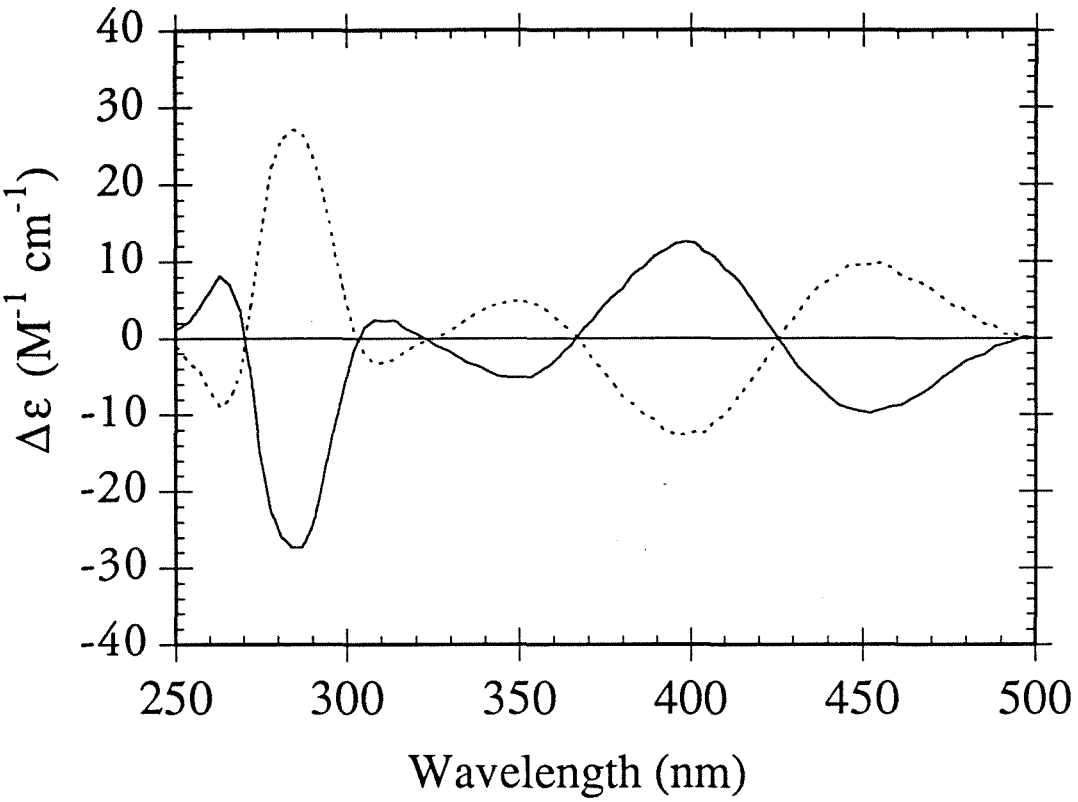

40
30
20
10
0
-10
-20
-30
-40
Δε (M-1 cm-1)
250
300
350
400
450
500
Wavelength (nm)

Figure 2.14. Circular dichroism of the diastereomers of $[Rh(phi)_2(phen')]^{3+}$-TQQSKKQLQNKAAA-$CONH_2$ in 10 mM Tris·HCl, pH 7.0 : (-------) Λ-diastereomer and (———) Δ-diastereomer.

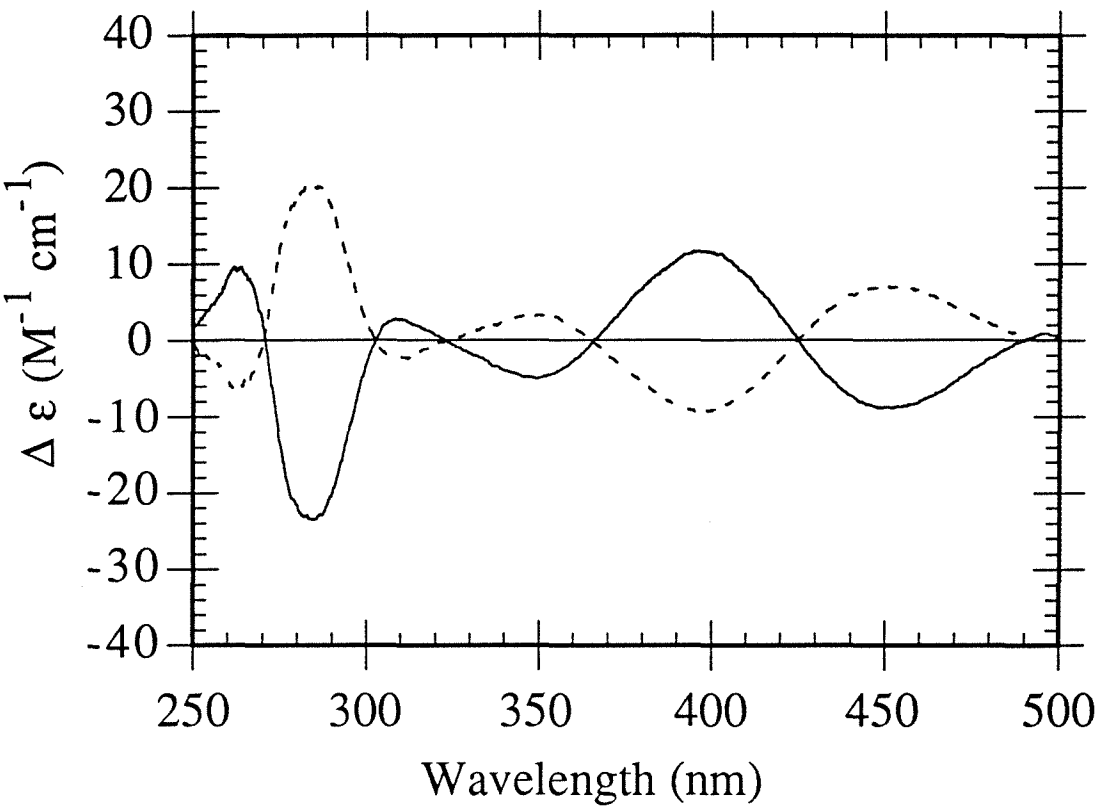
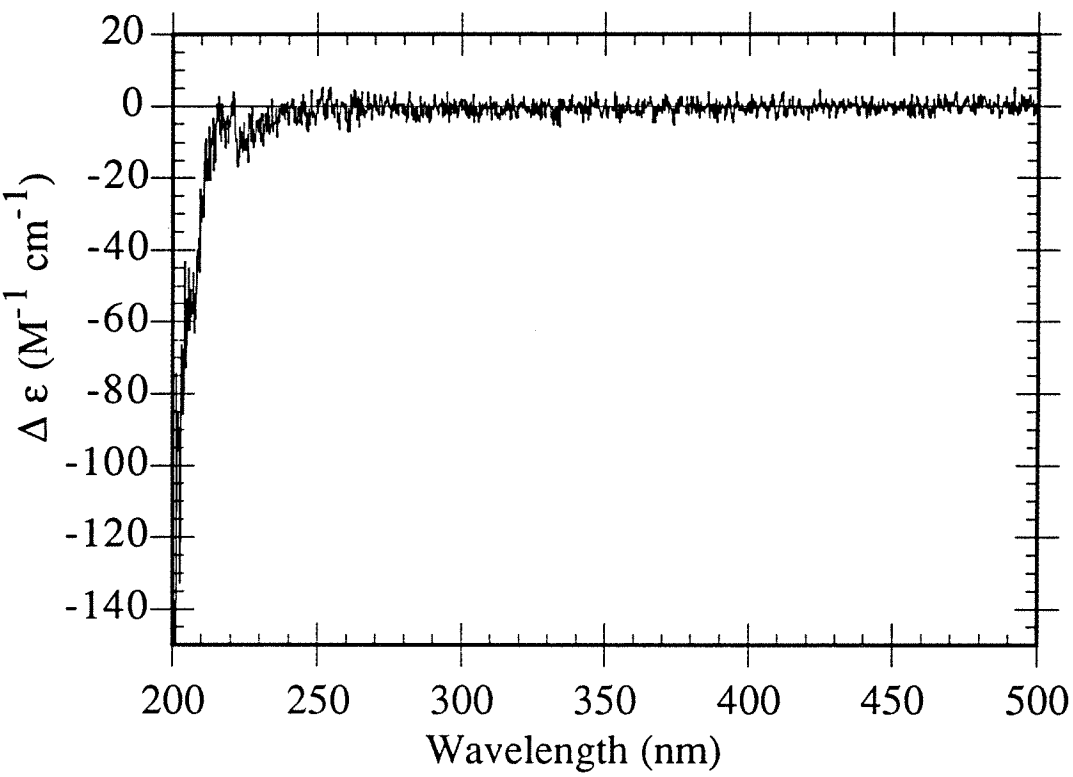

40
30
20
10
0
-10
-20
-30
-40
Δ ε (M-1 cm-1)
250
300
350
400
450
500
Wavelength (nm)

Figure 2.15. Circular dichroism of racemic $[Rh(phi)_2(phen')]^{3+}$-TQQSKKQLQNKAAA-$CONH_2$ in 20 mM Tris·HCl, pH 7.0.


20
0
-20
-40
-60
-80
-100
-120
-140
Δ ε (M-1 cm-1)
200
250
300
350
400
450
500
Wavelength (nm)

2.3.6. Mass Spectrometry.

Characterization of the Metal-Peptide Complexes. The PDMS spectrum of **6** is shown in Figure 2.4. The molecular ion peak is observed at 823.9 (824.7) and corresponds to $[Rh(phi)_2(phen')]^+$. Characteristic peaks corresponding to $[Rh(phi)(phen')]^+$ at 618.9 (618.5), $[Rh(phi)_2]^+$ at 515.4 (515.4), $[Rh(phen')]^+$ or $[Rh(phi)_2(phen')]^{2+}$ at 412.3 (412.2) and $[Rh(phi)]^+$ at 309.6 (309.2) as well as a series of peaks corresponding to sequential degradation of the linker from the phen' containing fragments are observed.

The mass spectrometry of the chimera establishes that the metal complex and peptide are covalently bound. As is evident in Figure 2.16, the PDMS spectrum of $[Rh(phi)_2(phen')]^{3+}$-LQQAIEQLQNAAAA-COOH shows the expected molecular ion peak at 2274.7 (2275.3), $[Rh(phi)_2]^+$ at 515.4 and $[Rh(phi)]^+$ at 309.2.

Importantly, the presence of the covalently attached metal complex promotes fragmentation of the attached peptide. As shown in Figure 2.16, a series of A_n fragments is observed which reflects cleavage of the C_α-CO bonds in the metal-peptide chimera. The calculated mass of the observed fragments is shown in Figure 2.17. In contrast, no fragmentation is evident for the peptide lacking the metal complex. In addition, for each A_n fragment a corresponding fragment of 206 lower mass/charge is observed. These pairs of fragments correspond to A_n fragments with the intact rhodium complex attached and A_n fragments containing metal complexes in which one of the phi ligands has been lost. For metal-peptide chimeras containing $\leq$ 14 residues, the complete series of A_n fragments is observed.

As with **6**, it is also noteworthy that the molecular ion peak corresponds to the singly charged species. This feature is observed even with highly positive charged metal-peptide complexes. The doubly or triply charged ion peaks are small or not observable. The $[Rh(phi)_2(phen'\text{-peptide})]^+$ fragments are consistently smaller than the expected mass by 2 or 3 mass units. The free phi ligand is also seen at its calculated mass of 206.

Figure 2.16. ^{252}Cf PDMS spectra of a representative metal-peptide conplex, $[Rh(phi)_2(phen')]^{3+}$-LQQAIEQLQNAAAA-COOH. PDMS spectrum of the metal-peptide complex (top) showing the molecular ion peak (M^+), fragments characteristic of cleavage at the metal center and A_n fragments arising from cleavage of the peptide C_α-CO bond. The inset above shows an enlarged view of the 800-1950 region with A_1-A_{10} fragments labeled. The PDMS spectrum of the free peptide, H_2N-LQQAIEQLQNAAAA-COOH, showing the M^+ peak and no other significant fragmentation is shown below.

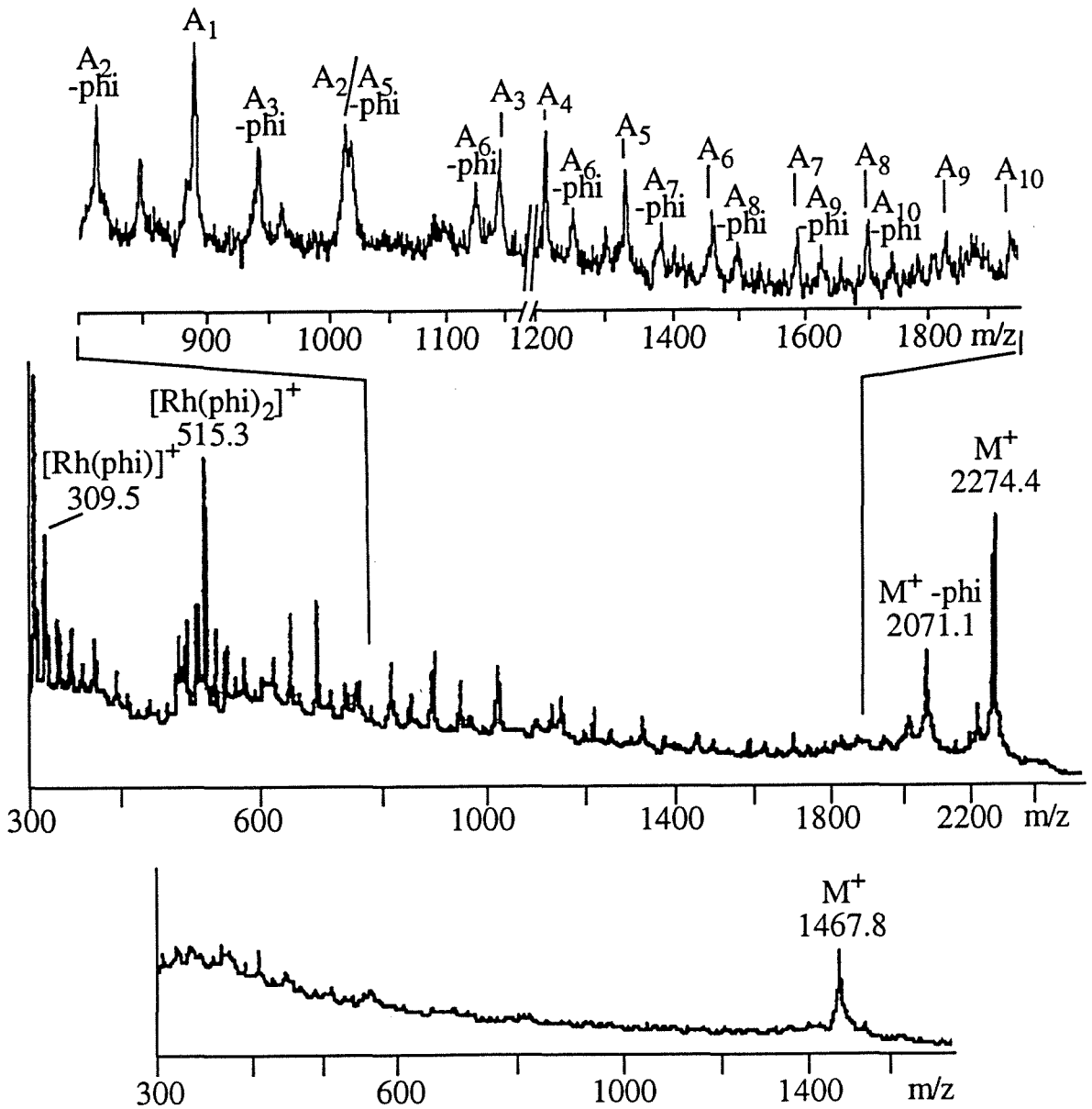
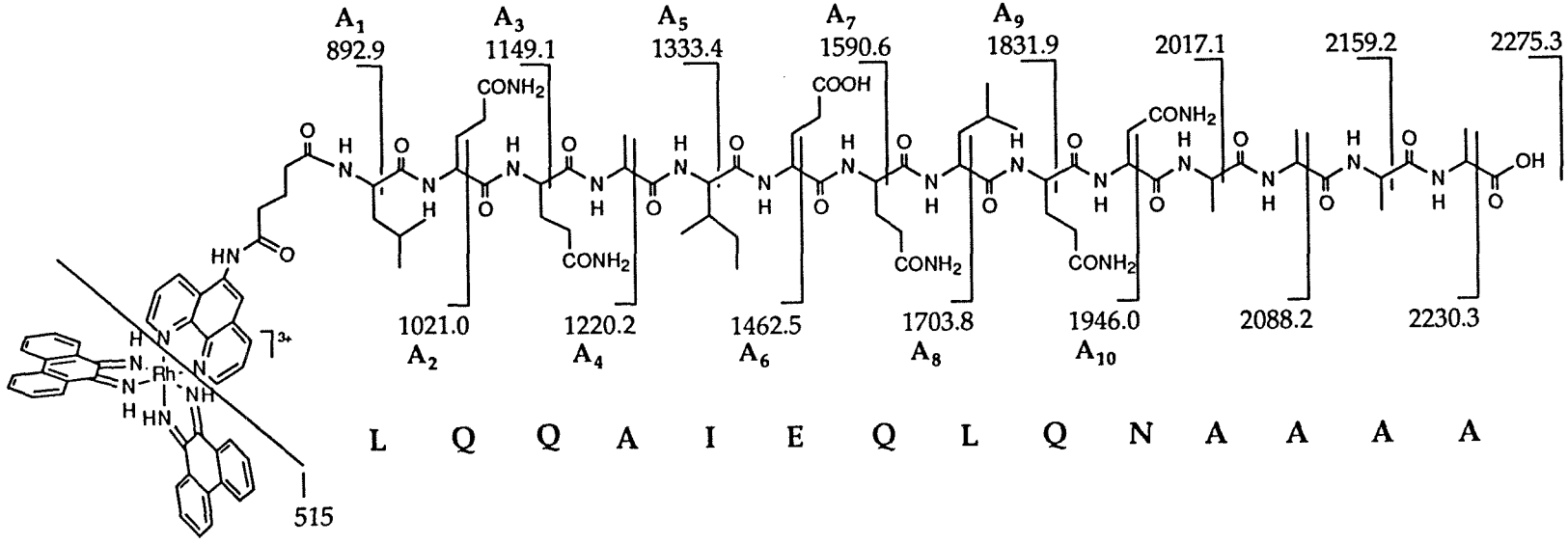

A1
A2 -phi
A3 -phi
A2/A5 -phi
A3
A4
A6 -phi
A6 -phi
A5
A7 -phi
A6
A8 -phi
A7
A9 -phi
A8
A10 -phi
A9
A10
900
1000
1100
1200
1400
1600
1800 m/z
[Rh(phi)2]+ 515.3
[Rh(phi)]+ 309.5
M+ 2274.4
M+ -phi 2071.1
300
600
1000
1400
1800
2200 m/z
M+ 1467.8
300
600
1000
1400
m/z

Figure 2.17. Schematic representation of the A_n fragments for the metal-peptide complex and their calculated mass. The observed mass seen in figure 2.16 is 1-3 mass units lower than the calculated mass.

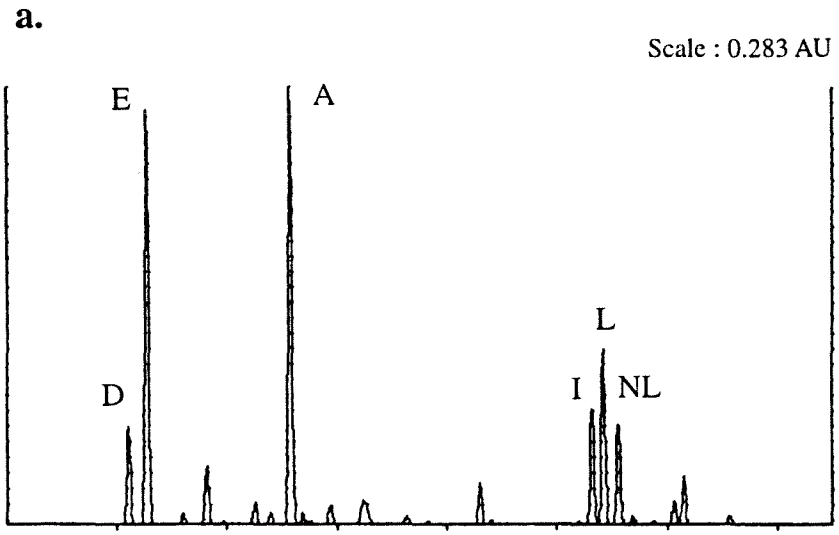

A1
892.9
A3
1149.1
A5
1333.4
A7
1590.6
A9
1831.9
2017.1
2159.2
2275.3
1021.0
A2
1220.2
A4
1462.5
A6
1703.8
A8
1946.0
A10
2088.2
2230.3
L Q Q A I E Q L Q N A A A A
515

Illustrations of Utility in Determination of Sequence Information by Mass Spectrometry. The HPLC elution profile for the purification of Sk-α_A indicated a major product and a second product (figure 2.9), both of which showed the characteristic UV-Visible spectrum for a metal-peptide complex. The amino acid analyses of the two HPLC peaks, Sk-α_A-4 and Sk-α_A, are shown in Figure 2.18. Sk-α_A has the expected amino acid composition while Sk-α_A-4 is missing 1 Glx residue. However, the sequence has 1 Glu and 4 Gln residues, any of which may have failed to couple during the peptide synthesis. The PDMS spectra of the two products are shown in Figures 2.19 and 2.20. Sk-α_A shows the expected molecular ion peak and the characteristic fragmentation pattern. Sk-α_A-4, in contrast, has a M^+ peak at 2145.6 which is 129.1 smaller than the M^+ peak observed for Sk-α_A. This may correspond to a deletion of Glu (129) or any of the four Gln (128) residues. However, comparison of the fragments produced in the two products reveals that the mass spec peaks match up well until 1329 and then are different. This indicates that the Sk-LQQAI piece and smaller fragments are identical, while the larger fragments differ because E^6 failed to couple some of the time during the initial peptide synthesis leading to an E^6 deletion mutant. Such specific sequence information is only available unambiguously because of the unique fragmentation pattern observed for the metal-peptide complexes.

In another series, two HPLC products were isolated following the coupling of α_C to **6**. The two products, Sk-α_C and Sk-α_C-9, have different HPLC and CE retention times but identical amino acid analyses (Figure 2.21). The PDMS spectrum, however, indicates that the M^+ ions for the two products differ by 152 mass units (Figures 2.22 and 2.23). Sk-α_C has the expected mass 2301.1 (calc. 2301.8), Sk-α_C-9 has an M^+ of 2453.0 and the difference of 152 suggests the presence of a Tos protecting group on Arg (calc. difference is 153). Again the two products exhibit an identical PDMS fragmentation for Sk-LQQAIE and smaller pieces, but fragments corresponding to Sk-LQQAIER and longer sequences differ by 153. The break in the fragmentation pattern corresponds

Figure 2.18. Amino acid analysis of **(a)** $[Rh(phi)_2(phen')]^{3+}$-LQQAIEQLQNAAAA-COOH (Sk-α_A) and **(b)** $[Rh(phi)_2(phen')]^{3+}$-LQQAI-QLQNAAAA-COOH (Sk-α_A-4). The expected ratios for amino acids in Sk-α_A are: Asx (1); Glx (5); Ala (5); Ile (1); Leu (2). The amino acid analysis data for Sk-α_A-4 indicates a deletion of 1 Glx residue. Nor-leucine (NL) is used as an internal standard during peptide hydrolysis for quantitation of the amino acids.

a.



b.

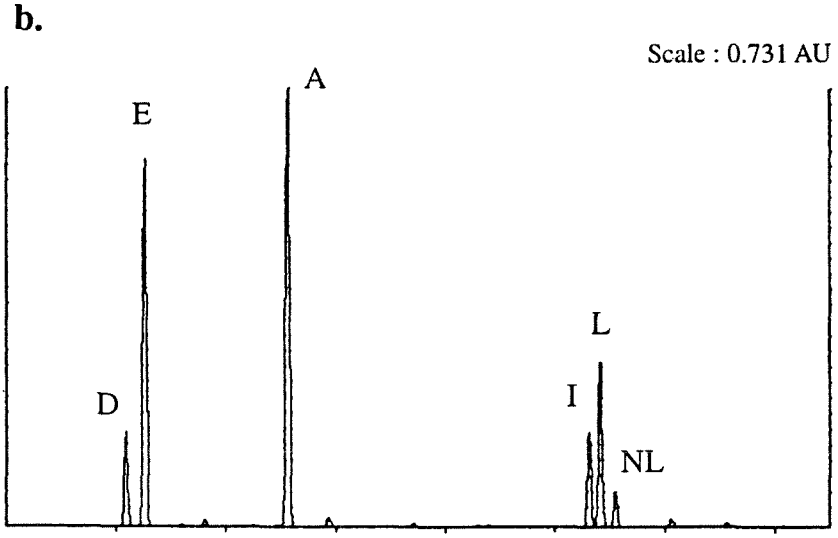
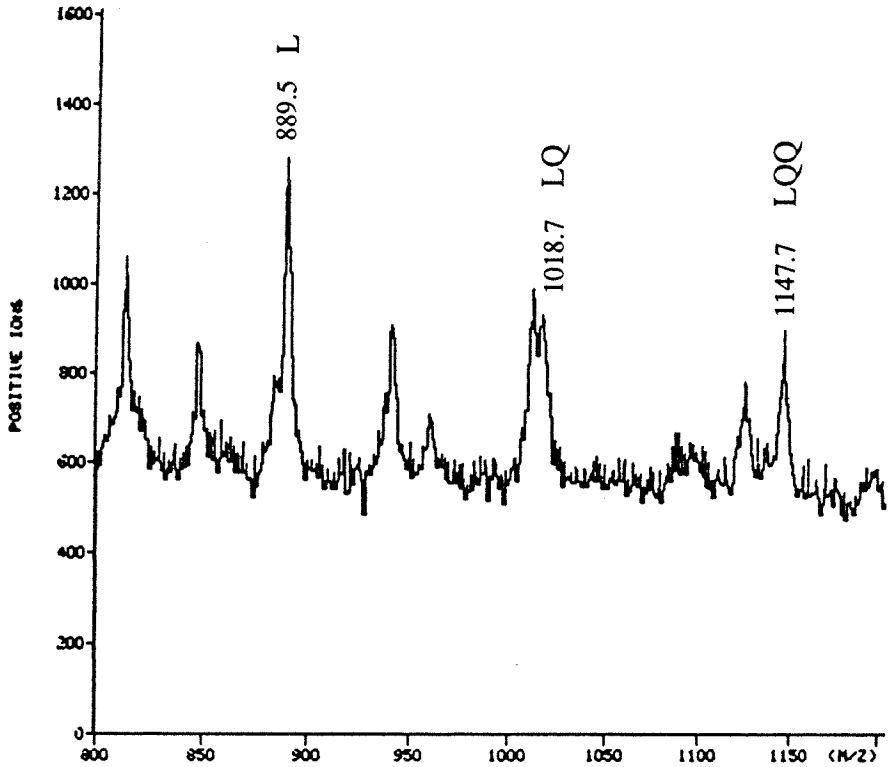


Figure 2.19. ^{252}Cf PDMS of $[Rh(phi)_2(phen')]^{3+}$-LQQAIEQLQNAAAA-COOH (Sk-α_A) showing the fragmentation pattern in detail. Only the fragments containing intact $[Rh(phi)_2(phen')]^{3+}$ have been indicated. The unlabeled peaks correspond to the fragments that have lost a phi ligand. The complete spectrum is shown in figure 2.16.


889.5 L
1018.7 LQ
1147.7 LQQ
POSITIVE IONS
(M/Z)

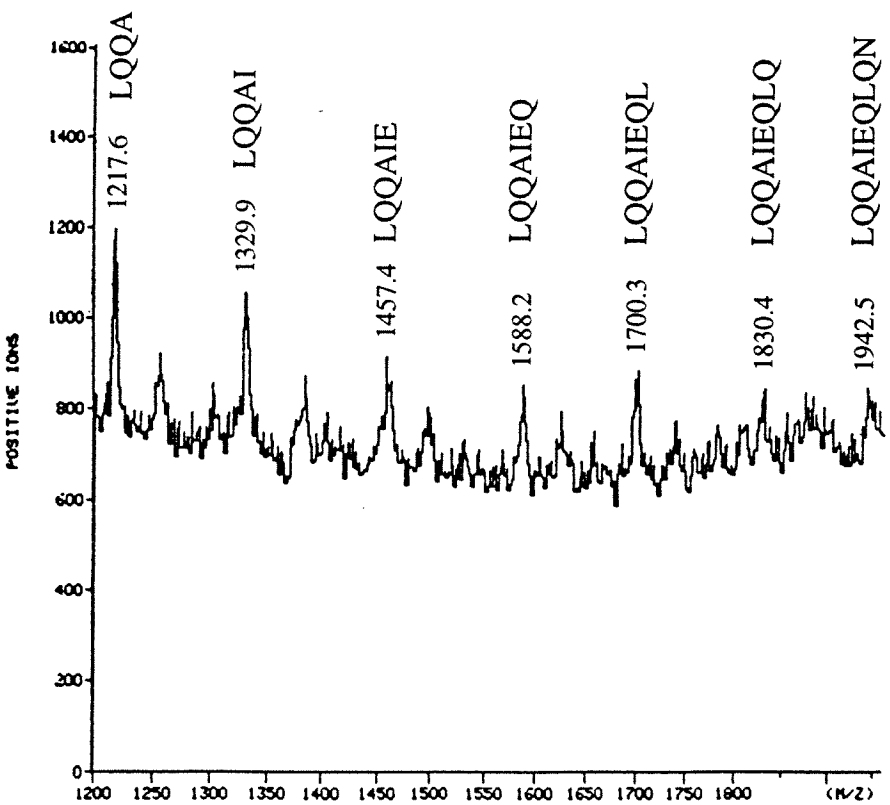
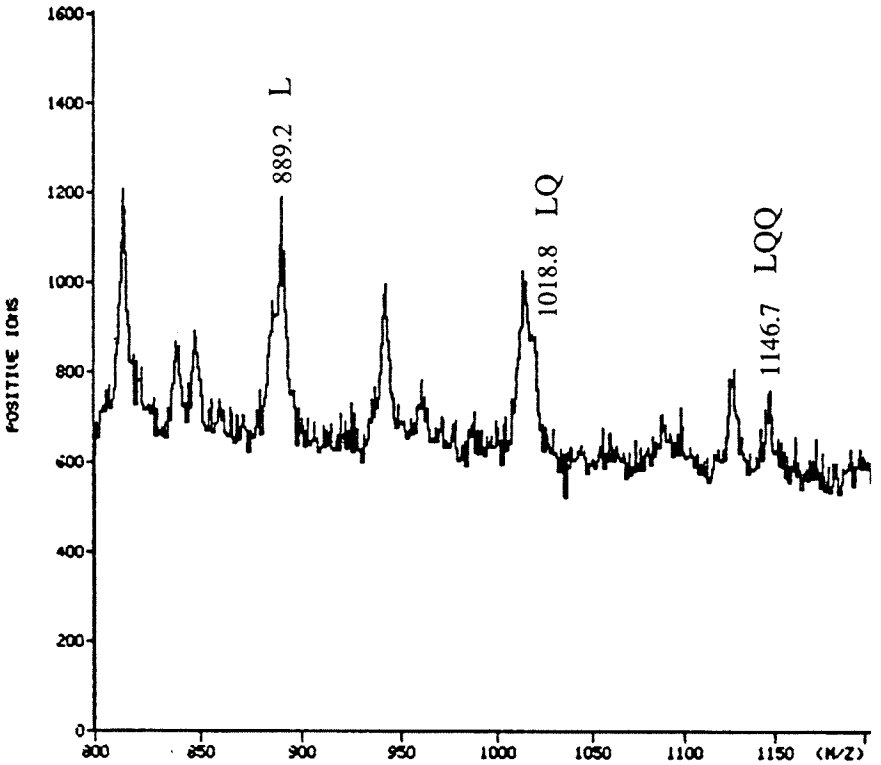

1217.6 LQQA
1329.9 LQQAI
1457.4 LQQAIE
1588.2 LQQAIEQ
1700.3 LQQAIEQL
1830.4 LQQAIEQLQ
1942.5 LQQAIEQLQN
POSITIVE IONS
(M/Z)

Figure 2.20. ^{252}Cf PDMS of $[Rh(phi)_2(phen')]^{3+}$-LQQAI-QLQNAAAA-COOH (Sk-α_A-4) showing the fragmentation pattern in detail and revealing unambiguously the location of the missing Glx residue to be Glu^6 (comparison to figure 2.19). Only the fragments containing intact $[Rh(phi)_2(phen')]^{3+}$ have been indicated. The unlabeled peaks correspond to the fragments that have lost a phi ligand. The complete spectrum is shown in the appendix to Chapter 2.


POSITIVE IONS
889.2 L
1018.8 LQ
1146.7 LQQ
(M/Z)

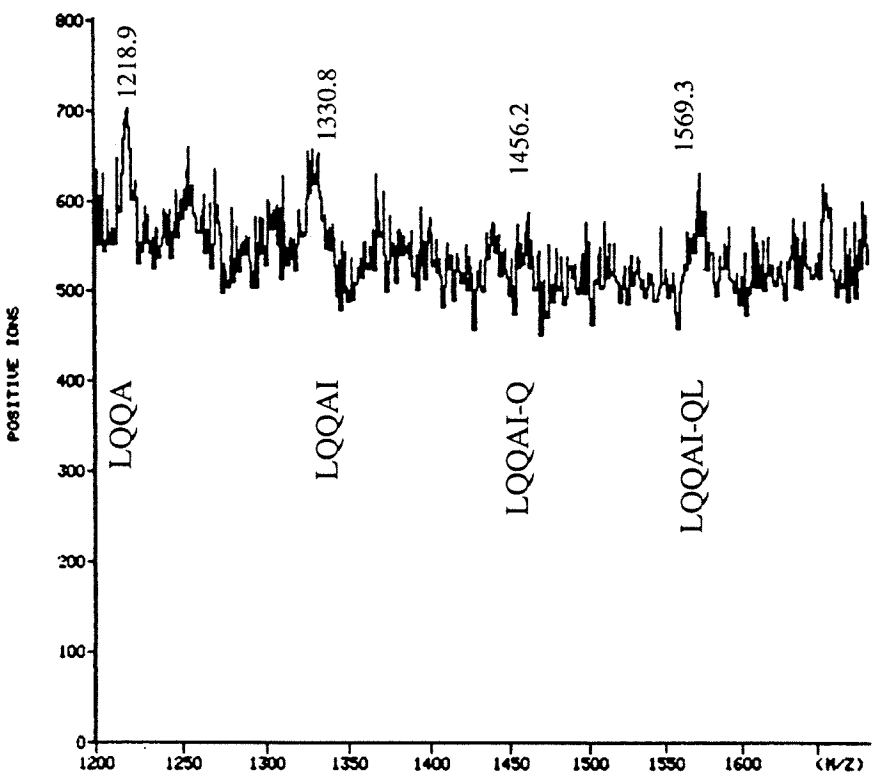
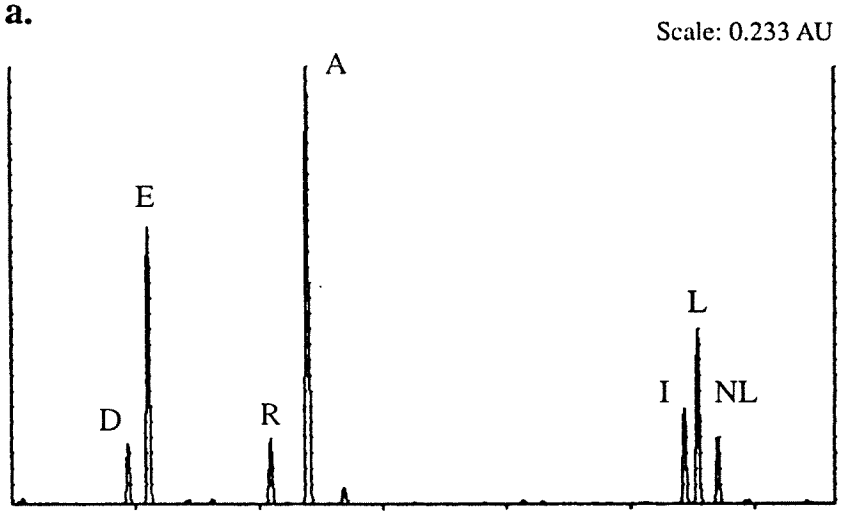

POSITIVE IONS
1218.9
1330.8
1456.2
1569.3
LQQA
LQQAI
LQQAI-Q
LQQAI-QL
(M/Z)

Figure 2.21. Amino acid analyses of **(a)** $[Rh(phi)_2(phen')]^{3+}$-LQQAIERLQNAAAA-COOH (Sk-α_C) and **(b)** $[Rh(phi)_2(phen')]^{3+}$-LQQAIER(Tos)LQNAAAA-COOH (Sk-α_C-7). The amino acid analyses for the two complexes are identical. The calculated ratios of the amino acids are: Asx (1); Glx(4); Arg (1); Ala (5); Ile (1); Leu (2). Nor-leucine (NL) is used as an internal standard during peptide hydrolysis for quantitation of the amino acids.

a.



b.

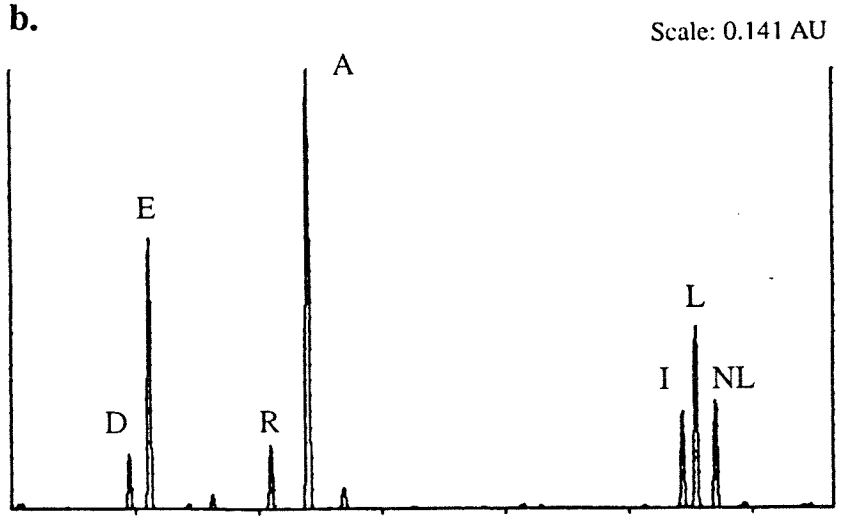
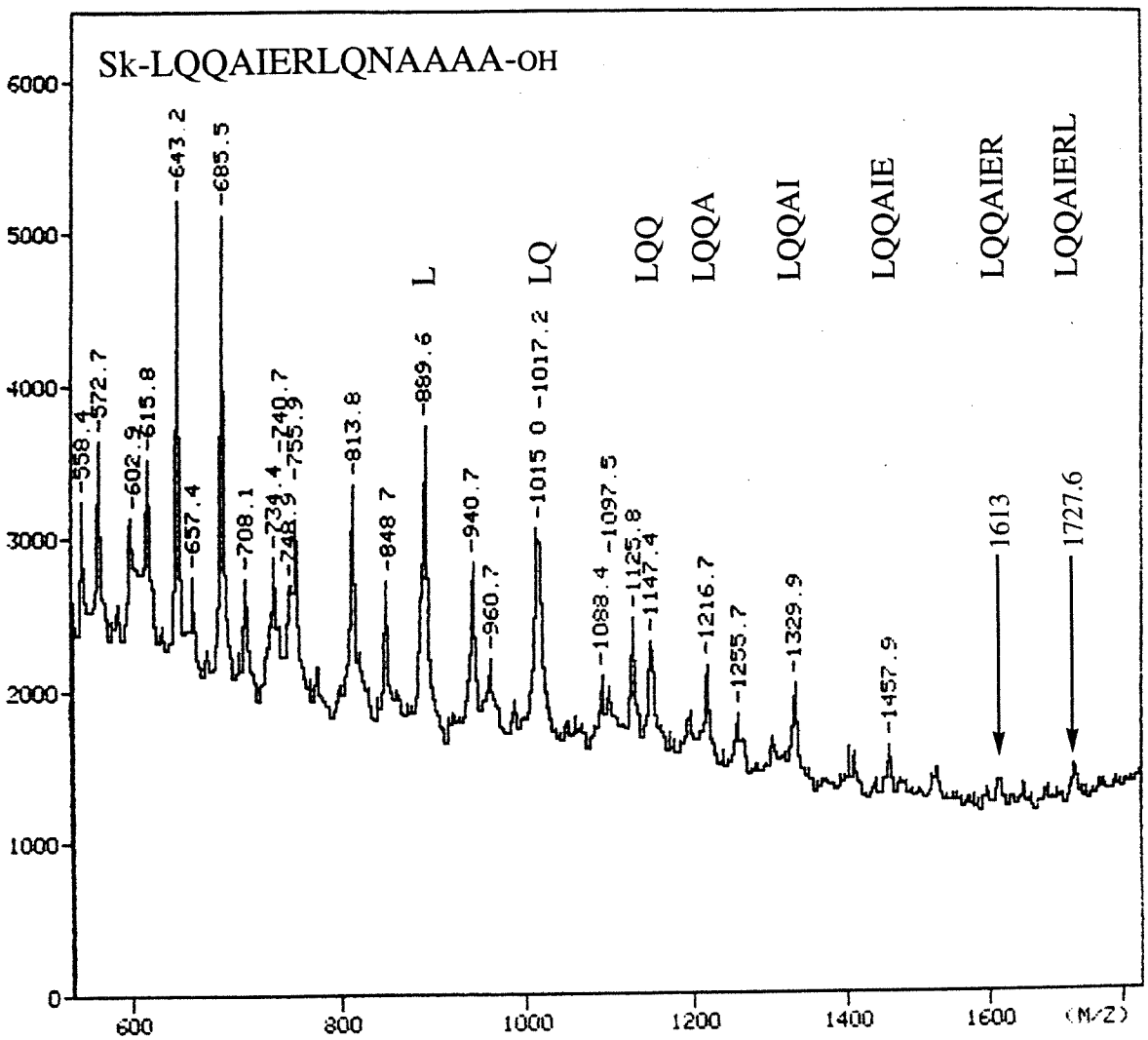


Figure 2.22. ^{252}Cf PDMS of $[Rh(phi)_2(phen')]^{3+}$-LQQAIERLQNAAAA-COOH (Sk-α_C) showing the fragmentation pattern in detail. Only the fragments containing intact $[Rh(phi)_2(phen')]^{3+}$ have been indicated. The unlabeled peaks correspond to the fragments that have lost a phi ligand. **a.** Spectrum from 550-1800 mass units; **b.** spectrum from 1400-2125 mass units. A spectrum showing the molecular ion peak is shown in the appendix to Chapter 2.


Sk-LQQAIERLQNAAAA-OH
6000
5000
4000
3000
2000
1000
0
600
800
1000
1200
1400
1600
(M/Z)
-558.4
-572.7
-602.2
-615.8
-643.2
-657.4
-685.5
-708.1
-734.4
-740.7
-755.9
-813.8
-848.7
-889.6
-940.7
-960.7
-1015.0
-1017.2
-1088.4
-1097.5
-1125.8
-1147.4
-1216.7
-1255.7
-1329.9
-1457.9
L
LQ
LQQ
LQQA
LQQAI
LQQAIE
LQQAIER
LQQAIERL
1613
1727.6

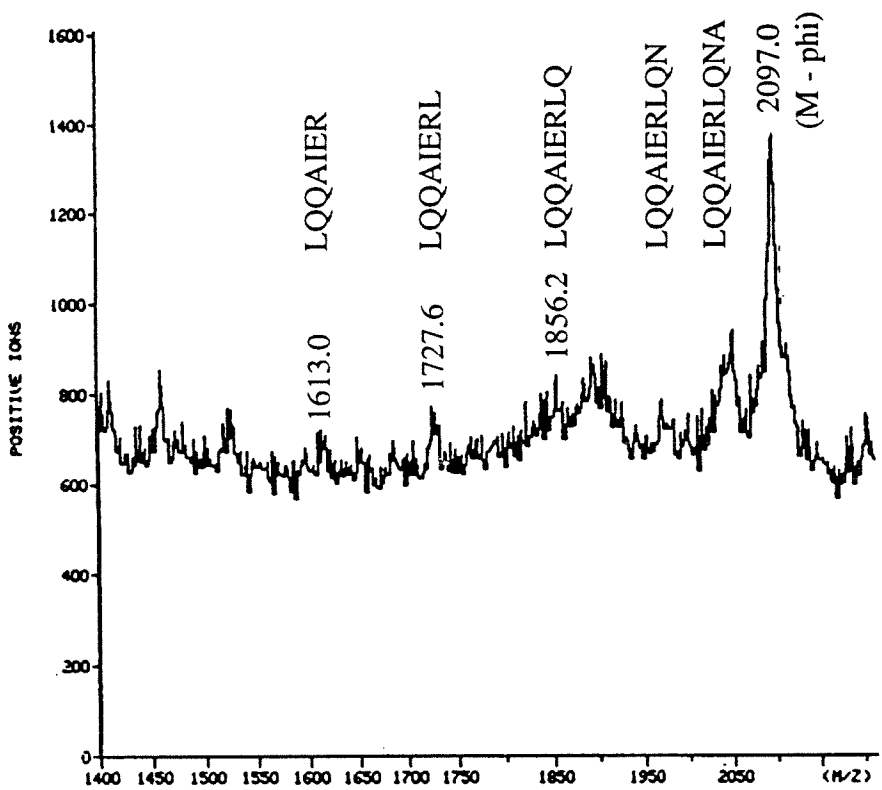
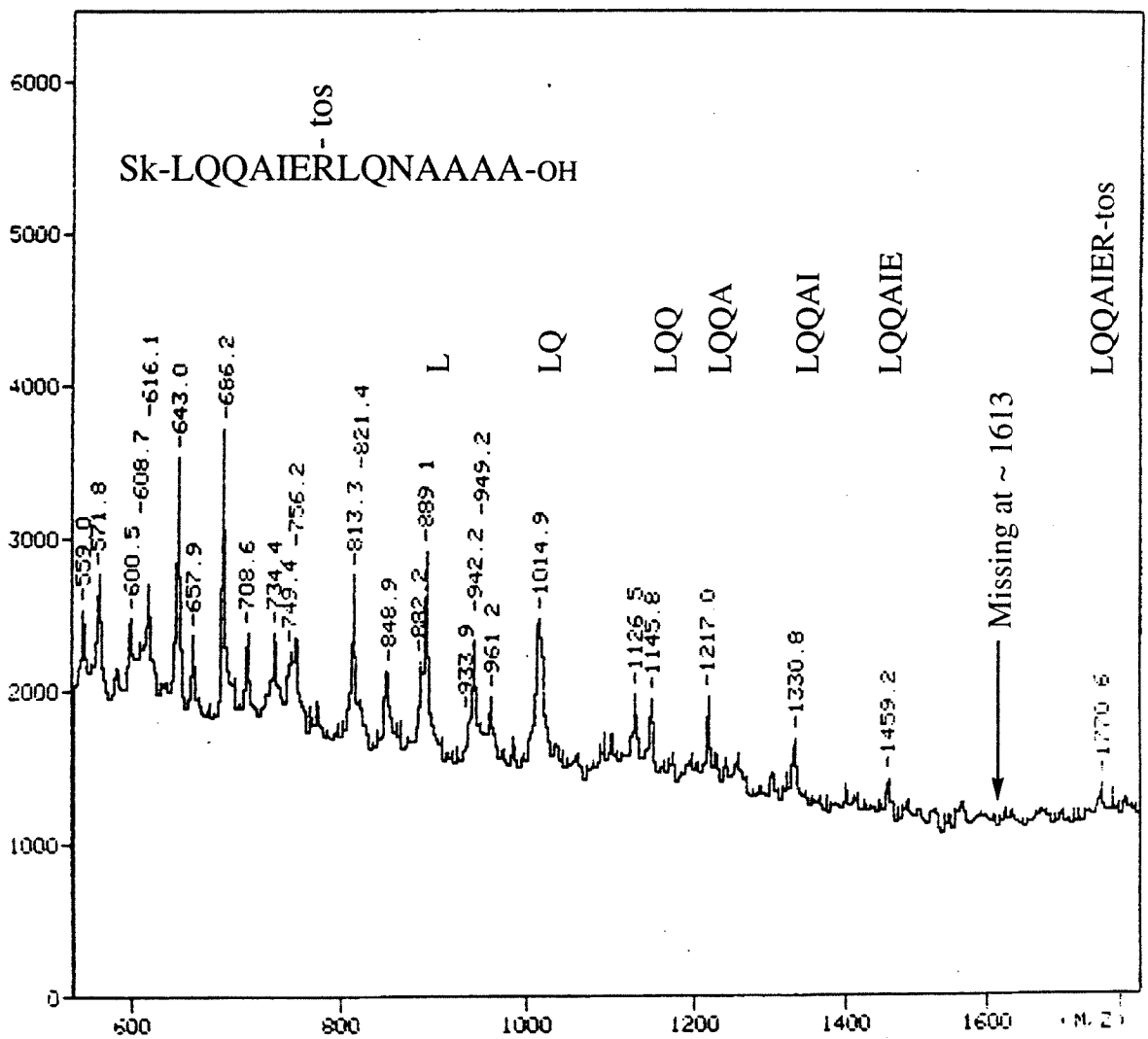
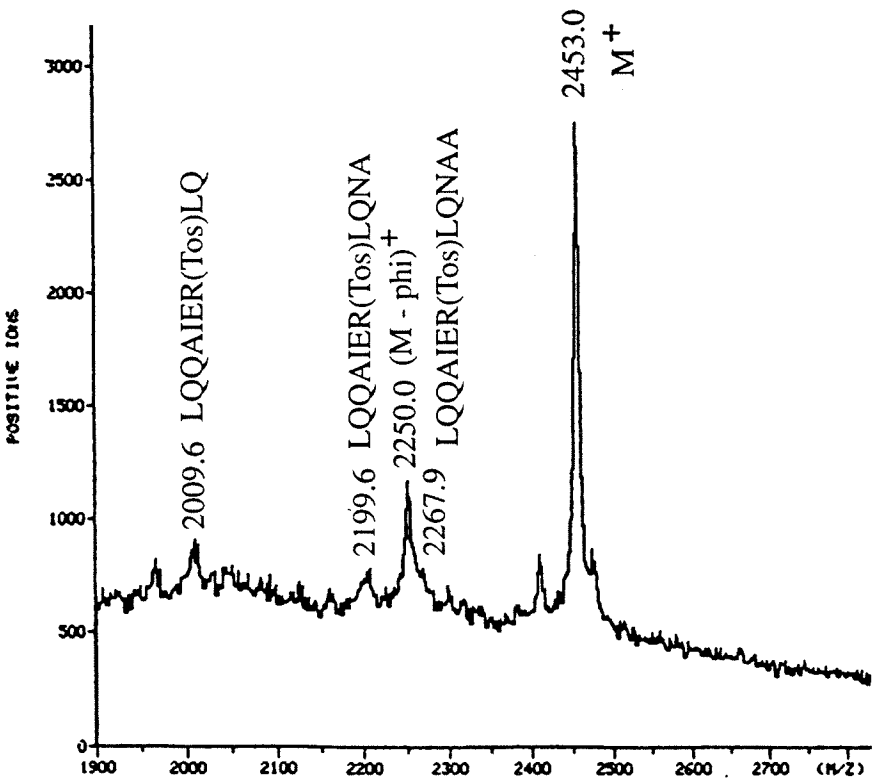

POSITIVE IONS
1600
1400
1200
1000
800
600
400
200
0
1400 1450 1500 1550 1600 1650 1700 1750 1850 1950 2050 (M/Z)
1613.0 LQQAIER
1727.6 LQQAIERL
1856.2 LQQAIERLQ
LQQAIERLQN
LQQAIERLQNA
2097.0 (M - phi)

Figure 2.23. ^{252}Cf PDMS of [Rh(phi)$_2$(phen')]$^{3+}$-LQQAIER(Tos)LQNAAAA-COOH (Sk-α_C-7) showing the fragmentation pattern in detail. Only the fragments containing intact [Rh(phi)$_2$(phen')]$^{3+}$ have been indicated. The unlabeled peaks correspond to the fragments that have lost a phi ligand. **a.** Spectrum from 550-1800 mass units; **b.** spectrum from 1900-2800 mass units.


Sk-LQQAIERLQNAAAA-OH
- tos
L
LQ
LQQ
LQQA
LQQAI
LQQAIE
LQQAIER-tos
Missing at ~ 1613
-559.0
-571.8
-600.5
-608.7
-616.1
-643.0
-657.9
-666.2
-708.6
-734.4
-749.4
-756.2
-813.3
-821.4
-848.9
-882.2
-889.1
-933.9
-942.2
-949.2
-961.2
-1014.9
-1126.5
-1145.8
-1217.0
-1330.8
-1459.2
-1770.6
6000
5000
4000
3000
2000
1000
0
500
800
1000
1200
1400
1600


POSITIVE IONS
3000
2500
2000
1500
1000
500
0
1900
2000
2100
2200
2300
2400
2500
2600
2700
(M/Z)
2009.6 LQQAIER(Tos)LQ
2199.6 LQQAIER(Tos)LQNA
2250.0 (M - phi)+
2267.9 LQQAIER(Tos)LQNAA
2453.0
M+

exactly to the residue containing the protecting group, i.e., Arg^7. Arg-Tos containing peptides are more facilely deprotected using HF. For TFMSA deprotection protocols, Arg-Mts is preferred because the Mts group comes off under milder conditions. This error in choosing appropriate protecting groups for the chemistry employed was discovered by the sequence information obtained from the PDMS spectra of Sk-α_C and Sk-α_C-9.

2.4. Discussion

2.4.1. Scope of the Synthesis. A general method has been developed for the covalent attachment of inert coordinatively saturated rhodium complexes to a specific site on synthetic peptides. The peptides tethered here are 11-15 amino acids in length, contain 15 out of the 20 naturally occurring residues and are tethered to a rhodium complex through a glutaryl linker pendant on a phen ligand. The methodology may be easily generalized to include peptides of any length synthetically accessible, contain any desired amino acid residues and be coupled to rhodium complexes with different chelating ligands. Work along these lines is in progress in the Barton laboratory.[15, 17] Other coordinatively saturated metal complexes which are stable to the conditions required may also be utilized.[3, 24]

2.4.2. Comparison of the Different Methods: Advantages and Limitations.

Solution Phase versus Solid Phase. The solid phase strategy allows for selective deprotection of one functional group on the peptide to serve as or to attach to a specific ligand for the metal complex, and this selectivity represents a clear advantage over solution phase strategies. In the syntheses described in this chapter, the amino terminus has been used as the attachment point for a chelator. Allowing the synthesis to proceed on a solid support also ensures the separation of unreacted reagents by filtration, making the subsequent chromatographic purification easier. Indeed, the chromatograms

accurately reflect the efficiency of the initial peptide synthesis. The obvious limitation of such an approach is that only metal complexes that are stable to the deprotection and cleavage conditions from the resin support can be used. The overall yields of the rhodium-peptide complex are modest and are dependent primarily on the yield of the initial peptide synthesis. The presence of the metal center does, however, further limit the yield, and in particular the presence of covalently bound metal complex tends to inhibit recovery of the product from the solid support.

The solution phase methods give very poor yields and are not regioselective if more than one point of attachment exists on the peptide. With seemingly deactivated (sterically or electronically) electrophiles like **6**, the peptides may self-condense to form multimers, and it becomes difficult to separate side products and unreacted reagents during multiple coupling cycles. Unlike the ruthenium complexes reported in the literature,[1-3] peptide coordination to rhodium precursors in the solution phase coordination method is a serious concern. However, in principle, the efficiency of the initial peptide synthesis has little to do with the overall yields by the solution phase methods, since the peptide has already been purified to homogeniety prior to the coupling reaction. Thus products corresponding to attachment of the rhodium complex to failure sequences and comigrating with authentic material, particularly with longer sequences, are not a concern. Moreover, the solution phase strategy may be the only viable method for metal complexes that cannot withstand the harsh peptide deprotection and cleavage conditions from the resin. The solution method proves useful in post-synthetic modifications of the metal-peptide complexes, such as methylation of side chain carboxylates, as discussed in Chapter 5.

Coordination versus Direct Coupling Methods. Two strategies for coupling together the metal center and the peptide on the resin have been described, the coordination method and direct coupling, and each offers certain advantages. No substantial differences in yield are observed. The presence of the metal center tends to

inhibit the coupling reaction, but coordination on the resin is of lower efficiency than the coordination of the metal complex alone in solution. The direct coupling method also preserves the integrity of the metal center and may be preferable under conditions where diastereomerically pure samples are needed. The conditions for synthesis do not lead to racemization about the metal center, if enantiomerically pure metal complexes are used in the coupling. Nonetheless, pure metal-peptide chimeras may also be resolved by chromatography with chiral eluents from samples which are racemic about the metal center. Furthermore, UV-visible spectroscopy offers a sensitive gauge of the coordination about the metal center. The presence of these two strategies therefore provides versatility and the preferred strategy will depend upon the application required.

2.4.3. Features of the Metal-Peptide Chimeras. The spectroscopic characteristics of the metal-peptide complexes are seen to be the sum of those of the isolated metal complex and the appended peptide. The electronic spectra for the chimeras are very similar to that of the parent $[Rh(phi)_2(phen')]^{3+}$. The coordination sphere of the rhodium is not perturbed by coupling to the peptide; any ligand exchange with side chain coordination would be sensitively detected spectroscopically. Besides providing a measure of concentration (because of higher extinction coefficients at longer wavelengths for the chimera than for the peptide alone), the electronic spectrum also is diagnostic of the integrity of the metal-peptide complex.

Significantly, no induced CD signals above 300 nm are observed despite having a chiral peptide attached to a diastereotopic metal complex. The chirality of the individual peptides is not strong enough to favor formation of one enantiomer about the metal center. Likewise, the CD spectra of the pure diastereomers of the metal-peptide complex is equal and opposite above 250 nm. However, below 250 nm the CD spectrum is the sum of the contribution of the peptide (negative ellipticity) and the enantiomerically pure metal complex making up each diastereomer.

Another important benefit of metal attachment to the peptide is evident in the mass spectral analysis. The PDMS spectra of the metal-peptide complexes show, in addition to the parent molecular ion peaks, two families of fragments that reflect the sequential analysis of the metal-peptides. Under conditions used in these experiments, no fragmentation of the peptide is evident without the metal being attached. Thus the covalently bound metal complex enhances the intensity of the A_n fragments substantially, possibly because of the nascent positive charge on the rhodium center. A_n fragments have not been commonly observed in the mass spectrometry of peptides and proteins, but it is noteworthy that limited A_n fragments have been seen in a study of proteins by ^{252}Cf PDMS.[25]

This sequential fragmentation may be powerfully exploited in the analysis of the metal-peptide chimera. Using PDMS as a diagnostic technique, we have been able to characterize mutant chimeras containing deleted amino acids and we have been able to establish sites where side chain protecting groups were still bound. This information would normally be available by peptide sequencing, but common sequencing techniques are not viable with N-terminally modified peptides such as these. Metal complex attachment may therefore prove to be valuable in applications where such sequence analysis is essential.[26]

2.4.4. Implications. Metal-peptide chimeras may therefore be readily constructed using solid-phase synthesis on a multi-milligram scale; a scale that is relevant for many biological applications. The spectroscopic characteristics of the coordination complexes provide a convenient handle to monitor the structure and reactivity, and the mass spectral characteristics provide a unique handle to analyze the peptide sequence. As with metalloproteins, features of coordination chemistry may be incorporated in these metal-peptide chimeras and exploited. One such application, the recognition of double helical DNA, is discussed in the chapters that follow.

2.5. References.

1. Ghadiri, M. R.; Fernholz, A. K. *J. Am. Chem. Soc.* **1990**, *112*, 9633. Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825. Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 4000. Ghadiri, M. R.; Choi, C. *J. Am. Chem. Soc.* **1990**, *112*, 1630. Lieberman, M.; Sasaki, T. *J. Am. Chem. Soc.* **1991**, *113*, 1470. Lieberman, M.; Tabet, M.; Sasaki, T. *J. Am. Chem. Soc.* **1994**, *116*, 5035.

2. Ruan, F. Q.; Chen, Y. Q.; Hopkins, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 9403. Ruan, F. Q.; Chen, Y. Q.; Itoh, K.; Sasaki, T.; Hopkins, P. B. *J. Org. Chem.* **1991**, *56*, 4347. Imperiali, B.; Fisher, S. L. *J. Am. Chem. Soc.* **1991**, *113*, 8527. Wuttke, D. S.; Gray, H. B.; Fisher, S. L.; Imperiali, B. *J. Am. Chem. Soc.* **1993**, *115*, 8455. Imperiali, B.; Kapoor, T. M. *Tetrahedron* **1993**, *49*, 3501.

3. Peek, B. M.; Ross, G. T.; Edwards, S. W.; Meyer, G. J.; Meyer, T. J.; Erickson, B. W. *Int. J. Peptide Protein Res.* **1991**, *38*, 114. Mecklenburg, S. L.; Peek, B. M.; Schoonover, J. R.; McCafferty, D. G.; Wall, C. G.; Erickson, B. W.; Meyer, T. J. *J. Am. Chem. Soc.* **1993**, *115*, 5479.

4. Shullenberger, D. F.; Eason, P. D.; Long, E. C. *J. Am. Chem. Soc.* **1993**, *115*, 11038.

5. Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; DeGrado, W. F.; Dutton, P. L. *Nature* **1994**, *368*, 425. Choma, C. T.; Lear, J. D.; Nelson, M. J.; Dutton, P. L.; Robertson, D. E.; DeGrado, W. F. *J. Am. Chem. Soc.* **1994**, *116*, 856. Handel, T. M.; Williams, S. A.; DeGrado, W. F. *Science* **1993**, *261*, 879. Handel, T.; DeGrado, W. F. *J. Am. Chem. Soc.* **1990**, *112*, 6710.

6. Krizek, B. A.; Amann, B. T.; Kilfoil, V. J.; Merkle, D. L.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4518. Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937.

7. Wade, W. S.; Koh, J. S.; Han, N.; Hoekstra, D. M.; Lerner, R. A. *J. Am. Chem. Soc.* **1993**, *115*, 4449. Barbas III, C. F.; Rosenblum, J. S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 6385. Wade, W. S.; Ashley, J. A.; Jahangiri, G. K.; McElhaney, G.; Janda, K. D.; Lerner, R. A. *J. Am. Chem. Soc.* **1993**, *115*, 4906. Pessi, A.; Bianchi, E.; Crameri, A.; Venturini, S.; Tramontano, A.; Sollazzo, M. *Nature* **1993**, *362*, 367.

8. Pyle, A. M.; Barton, J. K. *Prog. Inorg. Chem.* **1990**, *38*, 413. Chow, C. S.; Barton, J. K. *Methods Enzymol.* **1992**, *212*, 219. Dupureur, C. M; Barton, J. K. *Comprehensive Supramolecular Chemistry*, (**1995**), Pergamon Press, in press.

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

10. (a) Stewart, J. M.; Young, J. D. *Solid Phase Peptide Synthesis*, 2nd Ed., (**1984**), Pierce Chemical Company, Rockford, Illinois. (b) Edmondson, J. M.; Klebe, R. J.; Zardeneta, G.; Weintraub, S. T.; Kanda, P. *BioTechniques*, **1988**, *6*, 866.

11. Atherton, E.; Sheppard, R. C. *Solid Phase Peptide Synthesis: A Practical Approach*, (**1989**), IRL Press, Oxford.

12. Yoon, S. S.; Still W. C. *J. Am. Chem. Soc.* **1993**, *115*, 823. Nestler, H. P.; Bartlett, P. A.; Still, W. C. *J. Org. Chem.* **1994**, *59*, 4723. Nielson, J.; Brenner, S.; Janda, K. D. *J. Am. Chem. Soc.* **1993**, *115*, 9812. Baum, R. M. *Chem. and Engr. News* February 7, **1994**, 20.

13. Pyle, A. M.; Chiang, M. Y.; Barton, J. K. *Inorg. Chem.* **1990**, *29*, 4487.

14. Abbreviations for peptide synthesis reagents: Bzl: Benzyl; Cl-Z: 2-chlorobenzyloxycarbonyl; DCC: N, N-dicyclohexylcarbodiimide; DIEA: *N,N*-diisopropylethylamine; DIPC: N, N-diisopropylcarbodiimide; DMAP: 4-dimethylaminopyridine; DMF: *N,N*-dimethylformamide; DSC: *N,N*-disuccinylmidyl carbonate; FMOC: 9-fluorenylmethoxycarbonyl; HOBt: 1-hydroxybenzotriazole; MBHA: 4-methylbenzahydrylamine; Mts: mesitylene-2-

sulfonyl; NMP: 4-methylpyrrolidone; OBzl: benzyl; PAM: phenylacetamido-methyl; PDMS: plasma desorption mass spectrometry; PEG: polyethylene glycol; tBoc: *tert.*-butyloxycarbonyl; tBu: *tert.*-butyl; TBTU: O-(1H-benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate; Tos: p-toluylsulfonyl; TSTU: O-(N-succinimidyl)-N,N,N',N'-tetramethyluronium tetrafluoroborate; TFA: trifluoroacetic acid; TFMSA: trifluoromethanesulfonic acid.

15. Zimmermann, K. *Post Doctoral Research Report* (**1993**), Caltech.

16. Dollimore, L. S.; Gillard, R. D. *J.C.S Dalton Trans.*, **1973**, 933. Cartwright, P. S.; Gillard, R. D.; Sillanpaa, E. R. J. *Polyhedron,* **1987**, *6,* 105.

17. Lin, S. C. *Ph. D. Dissertation*, Work in Progress, Caltech.

18. Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Anal. Biochem.* **1981**, *117*, 147.

19. Swerdloff, M. D.; Anderson, S. B.; Sedgwick, R. D.; Gabriel, M. K.; Brambilla, R. J.; Hindenlang, D. M.; Williams, J. T. *Int. J. Pept. Prot. Res.* **1989**, *33* , 318.

20. King, D. S.; Fields, C. G.; Fields, G. B. *Int. J. Pept. Prot. Res.* **1990**, *36,* 255.

21. Pyle, A. M. *Ph. D. Dissertation* **1989**, Columbia University.

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

23. David, S. S.; Barton, J. K. *J. Amer. Chem. Soc.* **1993**, *115*, 2984. Krotz, A. H.; Kuo, L. Y.; Shields, T. P.; Barton, J. K. *J. Amer. Chem. Soc.* **1993**, *115*, 3877.

24. Jenkins, Y.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 8736.

25. Bunk, D. M.; Macfarlane, R. D. *Proc. Natl. Acad. Sci.* **1992**, *89*, 6215.

26. Stults, J. T. *Methods Biochem. Anal.* **1990**, *34*, 145. Chait, B. T.; Kent, S. B. H. *Science* **1992**, *257*, 1885. Metzger, J. W.; Kempter, C.; Wiesmüller, K. -H. Jung, G. *Anal. Biochem.* **1994**, *219*, 261.

2.6. Appendix: Characterization Data for the Different Metal-Peptide Complexes

General Notes: The parent peptides were synthesized by automated peptide synthesis using N-tBoc protected amino acids except Sk-L219 and Sk-L220 which were synthesized using N-FMOC amino acids. For the tBoc syntheses, PAM resin was used to obtain a C-terminal acid or MBHA resin was used for C-terminal amides. The MS parent ion peaks correspond to the respective molecular ions unless noted otherwise. The metal-peptide complexes generally exhibit a molecular ion peak corresponding to a $(M-H)^+$ or $(M-2H)^+$ species, while the free peptides usually show up as the $(M+H)^+$ species.

Table 2.1. List of the Metal-Peptide Complexes Synthesized and Characterized.

Name	Description of Metal-Peptide Complex
Sk-α_3	$[Rh(phi)_2(phen'\text{-}TQQSIEQLQNG\text{-}COOH)]^{+}$
Sk-α_A	$[Rh(phi)_2(phen'\text{-}LQQAIEQLQNAAAA\text{-}COOH)]^{+}$
Sk-α_A-4	$[Rh(phi)_2(phen'\text{-}LQQAIQLQNAAAA\text{-}COOH)]^{2+}$
Sk-α_A-OMe	$[Rh(phi)_2(phen'\text{-}LQQAIE(OMe)QLQNAAAA\text{-}COOMe)]^{+}$
Sk-α_B	$[Rh(phi)_2(phen'\text{-}LQAAIEQLQNAAAA\text{-}COOH)]^{+}$
Sk-α_C	$[Rh(phi)_2(phen'\text{-}LQQAIERLQNAAAA\text{-}COOH)]^{2+}$
Sk-α_C-9	$[Rh(phi)_2(phen'\text{-}LQQAIER(Tos)LQNAAAA\text{-}COOH)]^{+}$
Sk-α_C-9-OMe	$[Rh(phi)_2(phen'\text{-}LQQAIE(OMe)R(Tos)LQNAAAA\text{-}COOMe)]^{3+}$
Sk-α_E	$[Rh(phi)_2(phen'\text{-}ALQQSIAQLQNKAAA\text{-}CONH2)]^{4+}$
Sk-α_F	$[Rh(phi)_2(phen'\text{-}TQQSKKQLQNKAAA\text{-}CONH2)]^{6+}$
Sk-α_G	$[Rh(phi)_2(phen'\text{-}TQQSKKRLQNKAAA\text{-}CONH2)]^{7+}$
Sk-α_H	$[Rh(phi)_2(phen'\text{-}TQRSKKQLQNKAAA\text{-}CONH2)]^{7+}$
Sk-α_I	$[Rh(phi)_2(phen'\text{-}TQASKKQLQNKAAA\text{-}CONH2)]^{6+}$
Sk-P22I	$[Rh(phi)_2(phen'\text{-}AANVAIAAWDRAA\text{-}CONH2)]^{3+}$
Sk-P22K	$[Rh(phi)_2(phen'\text{-}NVAIAAWDRAA\text{-}CONH2)]^{3+}$
Sk-P22L	$[Rh(phi)_2(phen'\text{-}AANVAIAAWARAA\text{-}CONH2)]^{4+}$
Sk-P22M	$[Rh(phi)_2(phen'\text{-}AAQAAAQAWERAA\text{-}CONH2)]^{3+}$
Sk-L219	$[Rh(phi)_2(phen'\text{-}KAEGAKAFFKRSA\text{-}CONH2)]^{6+}$
Sk-L220	$[Rh(phi)_2(phen'\text{-}KAAKAFFKRSA\text{-}CONH2)]^{7+}$

2.6.1. Sk-α_3: [Rh(phi)$_2$(phen'-TQQSIEQLQNG-COOH)]$^+$

Synthesis Procedure:

(i). Solution phase coupling

(ii). Solution phase coordination (attempted)

(iii). Solid phase coordination

Amino Acid Analysis:

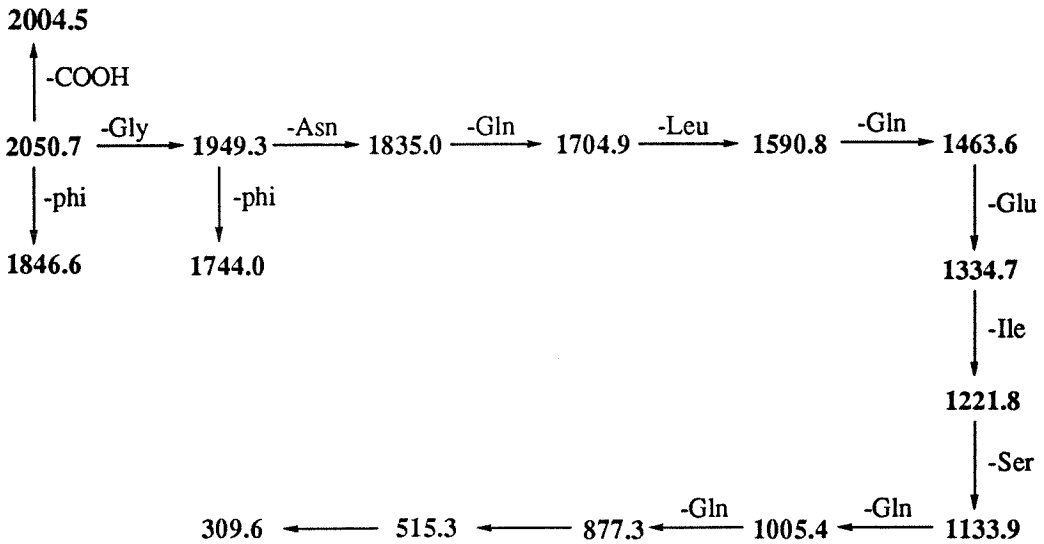
Amino Acid	Expected	Observed (free peptide)	Observed (Sk-peptide)
Asx	1	0.99	0.97
Glx	5	5.00	5.00
Ser	1	0.97	0.90
Gly	1	1.03	1.09
Thr	1	1.06	0.87
Ile	1	0.86	0.93
Leu	1	1.08	1.08

Mass Spectrometry:

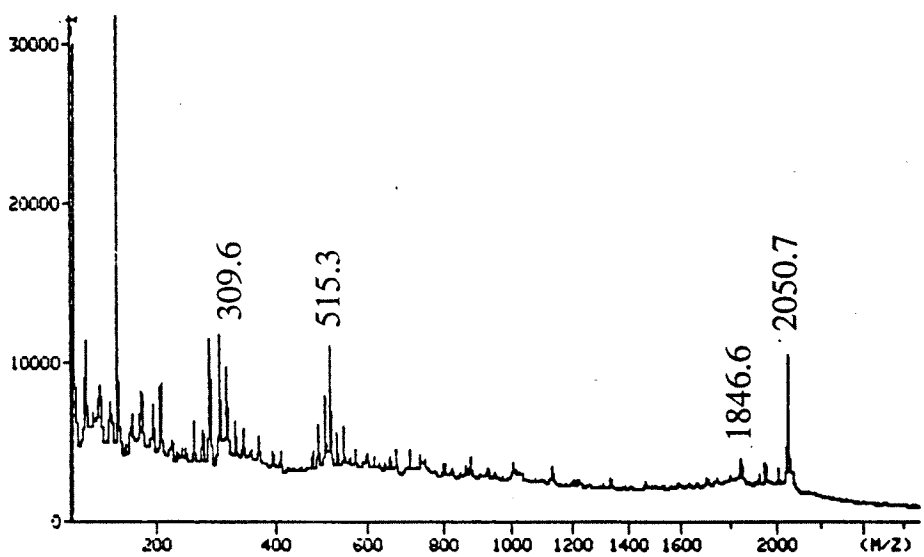
	Free Peptide (α_3)	**Coupled Peptide (Sk-α_3)**
Avg. Molecular Mass:	1246.3	2049.9
Monoisotopic Mass:	1245.6	2048.7
Observed Mass:	1245.8	2050.7

Fragmentation Pattern:

The observed masses of the fragments are as follows:



Peaks corresponding to loss of a phi ligand from the various fragments are also observed.



2.6.2. Sk-α_A: [Rh(phi)$_2$(phen'-LQQAIEQLQNAAAA-COOH)]$^+$

Synthesis Procedure:

(i). Solution phase coupling

(ii). Solution phase coordination (attempted)

(iii). Solid phase coordination

Amino Acid Analysis:

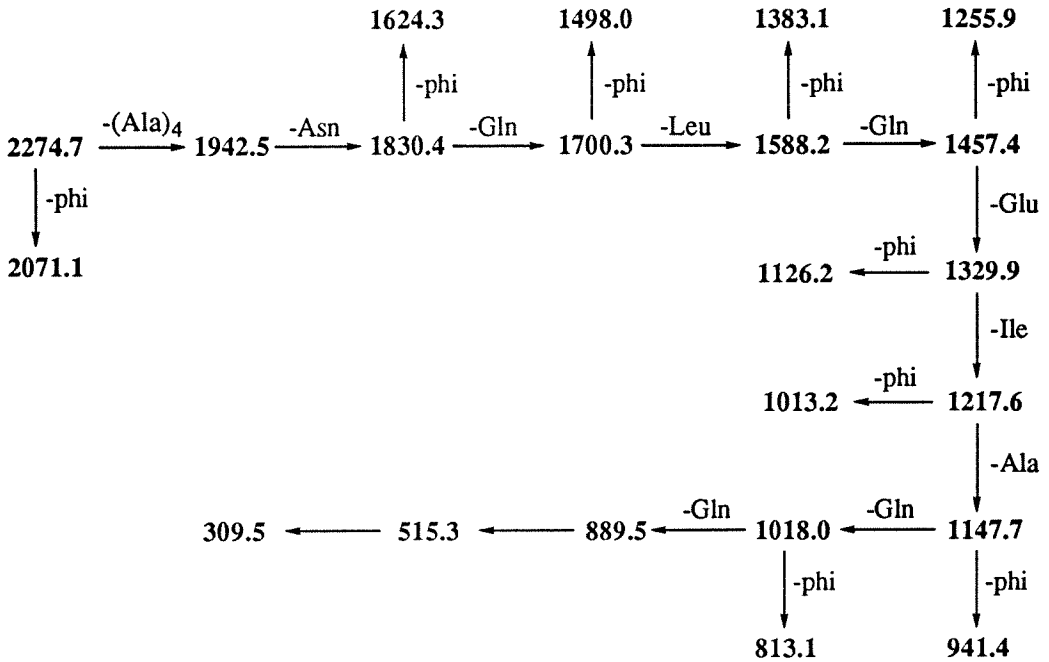
Amino Acid	Expected	Observed (free peptide)	Observed (Sk-peptide)
Asx	1	1.09	1.12
Glx	5	4.92	4.99
Ala	5	5.00	5.00
Ile	1	1.43	1.22
Leu	2	2.19	2.21

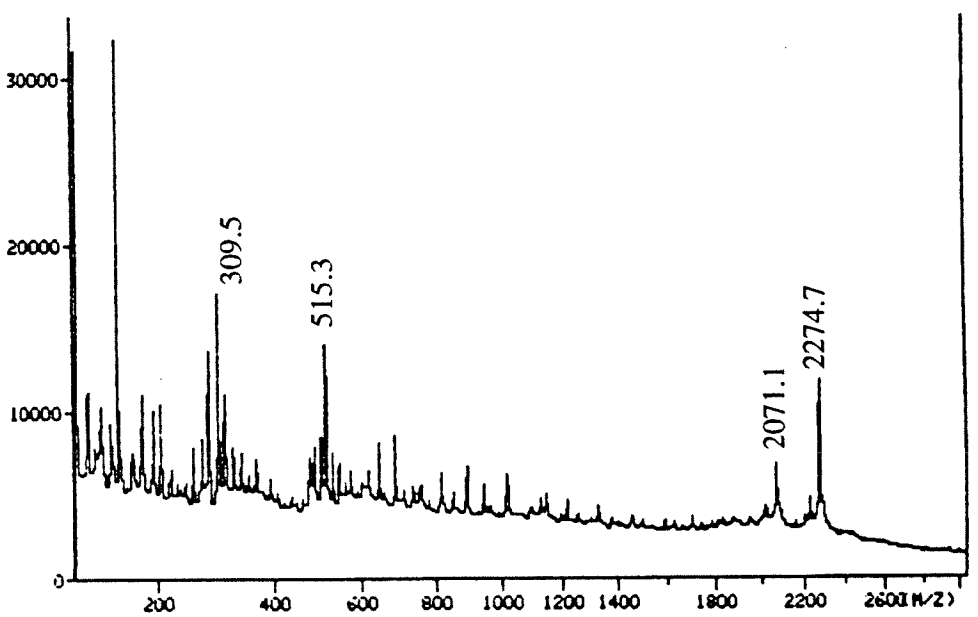
Mass Spectrometry:

	Free Peptide (α_A)	**Coupled Peptide (Sk-α_A)**
Avg. Molecular Mass:	1469.6	2275.3
Monoisotopic Mass:	1468.8	2273.8
Observed Mass:	1467.8	2274.7

Fragmentation Pattern:

The observed masses of the fragments are as follows:





2.6.3. Sk-α_A-4: $[Rh(phi)_2(phen'\text{-}LQQAIQLQNAAAA\text{-}COOH)]^{2+}$

Synthesis Procedure: Solid phase coordination. Isolated during the synthesis of Sk-α_A.

Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Asx	1	1.08
Glx	5	4.42
Ala	5	5.00
Ile	1	1.00
Leu	2	2.08

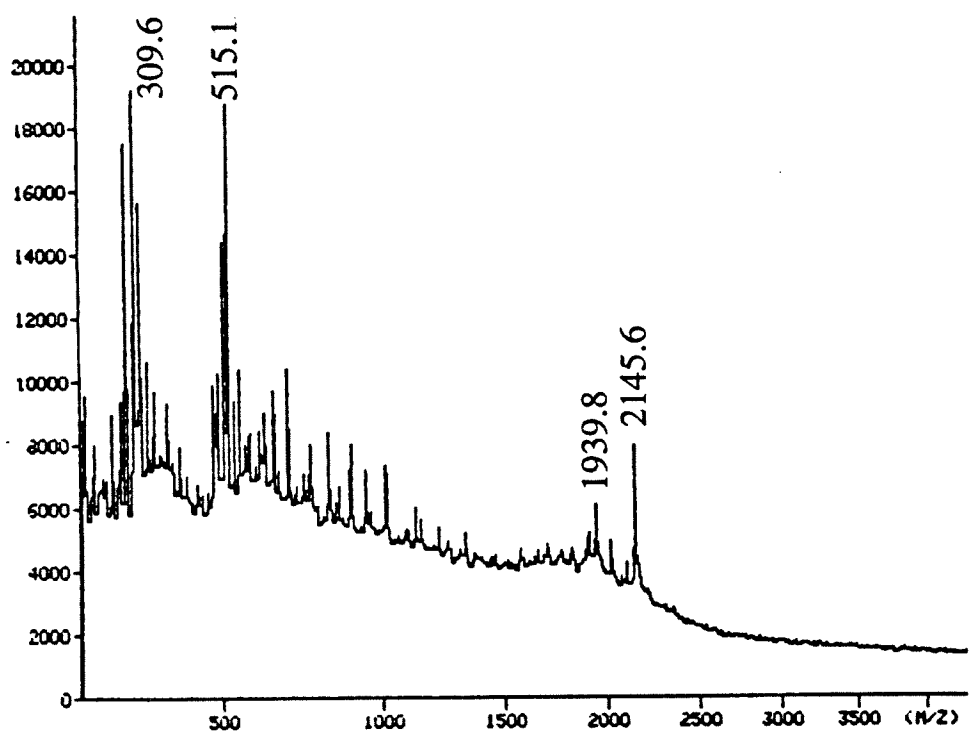
Mass Spectrometry:

Coupled Peptide (Sk-α_A-4)

Avg. Molecular Mass: 2146.2

Monoisotopic Mass: 2144.8

Observed Mass: 2145.6



2.6.4. Sk-α_A-OMe: [Rh(phi)$_2$(phen'-LQQAIE(OMe)QLQNAAAA-COOMe)]$^{3+}$

Synthesis Procedure: Sk-α_A (ca. 70 nmole) was dissolved in 200 μL MeOH in an eppendorf tube. DMAP (2 grains) and 10 μL of a 0.5 M solution of DCC in CH_2Cl_2 were added to the eppendorf tube. The eppendorf tube was sealed with parafilm and stirred on a mechanical shaker for 40 h. The reaction was quenched by addition of 1 mL H_2O (0.1% TFA) and purified by HPLC.

Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Asx	1	0.89
Glx	4 (+1)	4.22
Ala	5	5.00
Ile	1	1.02
Leu	2	2.19

Mass Spectrometry:

	Coupled Peptide (Sk-α_A-OMe)
Avg. Molecular Mass:	2303.3
Monoisotopic Mass:	2301.8
Observed Mass:	2301.7

2.6.5. Sk-α_B: [Rh(phi)$_2$(phen'-LQAAIEQLQNAAAA-COOH)]$^+$

Synthesis Procedure: Solution phase coordination (attempted). Phen'-α_B was synthesized and purified. The coordination reaction of [Rh(phi)$_2$]$^{3+}$ proved very messy and no distinct products were isolated or characterized definitively.

Amino Acid Analysis:

Amino Acid	Expected	Observed (phen'-peptide)
Asx	1	1.17
Glx	4	3.87
Ala	6	6.00
Ile	1	1.02
Leu	2	1.76

Mass Spectrometry:

	Free Peptide (α_B)	**Coupled Peptide (Phen'-α_B)**
Avg. Molecular Mass:	1412.6	1703.9
Monoisotopic Mass:	1411.8	1702.9
Observed Mass:		1703.0

2.6.6. Sk-α_C: $[Rh(phi)_2(phen'\text{-}LQQAIERLQNAAAA\text{-}COOH)]^{2+}$

Synthesis Procedure:

(i). Solution phase coordination (attempted)

(ii). Solid phase coordination

Amino Acid Analysis:

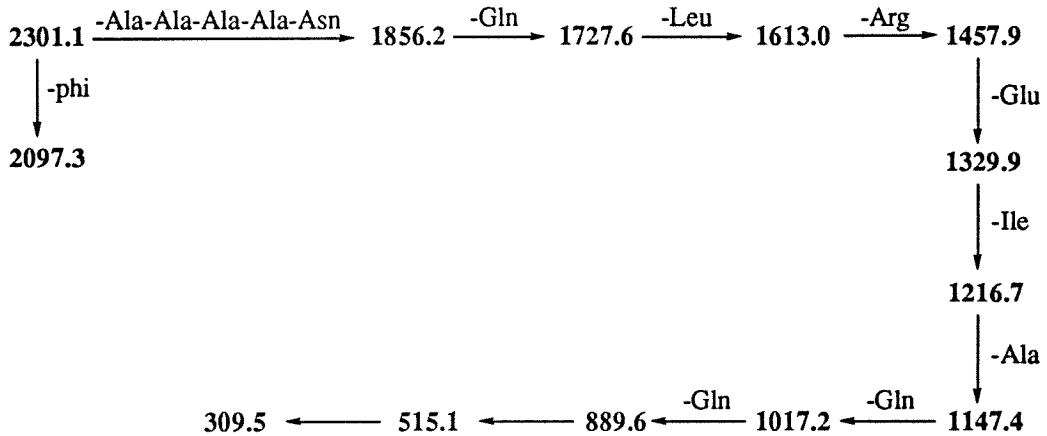
Amino Acid	Expected	Observed (Sk-peptide)
Asx	1	1.2
Glx	4	3.89
Arg	1	1.11
Ala	5	4.16
Ile	1	1.00
Leu	2	1.98

Mass Spectrometry:

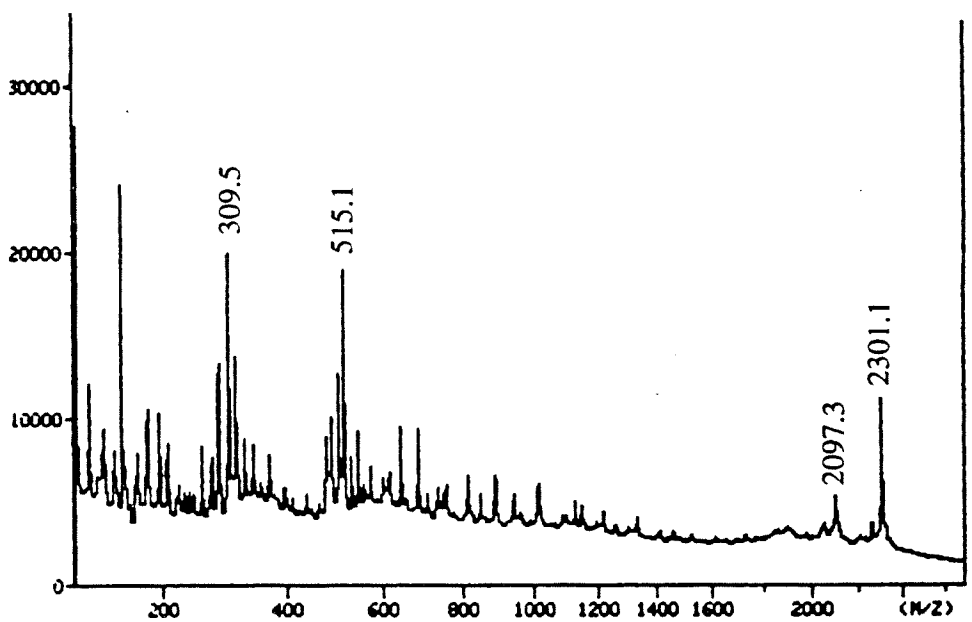
	Free Peptide (α_C)	**Coupled Peptide (Sk-α_C)**
Avg. Molecular Mass:	1496.7	2303.3
Monoisotopic Mass:	1495.8	2301.9
Observed Mass:		2301.1

Fragmentation Pattern:

The observed masses of the fragments are as follows:



Peaks corresponding to loss of a phi ligand from the various fragments are also observed.



2.6.7. Sk-α_C-9: [Rh(phi)$_2$(phen'-LQQAIER(Tos)LQNAAAA-COOH)]$^+$

Synthesis Procedure: Solid phase coordination. Isolated during the synthesis of Sk-α_C.

Amino Acid Analysis:

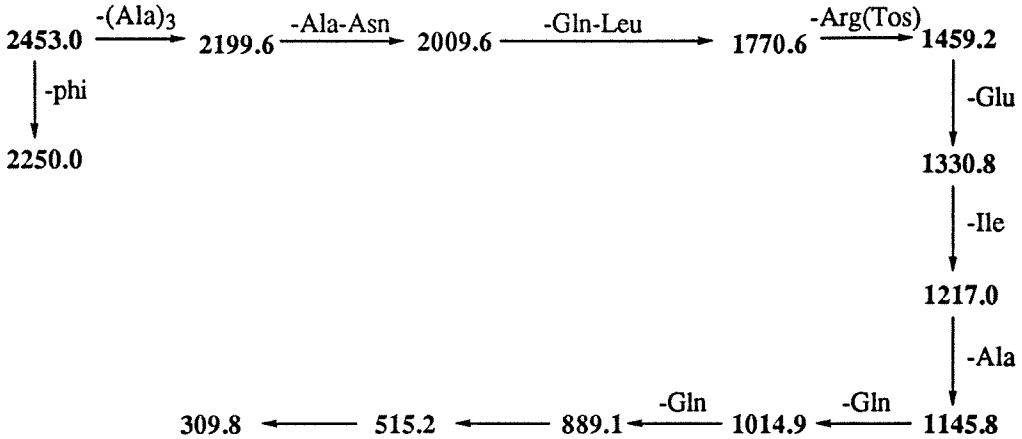
Amino Acid	Expected	Observed (Sk-peptide)
Asx	1	0.99
Glx	4	3.86
Arg	1	1.01
Ala	5	4.15
Ile	1	0.98
Leu	2	2.00

Mass Spectrometry:

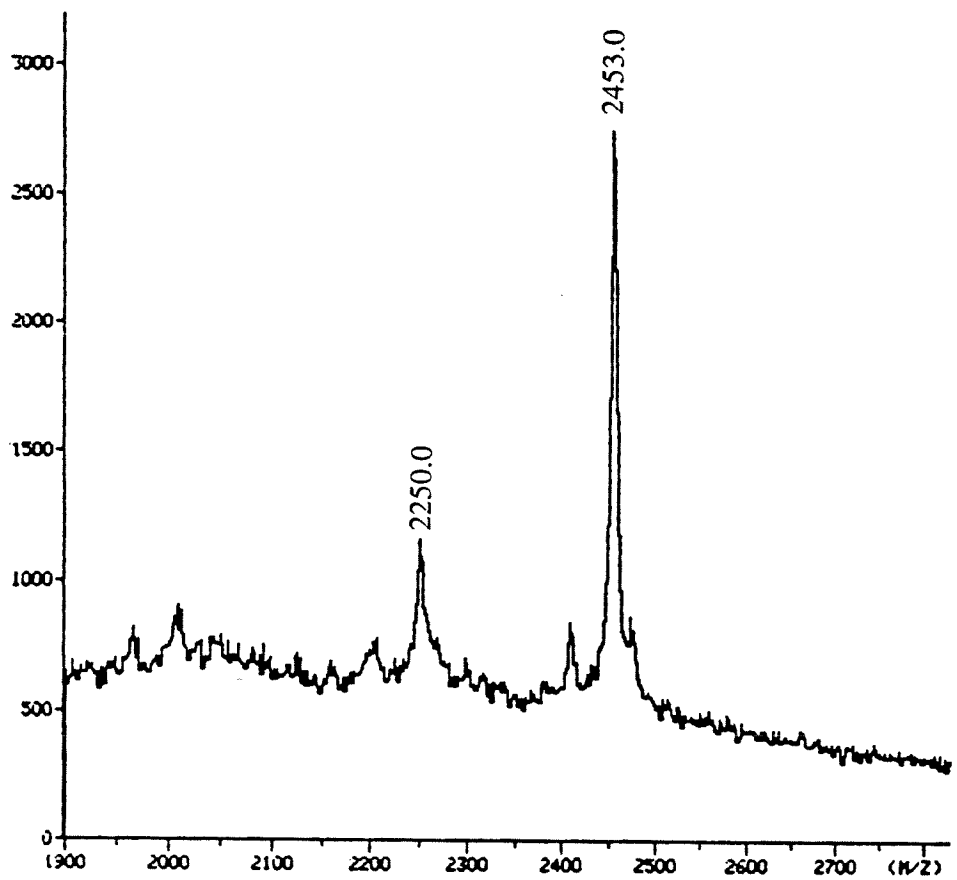
	Coupled Peptide (Sk-α_C-9)
Avg. Molecular Mass:	2456.3
Monoisotopic Mass:	2454.9
Observed Mass:	2453.0

Fragmentation Pattern:

The observed masses of the fragments are as follows:



Peaks corresponding to loss of a phi ligand from the various fragments are also observed.

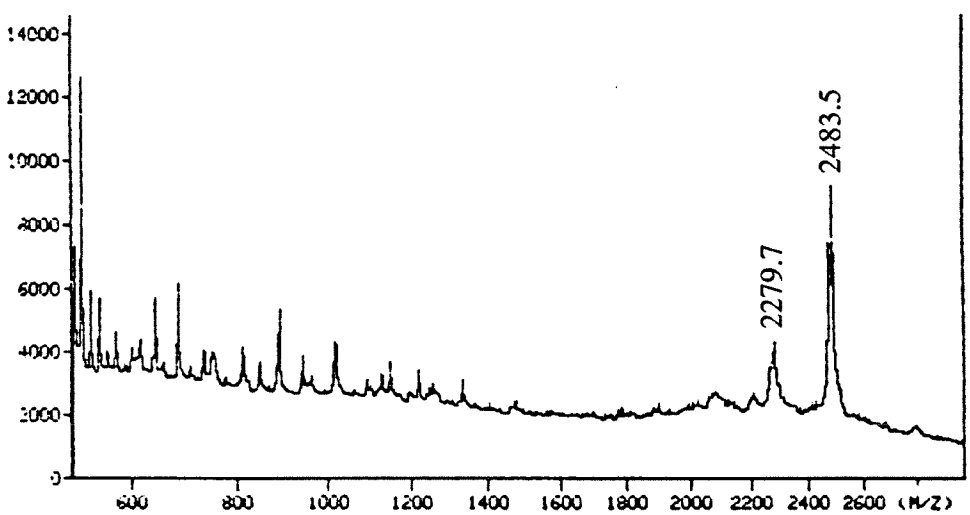


2.6.8. Sk-α_C-9-OMe: [Rh(phi)$_2$(phen'-LQQAIE(OMe)R(Tos)LQNAAAA-COOMe)]$^{3+}$

Synthesis Procedure: The carboxyl groups on glu and the C-terminus of Sk-α_C-9 were methylated by reacting with MeOH in the presence of DMAP and DCC/CH_2Cl_2. A procedure identical to that used for the methylation of Sk-α_A was followed (section 2.2.4). The corresponding methylation reaction of Sk-α_C only regenerated the starting complex.

Mass Spectrometry:

	Coupled Peptide (Sk-α_C-9-OMe)
Avg. Molecular Mass:	2484.3
Monoisotopic Mass:	2482.9
Observed Mass:	2483.5



2.6.9. Sk-α_E: $[Rh(phi)_2(phen'\text{-}ALQQSIAQLQNKAAA\text{-}CONH_2)]^{4+}$

Synthesis Procedure:

(i). Solid phase coordination

(ii). Solid phase direct coupling

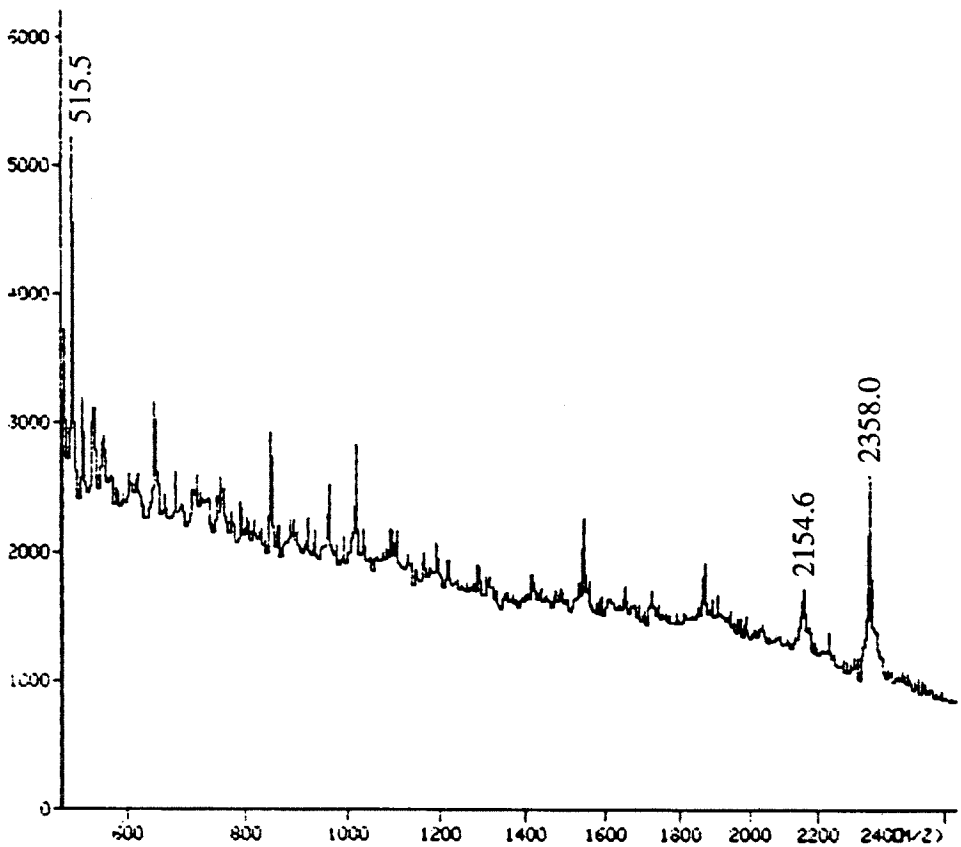
Amino Acid Analysis:

Amino Acid	Expected	Observed (free peptide)	Observed (Sk-peptide)
Asx	1	0.94	1.17
Glx	4	3.90	4.33
Ser	1	1.00	1.04
Ala	5	5.67	5.00
Ile	1	0.96	0.97
Leu	2	2.02	2.06
Lys	1	1.18	1.15

Mass Spectrometry:

	Free Peptide (α_E)	**Coupled Peptide (Sk-α_E)**
Avg. Molecular Mass:	1553.8	2360.4
Monoisotopic Mass:	1552.9	2358.9
Observed Mass:	1554.7	2358.0

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-α_E is shown overleaf.


6000
5000
4000
3000
2000
1000
0
515.5
2154.6
2358.0
600
800
1000
1200
1400
1600
1800
2000
2200
2400(M/Z)

2.6.10. Sk-α_F: [Rh(phi)$_2$(phen'-TQQSKKQLQNKAAA-CONH2)]$^{6+}$

Synthesis Procedure: Solid phase coordination.

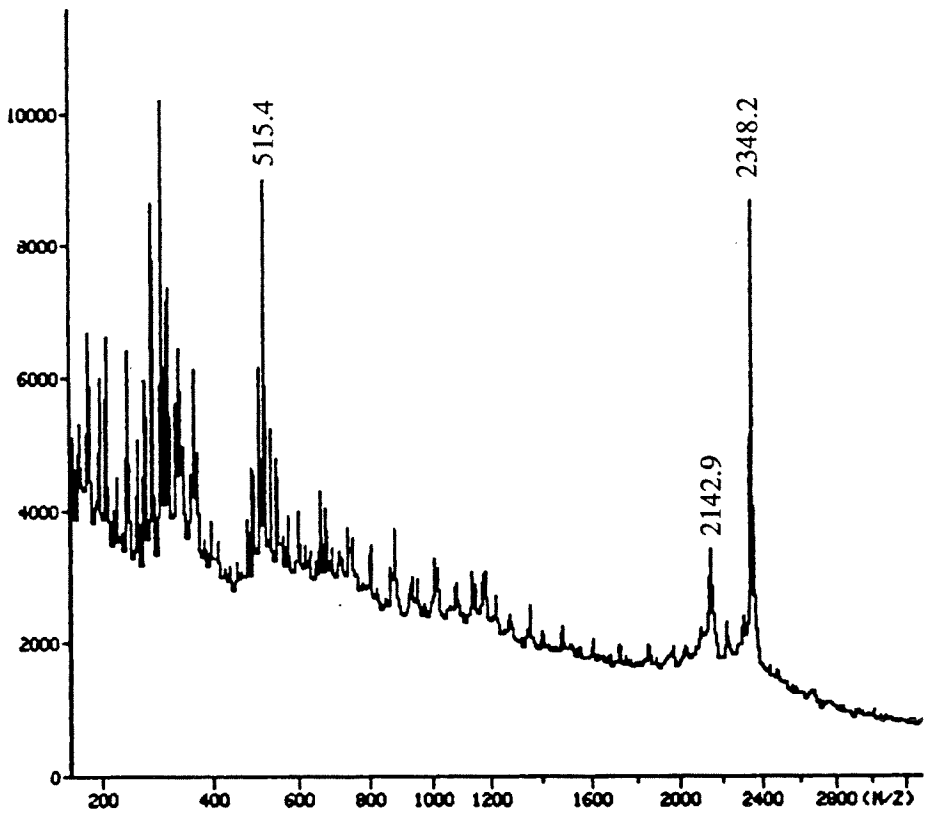
Amino Acid Analysis:

Amino Acid	Expected	Observed (free peptide)	Observed (Sk-peptide)
Asx	1	1.18	0.92
Glx	4	4.00	4.00
Ser	1	1.02	1.20
Thr	1	0.95	1.22
Ala	3	3.48	3.78
Leu	1	1.19	1.49
Lys	3	2.61	3.10

Mass Spectrometry:

	Free Peptide (α_F)	**Coupled Peptide (Sk-α_F)**
Avg. Molecular Mass:	1542.8	2349.4
Monoisotopic Mass:	1541.9	2347.9
Observed Mass:		2348.2

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-α_F is shown overleaf.


10000
8000
6000
4000
2000
0
515.4
2348.2
2142.9
200
400
600
800
1000
1200
1600
2000
2400
2800
(M/Z)

2.6.11. Sk-α_G: [Rh(phi)$_2$(phen'-TQQSKKRLQNKAAA-CONH2)]$^{7+}$

Synthesis Procedure: Solid phase coordination.

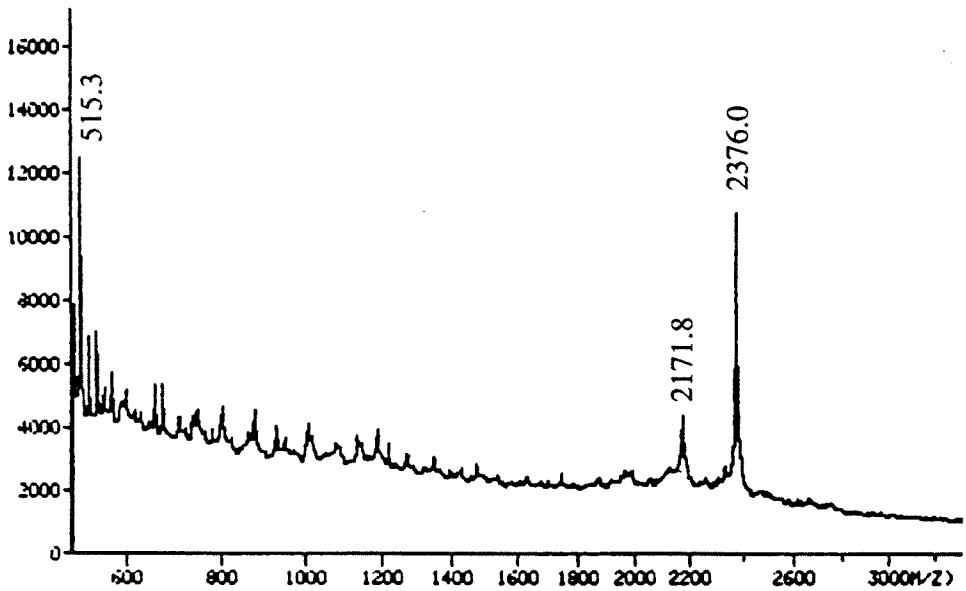
Amino Acid Analysis:

Amino Acid	Expected	Observed (free peptide)	Observed (Sk-peptide)
Asx	1	1.08	0.94
Glx	3	3.00	2.84
Ser	1	1.09	0.95
Arg	1	1.18	1.00
Thr	1	0.89	0.94
Ala	3	3.57	3.00
Leu	1	1.22	1.04
Lys	3	2.92	2.62

Mass Spectrometry:

	Free Peptide (α_G)	**Coupled Peptide (Sk-α_G)**
Avg. Molecular Mass:	1570.8	2377.4
Monoisotopic Mass:	1569.9	2376.0
Observed Mass:		2376.0

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-α_G is shown overleaf.


16000
14000
12000
10000
8000
6000
4000
2000
0
515.3
2171.8
2376.0
600
800
1000
1200
1400
1600
1800
2000
2200
2600
3000M/Z)

2.6.12. Sk-α_H: [Rh(phi)$_2$(phen'-TQRSKKQLQNKAAA-CONH2)]$^{7+}$

Synthesis Procedure: Solid phase coordination.

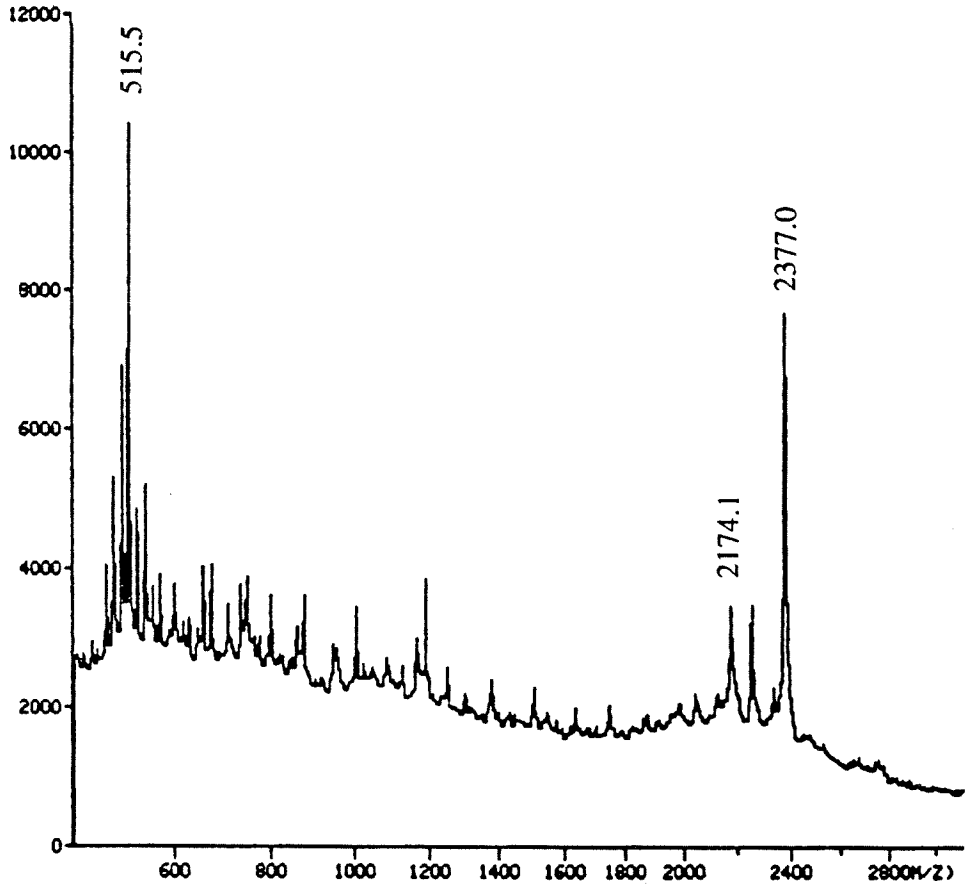
Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Asx	1	0.94
Glx	3	2.86
Ser	1	0.97
Arg	1	1.00
Thr	1	1.12
Ala	3	3.51
Leu	1	1.16
Lys	3	2.75

Mass Spectrometry:

	Free Peptide (α_H)	**Coupled Peptide (Sk-α_H)**
Avg. Molecular Mass:	1570.8	2377.4
Monoisotopic Mass:	1569.9	2376.0
Observed Mass:		2377.0

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-α_H is shown overleaf.


12000
10000
8000
6000
4000
2000
0
515.5
2174.1
2377.0
600
800
1000
1200
1400
1600
1800
2000
2400
2800M/Z)

2.6.13. Sk-α_I: [Rh(phi)$_2$(phen'-TQASKKQLQNKAAA-CONH2)]$^{6+}$

Synthesis Procedure: Solid phase coordination.

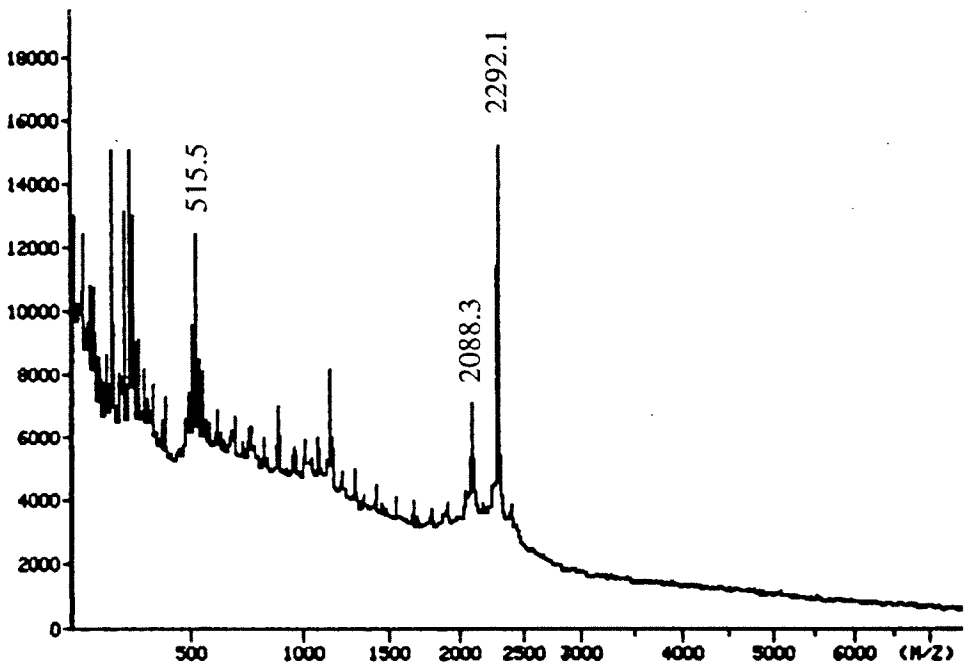
Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Asx	1	1.04
Glx	3	3.00
Ser	1	1.00
Thr	1	0.97
Ala	4	4.93
Leu	1	1.04
Lys	3	2.98

Mass Spectrometry:

	Free Peptide (α_I)	**Coupled Peptide (Sk-α_I)**
Avg. Molecular Mass:	1485.7	2292.3
Monoisotopic Mass:	1484.8	2290.9
Observed Mass:		2292.1

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-α_I is shown overleaf.


18000
16000
14000
12000
10000
8000
6000
4000
2000
0
515.5
2088.3
2292.1
500
1000
1500
2000
2500
3000
4000
5000
6000
(M/Z)

2.6.14. Sk-P22I: $[Rh(phi)_2(phen'\text{-}AANVAIAAWDRAA\text{-}CONH_2)]^{3+}$

Synthesis Procedure: Solid phase coordination.

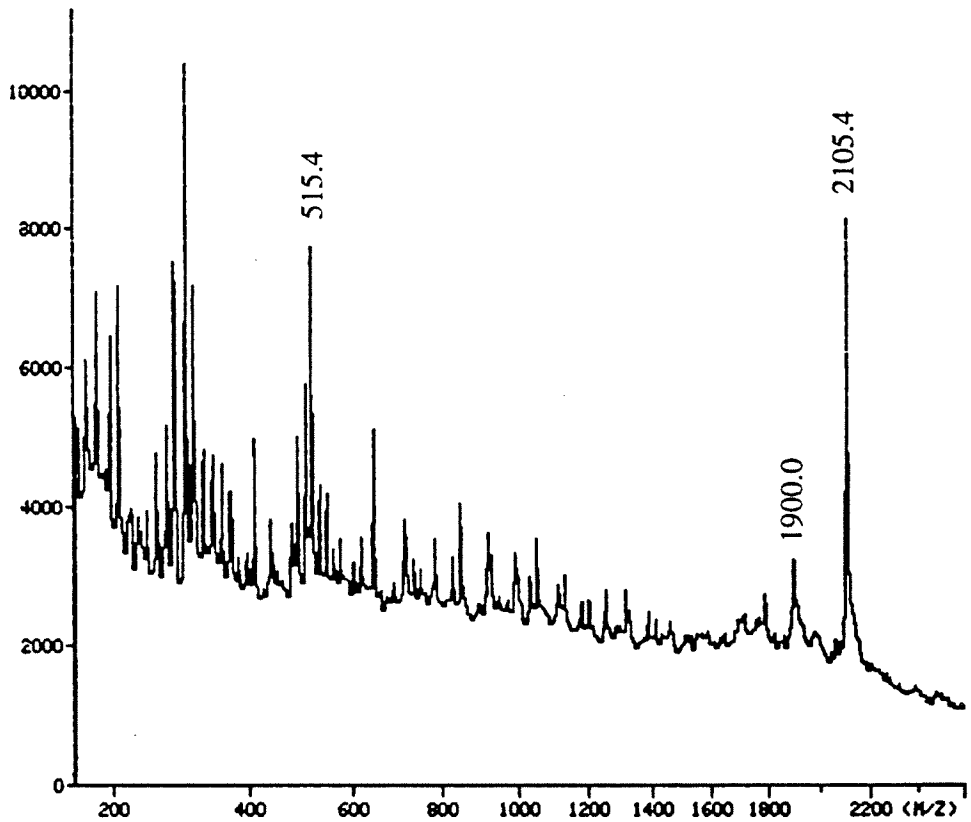
Amino Acid Analysis:

Amino Acid	Expected	Observed (Free peptide)	Observed (Sk-peptide)
Asx	2	1.76	1.82
Glx	0	0.00	0.00
Arg	1	1.25	1.19
Ala	7	7.00	6.56
Val	1	1.06	1.00
Ile	1	1.07	0.97
Trp	1	--	--

Mass Spectrometry:

	Free Peptide (P22I)	**Coup. Peptide (Sk-P22I)**
Avg. Molecular Mass:	1299.5 $(M+H)^+$	2105.1
Monoisotopic Mass:	1298.7 $(M+H)^+$	2103.8
Observed Mass:	1299.5	2105.4

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-P22I is shown overleaf.


10000
8000
6000
4000
2000
0
515.4
1900.0
2105.4
200
400
600
800
1000
1200
1400
1600
1800
2200 (M/Z)

2.6.15. Sk-P22K: $[Rh(phi)_2(phen'\text{-}NVAIAAWDRAA\text{-}CONH_2)]^{3+}$

Synthesis Procedure: Solid phase coordination.

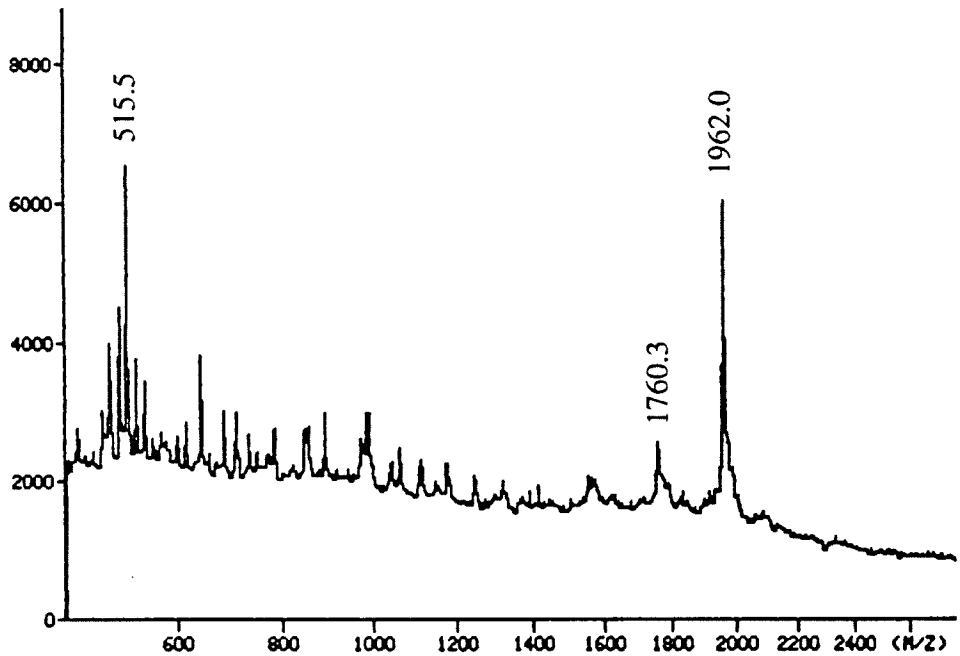
Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Asx	2	1.77
Glx	0	0.00
Arg	1	1.19
Ala	5	5.94
Val	1	1.00
Ile	1	1.03
Trp	1	--

Mass Spectrometry:

	Free Peptide (P22K)	**Coup. Peptide (Sk-P22K)**
Avg. Molecular Mass:	1157.3 $(M+H)^+$	1962.9
Monoisotopic Mass:	1156.6 $(M+H)^+$	1961.7
Observed Mass:		1962.0

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-P22K is shown overleaf.


8000
6000
4000
2000
0
515.5
1760.3
1962.0
600
800
1000
1200
1400
1600
1800
2000
2200
2400
(M/Z)

2.6.16. Sk-P22L: $[Rh(phi)_2(phen'\text{-}AANVAIAAWARAA\text{-}CONH2)]^{4+}$

Synthesis Procedure: Solid phase coordination.

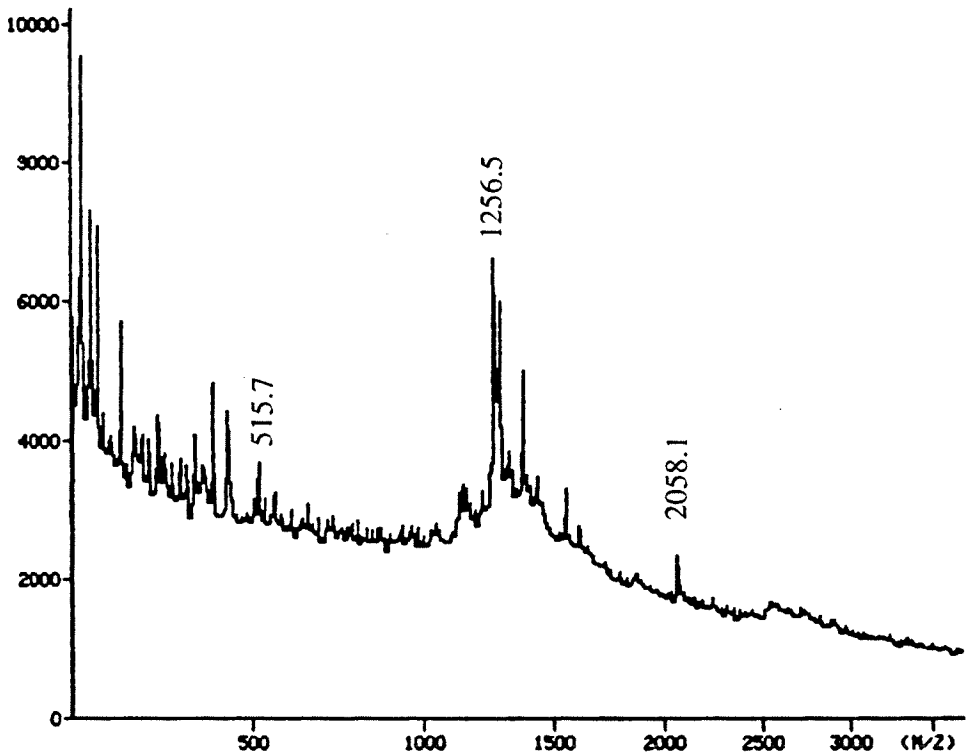
Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Asx	1	1.05
Glx	0	0.00
Arg	1	1.04
Ala	8	8.23
Val	1	1.00
Ile	1	0.99
Trp	1	--

Mass Spectrometry:

	Free Peptide (P22L)	**Coup. Peptide (Sk-P22L)**
Avg. Molecular Mass:	1255.5 $(M+H)^+$	2061.1
Monoisotopic Mass:	1254.7 $(M+H)^+$	2059.8
Observed Mass:	1255.4	2058.1

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-P22L is shown overleaf.


10000
9000
6000
4000
2000
0
500
1000
1500
2000
2500
3000
(M/Z)
515.7
1256.5
2058.1

2.6.17. Sk-P22M: $[Rh(phi)_2(phen'\text{-}AAQAAAQAWERAA\text{-}CONH2)]^{3+}$

Synthesis Procedure: Solid phase coordination.

Amino Acid Analysis:

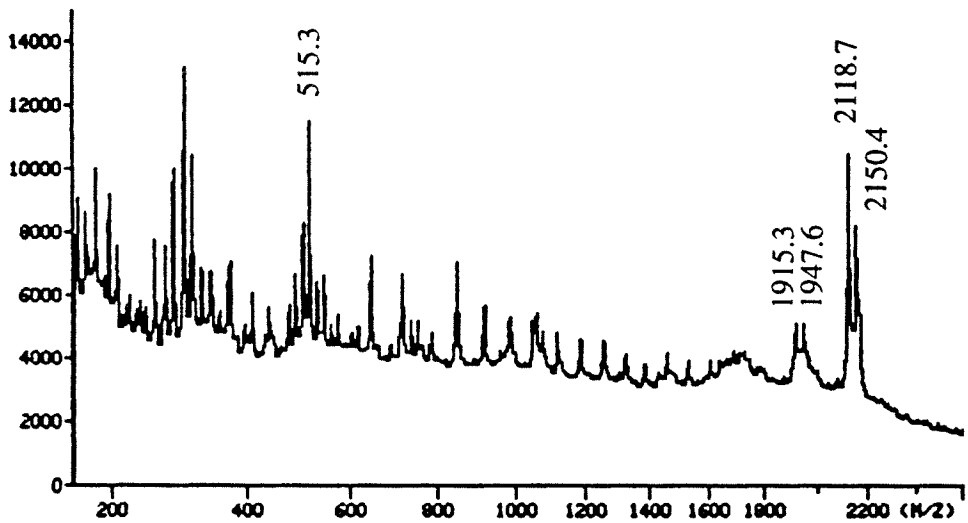
Amino Acid	Expected	Observed (Phen'-peptide)	Observed (Sk-peptide)
Asx	0	0	0.00
Glx	3	2.63	2.88
Arg	1	0.99	1.00
Ala	8	8.00	8.00
Trp	1	--	--

Mass Spectrometry:

	Free Peptide (P22M)	**Coup. Peptide (Sk-P22M)**
Avg. Molecular Mass:	1314.4 $(M+H)^+$	2120.1
Monoisotopic Mass:	1313.7 $(M+H)^+$	2118.7
Observed Mass:	1314.5	2118.7

Additionally, an HPLC peak corresponding to **Phen'-P22M** was isolated. The molecular mass of this peak is: $(M+H)^+$ 1605.4 (Obsd.) [1605.7 (Calc. avg.); 1605.0 (Calc. monoisotopic)].

The usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-P22M is shown overleaf.


14000
12000
10000
8000
6000
4000
2000
0
200
400
600
800
1000
1200
1400
1600
1800
2200 (M/Z)
515.3
1915.3
1947.6
2118.7
2150.4

2.6.18. Sk-L219: $[Rh(phi)_2(phen'\text{-}KAEGAKAFFKRSA\text{-}CONH_2)]^{6+}$

Synthesis Procedure: Solid phase coordination.

The peptides were synthesized by automated FMOC synthesis by Lederle Labs, American Cyanamid. The coupling of the rhodium complex to the N-terminus was carried out as described in Chapter 2.

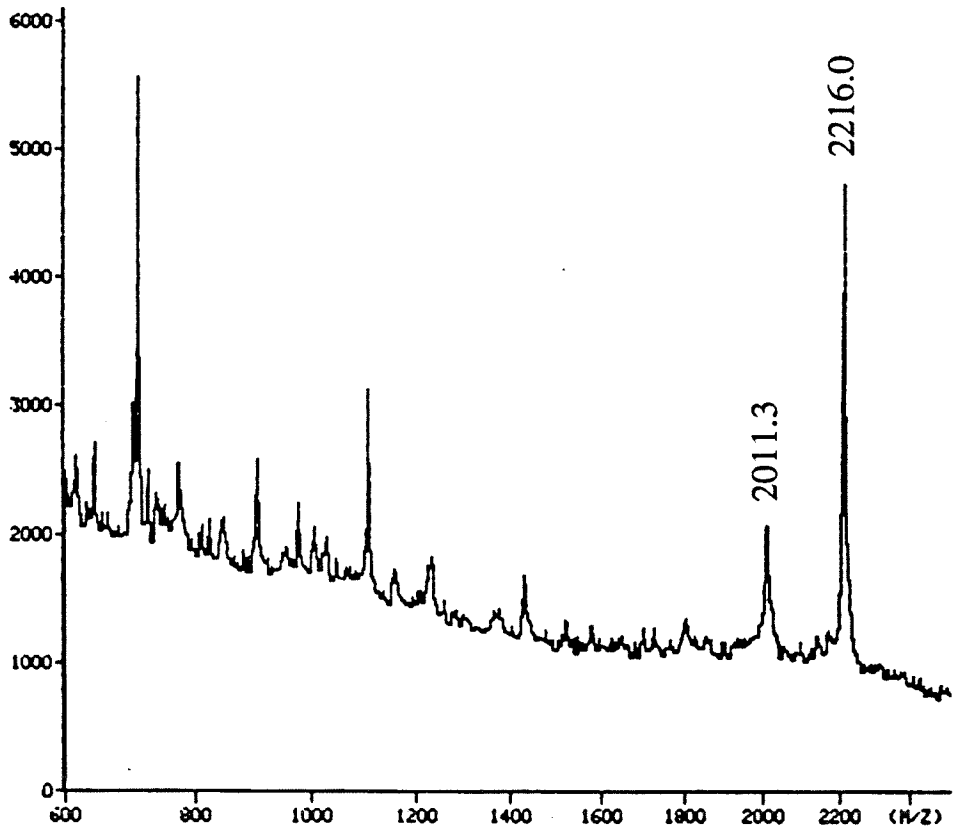
Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Glx	1	1.00
Ser	1	0.92
Gly	1	1.15
Arg	1	1.07
Ala	4	4.10
Phe	2	1.96
Lys	3	2.78

Mass Spectrometry:

	Free Peptide (L219)	**Coupled Peptide (Sk-L219)**
Avg. Molecular Mass:	1411.6 $(M+H)^+$	2217.3
Monoisotopic Mass:	1410.8 $(M+H)^+$	2215.9
Observed Mass:	471.05 $(M+3H)^{3+}$	2216.0

The mass spectrum of the free peptide is by electrospray MS. For the metal-peptide complex, the usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-L219 is shown overleaf.


6000
5000
4000
3000
2000
1000
0
600
800
1000
1200
1400
1600
1800
2000
2200
(M/Z)
2011.3
2216.0

2.6.19. Sk-L220: $[Rh(phi)_2(phen'\text{-}KAAKAFFKRSA\text{-}CONH_2)]^{7+}$

Synthesis Procedure: Solid phase coordination.

The peptides were synthesized by automated FMOC synthesis by Lederle Labs, American Cyanamid. The coupling of the rhodium complex to the N-terminus was carried out as described in Chapter 2.

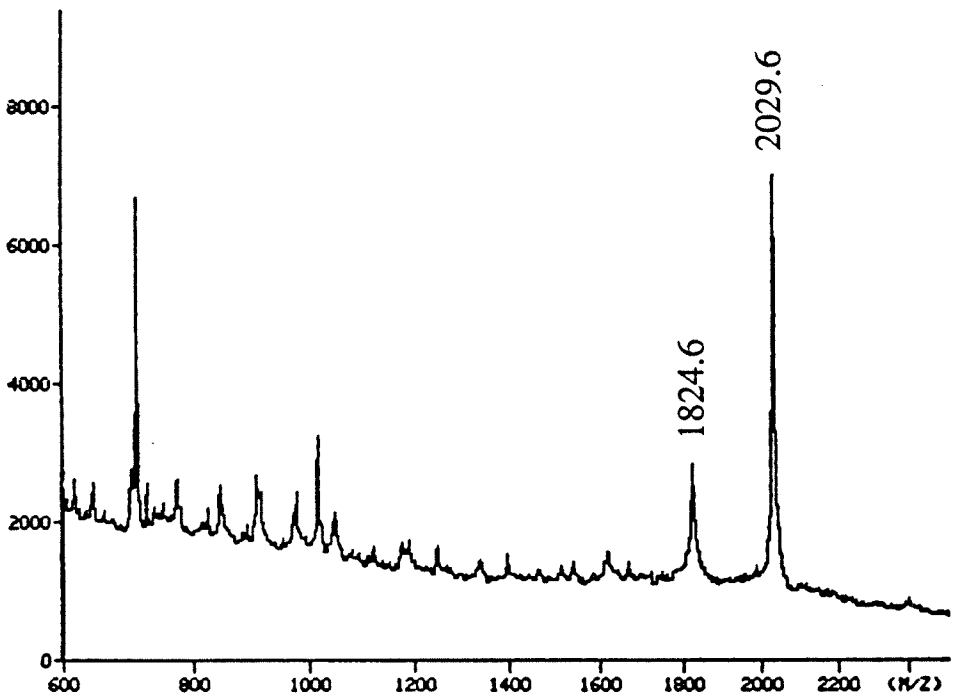
Amino Acid Analysis:

Amino Acid	Expected	Observed (Sk-peptide)
Ser	1	1.00
Arg	1	1.15
Ala	4	4.16
Phe	2	2.09
Lys	3	2.64

Mass Spectrometry:

	Free Peptide (L220)	**Coupled Peptide (Sk-L220)**
Avg. Molecular Mass:	1225.5 $(M+H)^+$	2030.1
Monoisotopic Mass:	1224.7 $(M+H)^+$	2028.8
Observed Mass:	409.05 $(M+3H)^{3+}$	2029.6

The mass spectrum of the free peptide is by electrospray MS. For the metal-peptide complex, the usual sets of pairs of A_n fragments are also observed. A ^{252}Cf PDMS spectrum of Sk-L220 is shown overleaf.


8000
6000
4000
2000
0
600
800
1000
1200
1400
1600
1800
2000
2200
(M/Z)
1824.6
2029.6

Chapter 3:
DNA Recognition by Peptide Complexes of Rhodium (III): Example of a Glutamate Switch†

3.1. Introduction

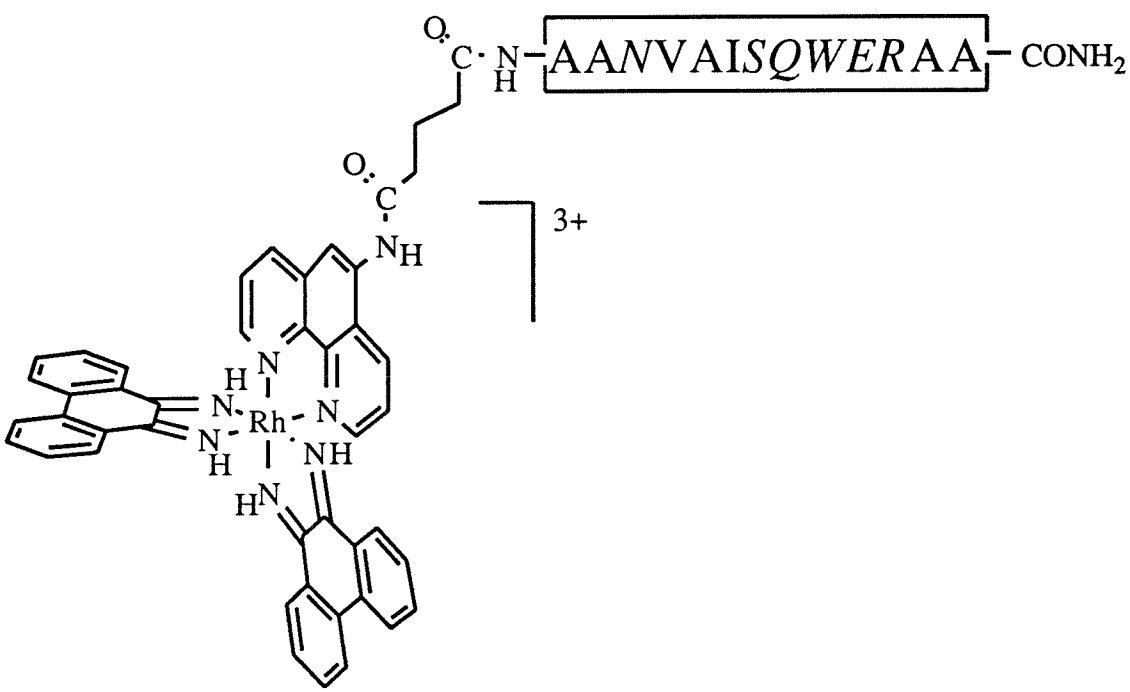
DNA recognition by small molecules and large peptides has evoked considerable interest in the last decade to establish essential elements for site-specificity.[1,2] The Barton laboratory has focused on the construction of coordinatively saturated octahedral metal complexes for site-specific DNA recognition.[3-6] In particular, 9,10-phenanthrene-quinonediimine (phi) complexes of rhodium(III) have been prepared which bind avidly through intercalation in the DNA major groove[4] ($K_a \geq 10^6\ M^{-1}$) and upon photoactivation promote strand scission via abstraction of the deoxyribose C3'-hydrogen atom.[5] Related work in the Barton group has involved the assembly of complexes which, like DNA binding proteins, pose several functionalities for site-specific, non-covalent interaction with DNA.[6] This chapter describes the design and DNA-recognition properties of metal-peptide complexes (Figure 3.1) based upon attachment of short oligopeptides (13 residues) onto the metallointercalating skeletal complex, $[Rh(phi)_2(phen')]^{3+}$ (Sk, phen' = 5-amidoglutaryl-1,10-phenanthroline). Our interest is in exploring whether the side chain functionalities of small peptides may be used to augment metal complex recognition.

Studies of peptide recognition of DNA have focused mainly on dimeric DNA binding domains of proteins. Different strategies to dimerize α-helical peptides have included the incorporation of leucine zippers or adamantane-β-cyclodextrin tethers in non-covalent dimeric assemblies[7] and disulfide or modified $Fe(trpy)_2^{3+}$ or biphenyl linkages in covalent dimers.[2] In each case, specific DNA binding has required dimerization of the peptides, likely because the monomeric peptides lack the non-specific

---

† Adapted from Sardesai, N. Y.; Zimmermann, K.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 7502-7508.

Figure 3.1. Representative metal-peptide complex, Sk-PA, showing linkage of the peptide to the skeletal complex. The italicized residues have been substituted in the family of metal-peptide complexes synthesized.


O
C
N
H
AAANVAISQWERAA
CONH2
O
C
NH
3+
N
H
N
Rh
N
N
H
NH
HN

binding affinity that is present in the intact protein. Several examples of tethering DNA cleaving moieties such as Fe-EDTA[2e] or [(1, 10-phenanthroline)Cu(I)][8] onto DNA binding proteins have also been reported in an effort to engineer new sequence-specific nucleases; this approach has been used elegantly to elucidate the binding modes and relative orientations of the DNA binding proteins.[9] However, an important distinction in all these examples is that the peptide dimers or protein fragments have been large (> 20 residues / monomer) and exhibit specific DNA affinity that is independent of the attached cleaving functionality.

The rationale behind the design of small metal-peptide complexes is the application of $[Rh(phi)_2(phen)]^{3+}$ as a delivery system for an array of peptide functionalities to the major groove of DNA. This design takes advantage of the octahedral metallointercalator as a small molecular anchor in the major groove.[4,5] The skeletal complex provides DNA binding affinity while the attached peptide functionalities may be varied to achieve site-specific DNA recognition. By varying systematically the sequence of short, monomeric peptides, the targeting of a variety of DNA sequences may be explored. The DNA photocleavage chemistry of the skeletal complex[5] also serves as a convenient tag for assaying DNA site selectivity. This chapter describes our investigation of the DNA recognition properties of the α_3-helix of the DNA binding phage P_{22} repressor[10] in isolation from the rest of the protein. The helix is small (8 residues) and contains an array of side-chain functionalities; the understanding of P_{22} targeting is unclear. A family of metal-peptide complexes was therefore constructed, and their DNA binding characteristics were examined. Remarkably, a single glutamate is found to be essential in directing site-specific recognition in these complexes. A model for DNA recognition by these metal-peptide complexes is presented in this chapter and further experiments to validate the model are described.

3.2. Experimental

3.2.1. Materials. Sonicated calf thymus DNA was purchased from Pharmacia, plasmid pUC 18 was purchased from Boehringer-Mannheim and all enzymes utilized were from commercial sources. [α-^{32}P]dATP and [γ-^{32}P]ATP were obtained from NEN-Dupont. Oligonucleotides were synthesized by the phosphoramidite method,[11] using 1.0 μmol columns on an ABI 391 DNA-RNA synthesizer.

3.2.2. Instrumentation. Absorption spectra and quantitation of the metal complexes were recorded on a Cary 219 spectrophotometer. HPLC was carried out on a Waters 600E system equipped with a Waters 484 tunable detector. All photocleavage experiments were carried out with an Oriel Model 6140 1000 W Hg/Xe lamp fitted with a monochromator and a 300 nm cut-off filter to avoid light damage to DNA. CD experiments were performed on a Jasco J-600 spectrometer. Gel electrophoresis experiments were quantified using a Molecular Dynamics phosphorimager and ImageQuant software. Computer models of the docked metal-peptide complexes in DNA were generated on a Silicon Graphics Indigo work station using Macromodel and InsightII programs.

3.2.3. Synthesis of the Metal-Peptide Complexes. The details of the synthesis and characterization of the metal-peptide complexes used in this study have been described in Chapter 2 and reference 12.

3.2.4. Photocleavage of DNA Restriction Fragments. Plasmid pUC-18 was digested with *EcoR I* restriction endonuclease. The digested plasmid was 3'-end-labeled by treatment with Klenow fragment of the DNA polymerase I and [α-^{32}P]dATP, dTTP, dCTP and dGTP.[13] A separate batch of the digested plasmid was treated with calf intestinal alkaline phosphatase and 5'-end labeled with T4 polynucleotide kinase and

[γ-^{32}P]ATP.[13] After labeling, the DNA was digested with *Pvu II* to yield a 180 and a 140 base pair fragment. The 180 base pair fragment was isolated by 8% nondenaturing preparative polyacrylamide gel electrophoresis followed by electroelution.

Cleavage reactions were carried out in 20 μL total volume contained in 1.7 mL presiliconized eppendorf tubes. The reaction mixtures contained 50 μM (nucleotide) calf thymus DNA, labeled restriction fragment (~ 40,000 cpm per sample) and 1.0 μM rhodium complex in 50 mM sodium cacodylate, pH 7.0. Reaction mixtures were incubated in the presence (or absence) of 5 mM $MnCl_2$ at 55° C (or room temperature) for 5 min and then irradiated at the same temperature at 313 nm as indicated. Controls without metal complex were also irradiated in parallel to test for light damage and cleavage by Mn^{2+}.

All samples were ethanol precipitated after irradiation by addition of 10 μL of 7.5 M NH_4OAc and 130 μL of ethanol. The precipitated DNA was dried and resuspended in 3 μL of 80% formamide loading buffer. The samples along with Maxam-Gilbert G+A and C+T sequencing reactions[13, 14] were loaded on to an 8% denaturing polyacrylamide gel and electrophoresed at 1600 V for approximately 135 min. The gel was transferred to paper and dried prior to autoradiography.

3.2.5. HPLC Analysis of Photocleavage Products. Reaction samples (40 μL total volume) were made to contain 1.0 mM (in base pairs) of a 32 base pair oligonucleotide duplex (5'-TGACTTTAAAGGTACCAATATTCCTAGGCAGT-3' and the complementary strand) in 50 mM ammonium formate. The samples were incubated in the presence or absence of 50 μM Sk-PG for 5 min. and irradiated at 313 nm for 10 min. at ambient temperature. After irradiation, the samples were analyzed by HPLC on a Cosmosil 5 μ, 15 cm C-18 reverse phase column under isocratic elution condition of 50 mM ammonium formate (flow rate 1.5 mL/min). The products were detected at 260 nm and 300 nm.[5]

3.2.6. Photocleavage of Oligonucleotides. A 31 base pair oligonucleotide with the sequence 5'-GCCACGAGCCACGACGAGCCACGAACGAGCC-3' was 5'-end labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP. The cold single strands corresponding to the 31 base-pair duplex and the labeled strand were annealed by heating up to 90° C followed by slow cooling to room temperature. Reaction samples contained 50 μM base pair oligonucleotide (including ~120,000 cpm of labeled oligonucleotide), 1.0 μM rhodium complex and 5 mM $MnCl_2$ in 50 mM sodium cacodylate, pH 7.0. Samples were incubated at 55° C for 5 min and irradiated at 55° C for 15 min as before. Approximately 20,000 cpm/sample were transferred into fresh eppendorf tubes, dried and then resuspended in 3 μL loading buffer. The samples were loaded onto a 20% denaturing polyacrylamide gel and electrophoresed at 2000 V for 3 h 15 min. The gel was then transferred to a film and stored at -70° C during autoradiography.

In another series of experiments, a 30 bp oligonucleotide duplex having the sequence, 5'-GCGCGATACAAGAAGCAGTTTGTATCCGCG-3', was [5'-^{32}P]-end-labeled as before and annealed to its complementary strand. Photocleavage reaction samples contained 0.1 μM oligonucleotide duplex (including ca. 110,000 cpm of labeled oligonucleotide) and 0.1 μM rhodium complex in 50 mM sodium cacodylate, pH 7.0 buffer. The samples were incubated for 5 min at 55° C and irradiated (313 nm) for 15 min at 55° C. Approximately 17,000 cpm/sample were dried and then resuspended in 3 μL loading buffer. The samples were loaded onto a 20% denaturing polyacrylamide gel, electrophoresed at 2200 V for 3 h 30 min, and autoradiographed as above.

3.2.7. Circular Dichroism Studies. Samples of the metal peptide complexes were prepared in 10 mM sodium phosphate or 10 mM Tris-HCl, pH 7.0. CD spectra were recorded in a 1.0 or 0.1 cm path length cell at room temperature and at 55° C using a temperature controlled water circulating bath. The effect of temperature and presence of 5 mM Mn^{2+} was measured over the range 350-200 nm.

3.3. Results

3.3.1. DNA Cleavage by a Metal-Peptide Complex. Photocleavage of DNA restriction fragments by the first member of the family of metal-peptide complexes, Sk-PA, was investigated under varying reaction conditions. As shown in Figure 3.2, Sk-PA promotes DNA strand scission upon photoactivation similar to the skeletal complex lacking the peptide. The efficiency of photocleavage is comparable to that for other phi complexes of rhodium.[5] Importantly, highly specific cleavage is apparent at the site 5'-CCA-3', with cleavage at the highlighted A. Similar specific cleavage is not evident with the skeletal complex lacking the peptide. This site-specificity is enhanced upon the addition of reagents such as trifluoroethanol or in the presence of divalent cations (Mn^{2+}, Mg^{2+}) at elevated temperature (55° C).

3.3.2. Site Selectivity on DNA Restriction Fragments by the Family of Metal-Peptide Complexes. Photocleavage of DNA with Sk-PA indicates that the peptide is essential in directing the metal complex to recognize the 5'-CCA-3' site. In order to identify the residues involved in this selectivity, residues with functional side chains were mutated to alanine individually or in pairs. The family of metal-peptide complexes thus constructed, as shown in Table 3.1, includes Sk-PB (N3A), Sk-PC (R11A), Sk-PD (S7A / Q8A) and Sk-PC-OMe (E10-methyl ester). We also synthesized Sk-PE, Sk-PF and Sk-PG to ascertain the role of Glu10 and Trp9 in recognition. These three complexes have, in addition, two lysine substitutions which we envisioned would increase the overall binding affinity of the metal-peptide complexes.

All the metal-peptide complexes promote DNA cleavage when photoactivated, as shown in Figure 3.3. A subset of the complexes show high specificity for 5'-CCA-3' sites, with cleavage at the highlighted adenosine. The cleavage is also severely canted to one strand with no specific cleavage being detected across the site on the complementary strand. Strong specific cleavage at any other 5'-pyr-pyr-pur-3' sequence, for example,

Figure 3.2. Autoradiogram of an 8% denaturing polyacrylamide gel showing photocleavage of the 180 base-pair [5'-^{32}P]-end-labeled *EcoR I / Pvu II* fragment of plasmid pUC18 by Sk and Sk-PA in the presence of various reagents. Untreated fragment (lane 1). Fragment irradiated in the absence of rhodium complex (lane 2). Fragment irradiated in the presence of 1 μM Sk and 5 mM Mg^{2+} (lane 3), 5 mM Mn^{2+} (lane 6) or 10 vol% trifluoroethanol (lane 9). Fragment irradiated in the presence of 1 μM Sk-PA and 5 mM Mg^{2+} (lane 4), 10 mM Mg^{2+} (lane 5), 5 mM Mn^{2+} (lane 7), 10 mM Mn^{2+} (lane 8), 10 vol% trifluoroethanol (lane 10) or 25 vol% trifluoroethanol (lane 11). Maxam-Gilbert A+G sequencing reaction (lane 12). The samples containing Mg^{2+} or Mn^{2+} were incubated and irradiated at 55° C, while those containing trifluoroethanol were at room temperature. Incubation time in every case was 5 minutes and the samples were irradiated for 20 minutes.

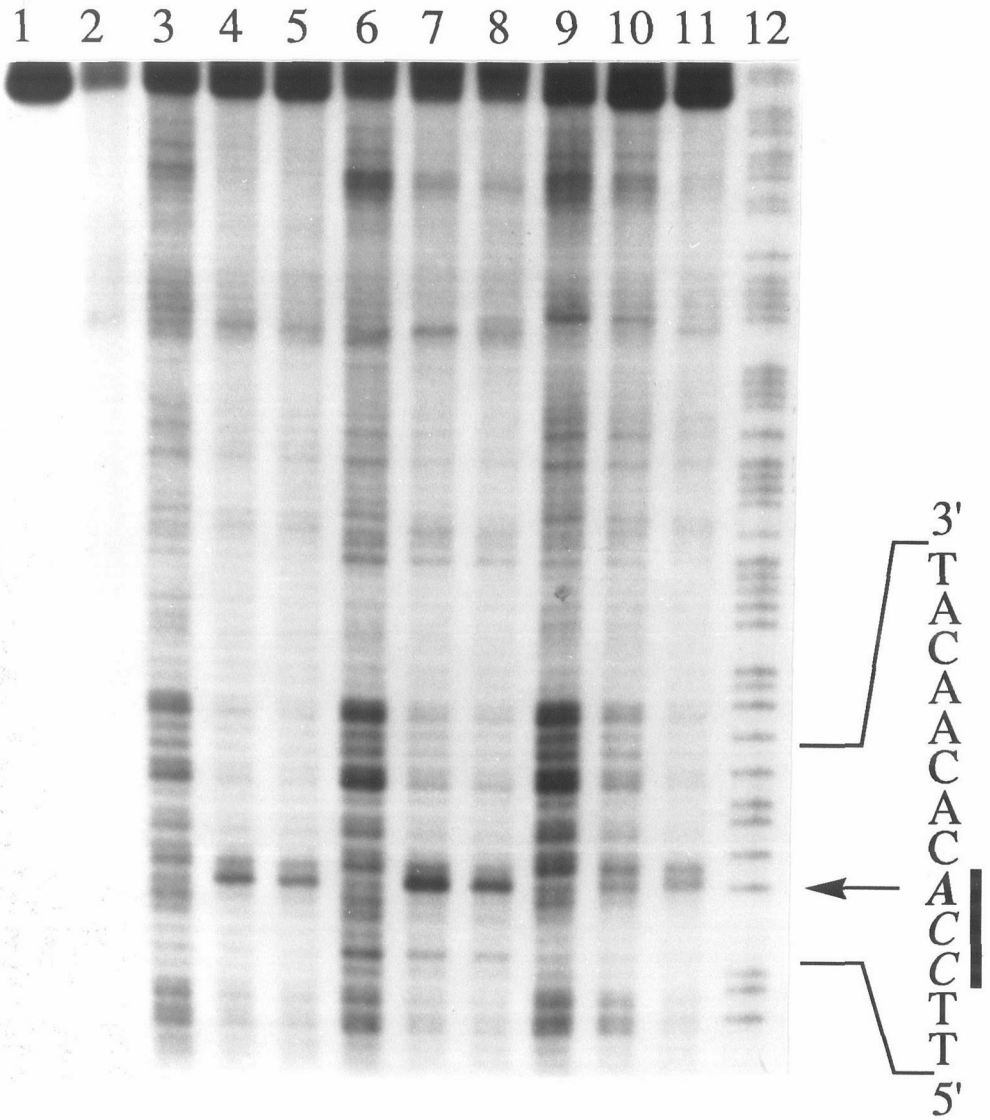

1 2 3 4 5 6 7 8 9 10 11 12
3'
TACAACACACCTT
5'

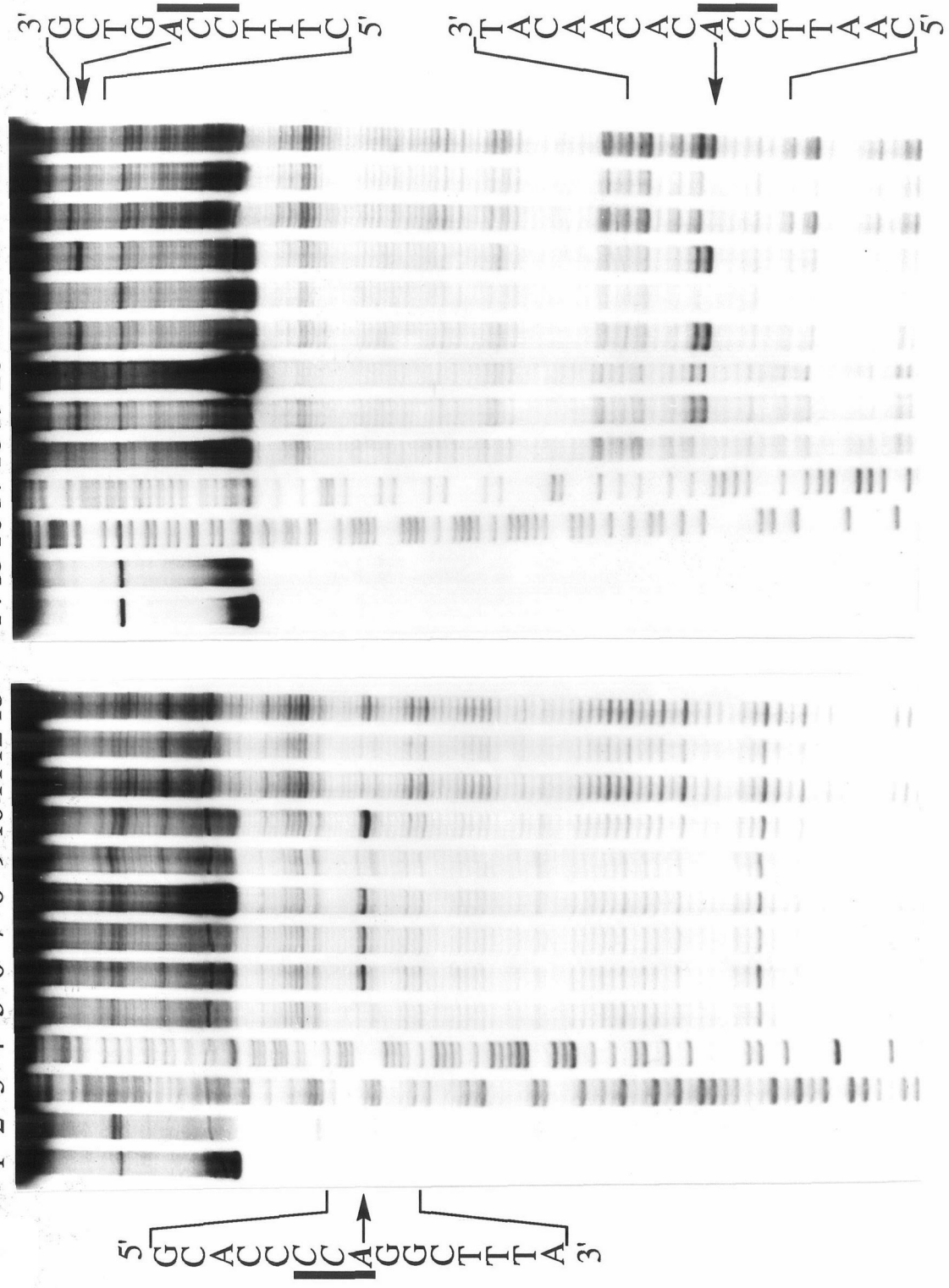
Table 3.1. Family of Metal-Peptide Complexes and Their DNA Recognition Sequences.

Metal Complex	Alternate Name[a]	Metal-Peptide Sequence[b, c]	DNA Site[d]
Sk[e]	Sk	[Sk- COO^-]$^{2+}$	-
Sk-PA	Sk-P22A	[Sk- AANVAISQW**E**RAA-$CONH_2$]$^{3+}$	5'-CCA-3'
Sk-PB	Sk-P22B	[Sk- AAAVAISQW**E**RAA-$CONH_2$]$^{3+}$	5'-CCA-3'
Sk-PC	Sk-P22C	[Sk- AANVAISQW**E**AAA-$CONH_2$]$^{2+}$	5'-CCA-3'
Sk-PC-OMe[f]	Sk-P22COMe	[Sk- AANVAISQW$\underline{E}$AAA-$CONH_2$]$^{3+}$	-
Sk-PD	Sk-P22E	[Sk- AANVAIAAW**E**RAA-$CONH_2$]$^{3+}$	5'-CCA-3'
Sk-PE	Sk-P22F	[Sk- AAKVAISQKQRAA-$CONH_2$]$^{6+}$	-
Sk-PF	Sk-P22G	[Sk- AAKVAISQWKRAA-$CONH_2$]$^{6+}$	-
Sk-PG	Sk-P22H	[Sk- AAKVAISQK**E**RAA-$CONH_2$]$^{5+}$	5'-CCA-3'
Sk-PH	Sk-P22I	[Sk- AANVAIAAWDRAA-$CONH_2$]$^{3+}$	-
Sk-PI	Sk-P22K	[Sk- --NVAIAAWDRAA-$CONH_2$]$^{3+}$	-
Sk-PJ	Sk-P22L	[Sk- AANVAIAAWARAA-$CONH_2$]$^{4+}$	-
Sk-PK	Sk-P22M	[Sk- AAQAAAQAW**E**RAA-$CONH_2$]$^{3+}$	5'-CAAN-3', 5'-CCA-3'

(Table 3.1 contd.)

[a] The metal-complex denotes the sequential naming of the metal-peptide complexes studied and referred throughout this chapter and in Sardesai, N. Y.; Zimmermann, K.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 7502-7508. The alternate name denotes the nomenclature followed in Chapter 2 and reference 12 describing the synthesis and characterization of the metal-peptide complexes. [b] The skeletal complex has a free carboxylate. The carboxy terminus of all the peptides is present as an amide. The overall charge shown on each metal-peptide complex is that expected at pH 7.0 assuming that the imine-nitrogen atoms on the phi ligands are fully protonated. [c] When present, the essential glutamate at position 10 is indicated in bold. [d] The primary photocleavage site is indicated in bold. [e] Sk denotes the skeletal complex, $[Rh(phi)_2(phen')]^{3+}$. [f] The underlined glutamate has its side chain carboxylate as a methyl ester.

Figure 3.3. DNA photocleavage by the family of metal-peptide complexes. Autoradiograms showing cleavage reactions on the 180 base-pair fragment [3'-^{32}P]- (lanes 1-13) or [5'-^{32}P]- (lanes 14-26) end-labeled at the *EcoR I* site. The reaction samples were incubated in the presence of 5 mM $MnCl_2$ for 5 minutes and irradiated (313 nm) for 10 minutes at 55° C. Lanes 1 and 14: DNA control. Lanes 2 and 15: fragment irradiated in the absence of rhodium complex. Lanes 3 and 16: Maxam-Gilbert A+G sequencing reaction. Lanes 4 and 17: Maxam-Gilbert C+T sequencing reaction. Lanes 5-13 and 18-26: fragment irradiated in the presence of 1μM of Sk, Sk-PA, Sk-PB, Sk-PC, Sk-PC-OMe, Sk-PD, Sk-PE, Sk-PF and Sk-PG, respectively. The DNA recognition sequence is shown in italics. The cleavage site is in bold and indicated by an arrow. Note that the 5'-CCA-3' recognition sequence is preferentially cleaved by the family of metal-peptide complexes except by Sk-PC-OMe, Sk-PE and Sk-PF, all of which lack a free carboxylate at position 10.


3' GCTGACCTTTC 5'
3' TACAACAACACCTTAAC 5'
14 15 16 17 18 19 20 21 22 23 24 25 26
1 2 3 4 5 6 7 8 9 10 11 12 13
5' GCACCCCCAGGCTTTA 3'

5'-TCA-3', 5'-CCG-3', or 5'-TCG-3', is not apparent. Importantly, as illustrated in Table 3.1, *the site-specificity requires glutamic acid at position 10 in the peptide.* Transposing the glutamate to position 6 in the context of a different metal-peptide complex does not afford selectivity for 5'-CCA-3' (see examples in Chapter 4). Methylation of the glutamate side chain, as with Sk-PC-OMe, or a direct E10Q substitution, as in Sk-PG and Sk-PE, abolishes the 5'-CCA-3' selectivity. Additionally, the skeletal complex, which also has a free terminal carboxylate, does not target 5'-CCA-3', implying that a carboxylate in the vicinity of the DNA helix is not sufficient to impart specificity.

Substitutions at other residues with functional side chains that may bind to DNA[15, 16] do not alter the site-selectivity appreciably. These substitutions perturb the intensity of cleavage at the 5'-CCA-3' site, which may correlate with structural variations that arise with the mutations. Of particular note is Trp9 which does enhance specificity. Sk-PG, which lacks the tryptophan but contains the glutamate, targets 5'-CCA-3', but more moderately. However, tryptophan alone, without the neighbouring glutamate (as with Sk-PF), is not sufficient for weak recognition of the target sequence. These peptide complexes therefore share a distinct DNA recognition motif, governed by a single specificity determinant, but are otherwise capable of tolerating a variety of substitutions.

It should be noted here that the consensus half-site sequence for binding of the P22 repressor protein is 5'-ANTNAAG-3'. The 5'-CCA-3' specificity observed for the metal-peptide complexes is, therefore, different from that predicted for P22R. However, weak cleavage is observed at 5'-CATAAAG-3' which has the putative P22R operator sequence and is strongest for Sk PD.

3.3.3. Differences Between Aspartate and Glutamate at Position 10. To investigate further the role of the glutamate at position 10, we synthesized an E10D mutant of Sk-PD, Sk-PH. Sk-PD shows the strongest recognition of the 5'-CCA-3' sequence amongst this family of metal-peptide complexes, and hence we believed that this complex

would be the most tolerant of minor variations in peptide sequence. Figure 3.4 shows a direct comparison of Sk, Sk-PD and Sk-PH (Sk-PD-E10D) under different reaction conditions. The autoradiogram clearly shows the recognition to be contingent on the presence of a glutamate. Even a very conservative substitution to an aspartate is not tolerated.

Not surprisingly, Sk-PI, a truncated version of Sk-PH with the N-terminal two residues deleted and containing an E10D mutation does not recognize the 5'-CCA-3' sequence either. Furthermore, no other alternate sites specific to this metal-peptide complex are observed. Thus it is not sufficient to merely have a carboxylate in the vicinity of DNA to achieve the observed recognition, but rather the positioning and the sequence context of the carboxylate is key in selecting the DNA sequence.

It should be pointed out here that Sk-PD shows modest 5'-CCA-3' selectivity even at room temperature in the absence of 5 mM Mn^{2+} in contrast to Sk-PA. Temperature and Mn^{2+} ions increase the cleavage individually, while photocleavage in the presence of Mn^{2+} at elevated temperature leads to the strongest selectivity. This observation suggests that the peptide structure, dictated by the sequence, is important for recognition.

It is noteworthy also that free peptide, untethered to the metal complex, does not bind comparably, even at much higher concentrations than the metal-peptide complex. In competition experiments with free peptide, it is found that 100 μM PD does not inhibit binding of 1.0 μM Sk-PD (Figure 3.4, lane 1). Also, in control experiments, PD in the presence of 5 mM Mn^{2+} did not cleave DNA upon irradiation.

3.3.4. Photocleavage of Oligonucleotides by Metal-Peptide Complexes. HPLC analysis of the photocleavage of an oligonucleotide duplex substrate by Sk-PG shows the production of base propenoic acid and free nucleic acid base products (Figure 3.5). Peaks eluting at 1.4 min (base propenoic acids), 2.0 min (cytosine), 4.0 min (guanine) and 4.7 min (thymine) under the HPLC conditions, were observed. The production of the base

Figure 3.4. Photocleavage of the [5'-^{32}P]- end-labeled 180-mer DNA restriction fragment by Sk, Sk-PD and Sk-PH comparing E^{10} and D^{10} in the presence or absence of 5 mM $MnCl_2$ and elevated temperature. Lanes 1-3, 7-9, 13-15 contained 5 mM $MnCl_2$. Lanes 1-3 and 10-15 were incubated and irradiated at 55° C, while lanes 4-9 were treated at r.t. The reaction samples were incubated for 5 minutes and irradiated (313 nm) for 8 minutes at the appropriate temperature in the presence of 1.0 μM Sk (lanes 4, 7, 10, 13); 1.0 μM Sk-PD (lanes 1, 5, 8, 11, 14) or 1.0 μM Sk-PH (lanes 6, 9, 12, 15). Untreated fragment is shown in lane 2 and fragment irradiated in the absence of rhodium complex is shown in lane 3. Lane 1 shows cleavage by Sk-PD in the presence of 100 μM PD (free peptide) that serves as a competitor for Sk-PD binding to DNA. The 5'-CCA-3' cleavage sites are indicated on the side.

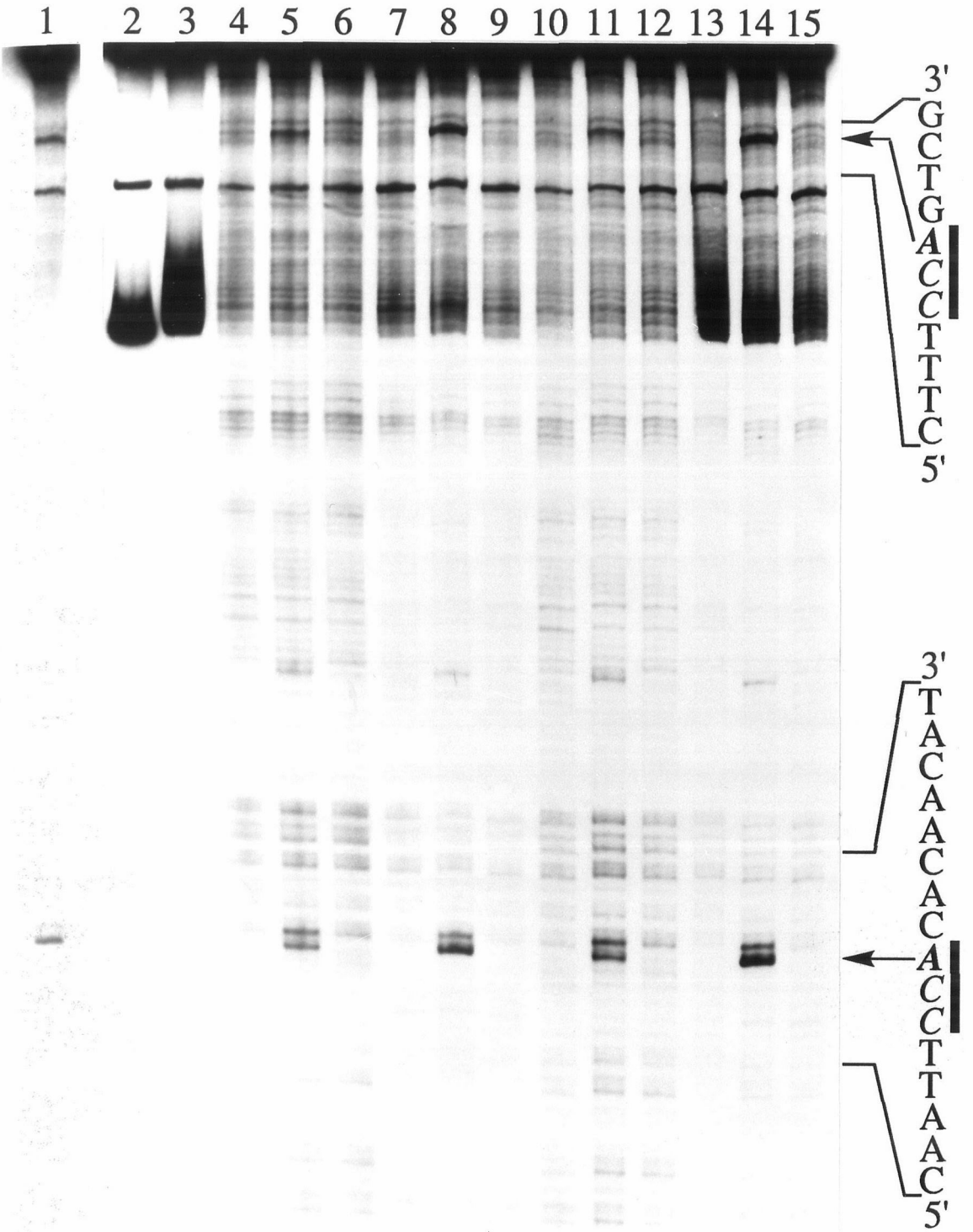
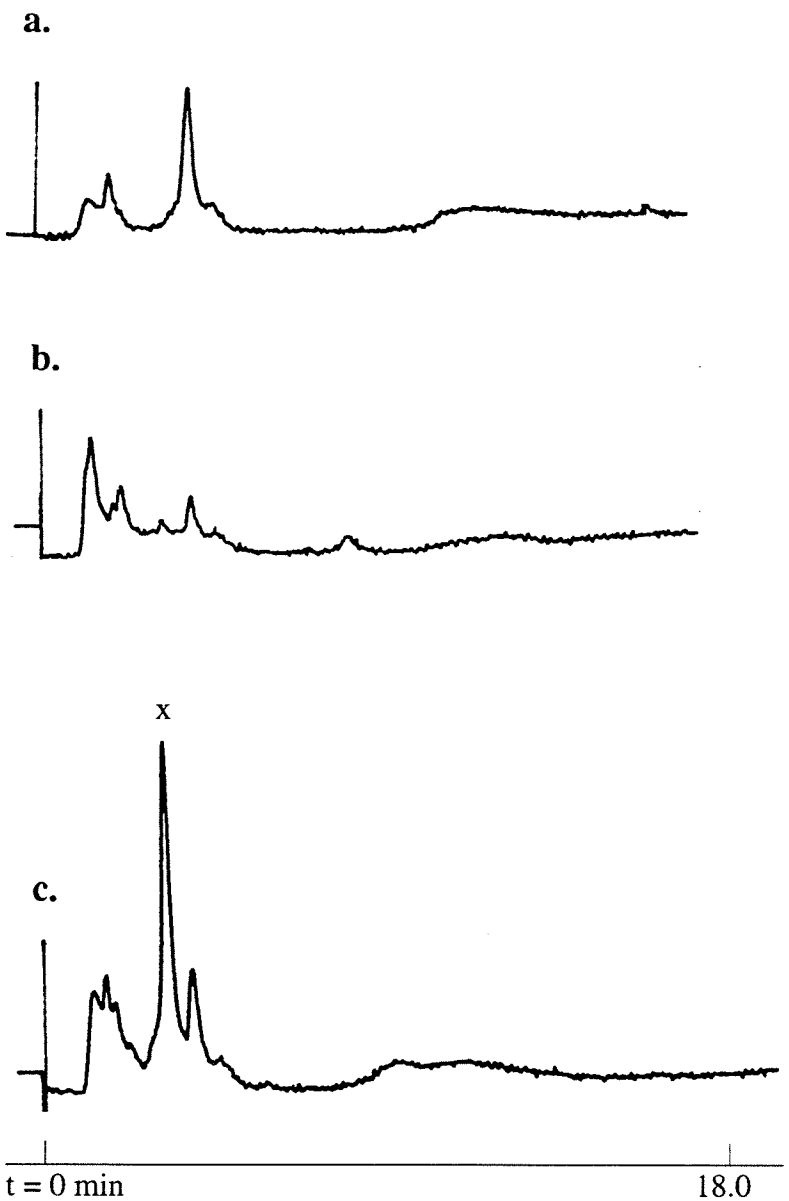
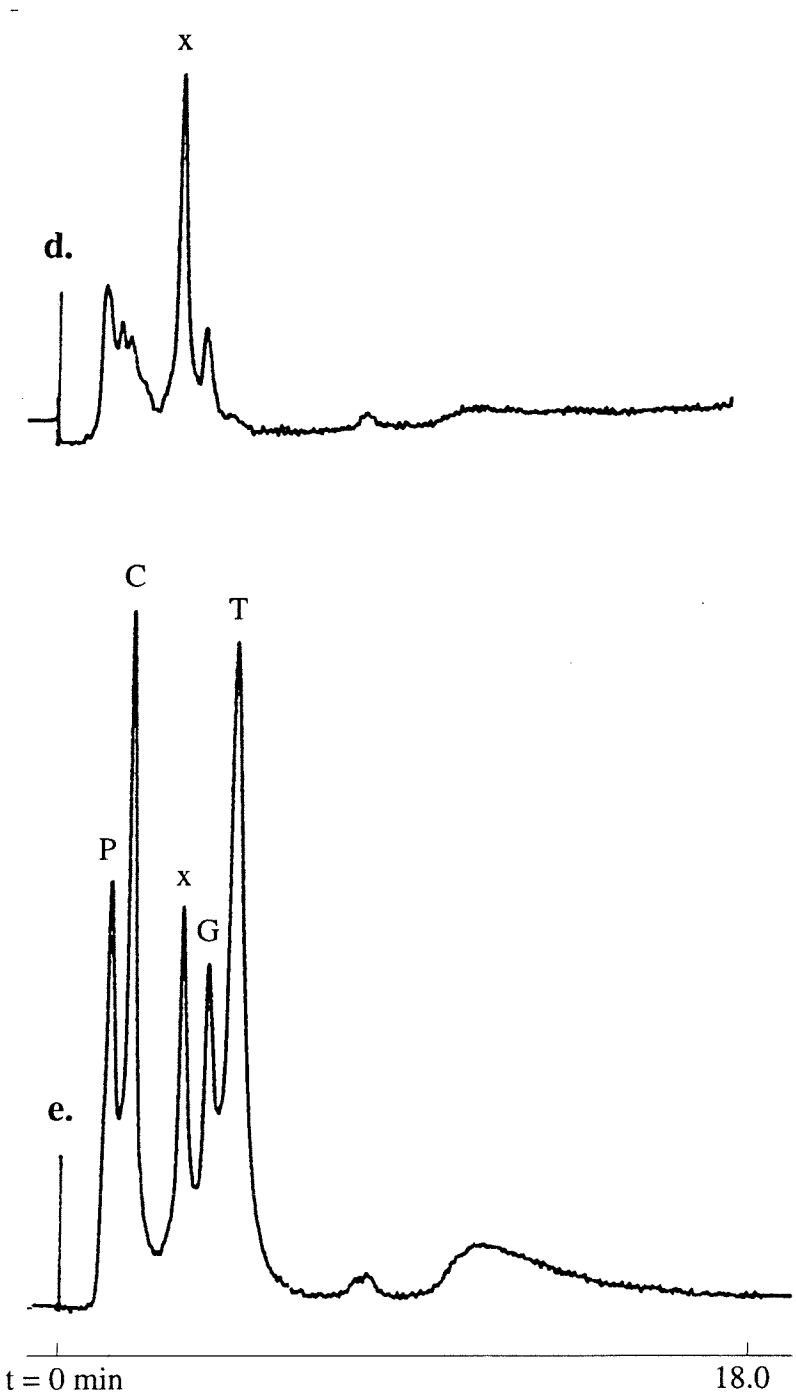

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
3'
GCTGACCTTTTC
5'
3'
TACAACACACCTTAAC
5'

Figure 3.5. HPLC traces showing the production of base propenoic acids. The products were detected at 260 nm (AUFS 0.005, data shown) or 300 nm (AUFS 0.003, data not shown). **(a)** Buffer blank. 0.05 M ammonium formate (40 μL); **(b)** Sk-PG + hν (40 μL); **(c)** DNA + hν (40 μL); **(d)** DNA + Sk-PG (40 μL); **(e)** DNA + Sk-PG + hν (40 μL). (P) denotes the base propenoic acids while (C), (G) and (T) denote the free bases cytosine, guanine and thymine respectively. Adenine free base is not eluted under these conditions. The peak labeled (x) is observed in the DNA control also. Note the enhancement in the production of base propenoic acids and free bases only when the DNA is irradiated in the presence of Sk-PG.

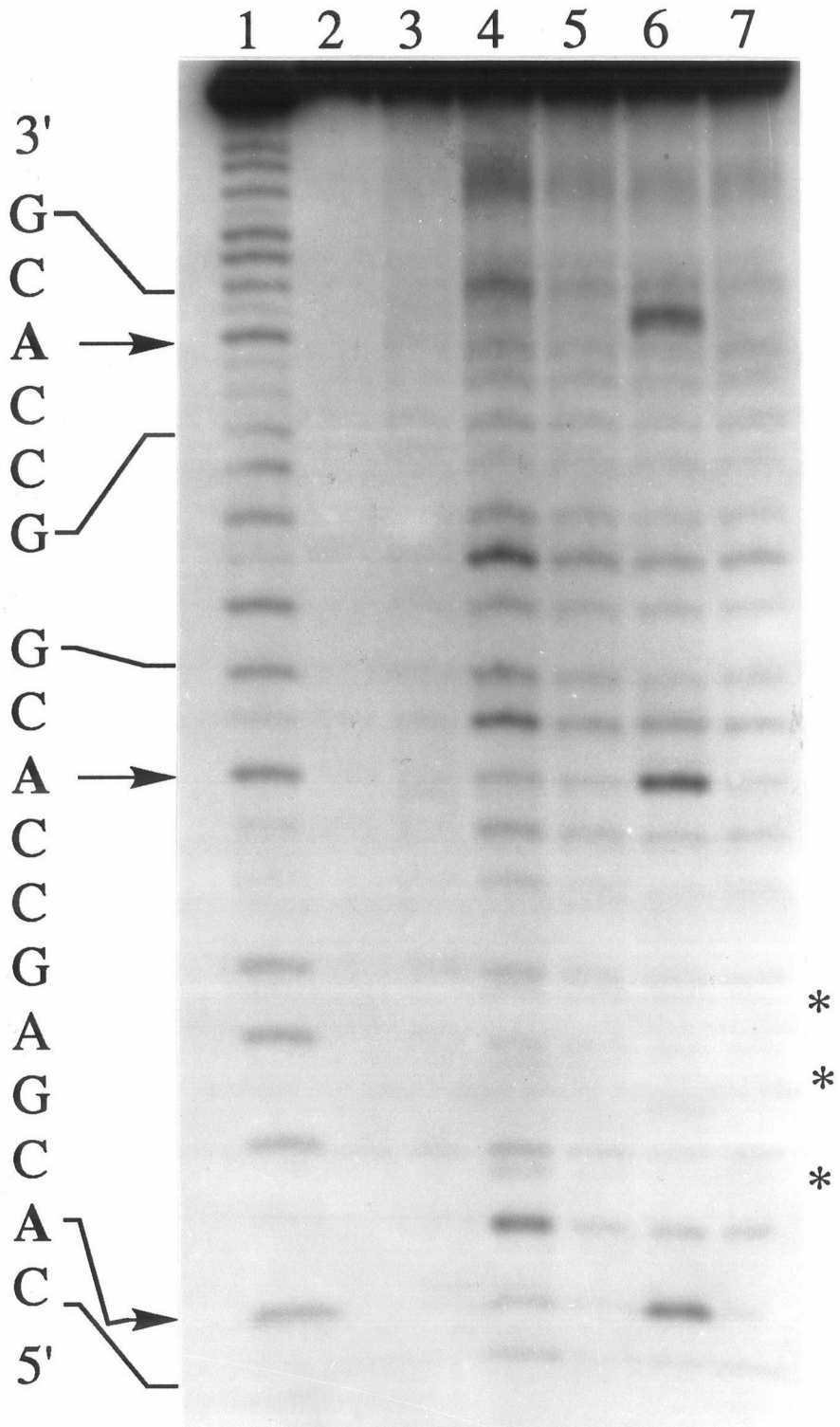

a.
b.
x
c.
t = 0 min
18.0


x
d.
C
T
P
x
G
e.
t = 0 min
18.0

propenoic acids (evidence for major groove chemistry)[5] was substantiated by monitoring the HPLC elution trace at 260 nm and 300 nm. Increasing the detection wavelength from 260 nm to 300 nm increases the ratio of the height of the propenoic acids peak to the nucleic acid base peaks.

Photocleavage of restriction fragments by the metal-peptide complexes is observed at the adenosine of 5'-CCAN-3'. This cleavage is consistent with intercalation in the 5'-AN-3' base step from the major groove by analogy to cleavage observed with other related phi complexes of rhodium.[5] However, the 5'-end labeled DNA cleavage experiments (Figure 3.2, 3.3 and 3.4) also reveal a second cleavage band migrating approximately with the base to the 3'- side of the cleavage site. The relative intensity of this secondary band was found to vary across the series of metal-peptide complexes. Since the resolution of these gels does not allow one to discriminate whether the band results from cleavage at an adjacent site or may correspond to a slower migrating product with a different 3'-terminus, a 31-base pair oligonucleotide duplex was synthesized containing three 5'-CCA-3' sites to distinguish between these possibilities. Photocleavage experiments were then carried out with Sk, Sk-PD and Sk-PH and compared to cleavage by $[Rh(phi)_2(bpy)]^{3+}$ (Figure 3.6). Again, only Sk-PD showed strong cleavage at the 5'-CCA-3' sites. The four complexes all showed cleavage bands corresponding to 3'-phosphate and 3'-phosphoglycaldehyde termini on a 20% denaturing polyacrylamide gel following photocleavage of 5'-end labeled oligonucleotides.[5] The formation of the 3'-phosphoglycaldehyde terminus results from an oxygen dependent strand scission pathway and is sensitive to the accessibility of dioxygen to the cleavage site. Consistent with this proposal is the observation that $[Rh(phi)_2(bpy)]^{3+}$ yields a greater amount of the 3'-phosphoglycaldehyde terminus (relative to the 3'-phosphate) compared to $[Rh(phi)_2(phen')]^{3+}$. Sk-PD shows the least amount of the 3'-phosphoglycaldehyde product suggesting that the incipient C3' radical produced is better shielded from dioxygen by the attached peptide. In the case of Sk-PD, moderately strong cleavage was

Figure 3.6. Photocleavage of a [5'-^{32}P]- end-labeled 31 bp oligonucleotide duplex with $[Rh(phi)_2(bpy)]^{3+}$, Sk, Sk-PD and Sk-PH to compare 3'-terminii produced. Maxam-Gilbert A+G sequencing reaction (lane 1). Untreated oligonucleotide duplex (lane 2). Oligonucleotide duplex irradiated in the absence of rhodium complex (lane 3). Duplex irradiated in the presence of 1.0 μM $[Rh(phi)_2(bpy)]^{3+}$ (lane 4), Sk (lane 5), Sk-PD (lane 6) and Sk-PH (lane 7). The major cleavage products are 3'-phosphate termini which co-migrate with the Maxam-Gilbert sequencing products while the secondary products corresponding to 3'-phosphoglycaldehyde terminii are indicated by the asterisks. The arrows represent cleavage at the three 5'-CCA-3' sites present on the duplex.


1 2 3 4 5 6 7
3'
G
C
A
C
C
G
G
C
A
C
C
G
A
G
C
A
C
5'
*
*
*

also observed at the base 3' to the 5'-CCA-3' cleavage site that appeared to co-migrate with the corresponding 3'-phosphate of the oligonucleotide containing an additional nucleotide. This observation may be the result of intercalation of the metal complex in two neighboring base steps, both of which allow for proper orientation of the tethered peptide along its target site.

3.3.5. Circular Dichroism of Metal-Peptide Complexes. Circular dichroic spectra of Sk-PD (Figure 3.7a), which specifically targets 5'-CCA-3', and Sk-PH (Figure 3.7b), which does not, were recorded in the presence and absence of 5 mM Mn^{2+} and heat to explore correlations between site-recognition and peptide folding. Both the metal-peptide complexes show slight increases in helicity (value of θ_{222}) upon addition of 5 mM Mn^{2+} and heating to 55° C. However, the striking difference between the two complexes lies in their inherent helicity.* Sk-PD is 72% helical in buffer at room temperature (89% in the presence of Mn^{2+} and heat) while Sk-PH is calculated to be only 8% helical (increasing to 18%) under identical conditions. This conformational difference between the two complexes may account for the observed sensitivity of the glutamate in DNA recognition.

CD spectra were also recorded of the free peptides PD (H_2N-AANVAIAAWERAA-$CONH_2$) and PH (H_2N-AANVAIAAWDRAA-$CONH_2$) in 10 mM Tris-HCl, pH 7.0 at ambient temperature with and without 5 mM Mn^{2+} (Figure 3.7c). PH appears to adopt a random conformation under these conditions, while PD may show some helical content. It is noteworthy that coupling the skeleton onto the

---

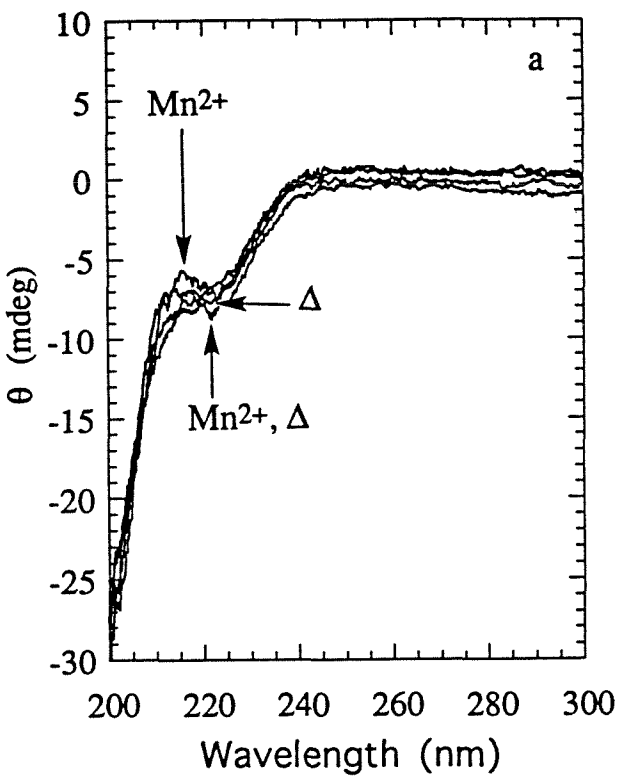
* The percent helicity is calculated from the mean residue ellipticity ($[\theta]_{222}$) by assuming that a 100% helical peptide has $[\theta]_{222}$ = -31500 deg.cm^2.dmol^{-1}.[17] The mean residue ellipticity is given by the relationship:

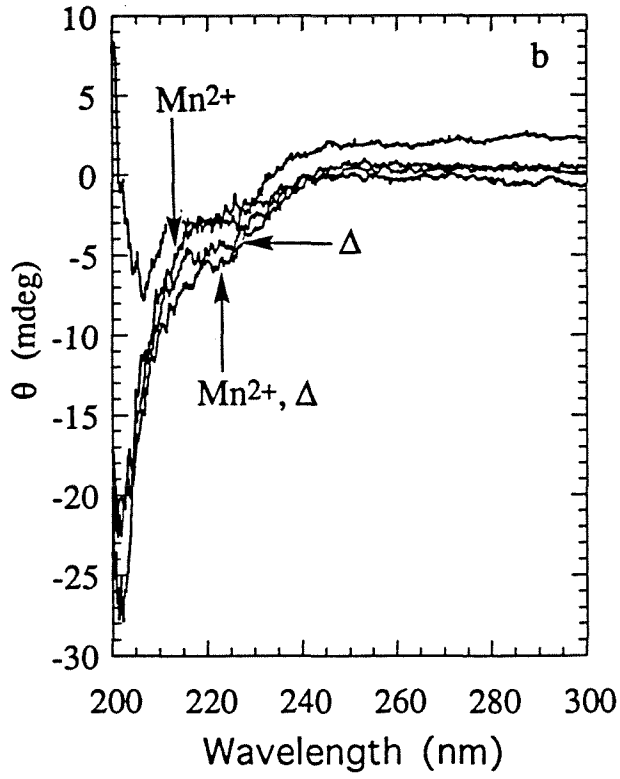
$$[\theta]_{222} = 100\,\theta_{222} / cnl$$

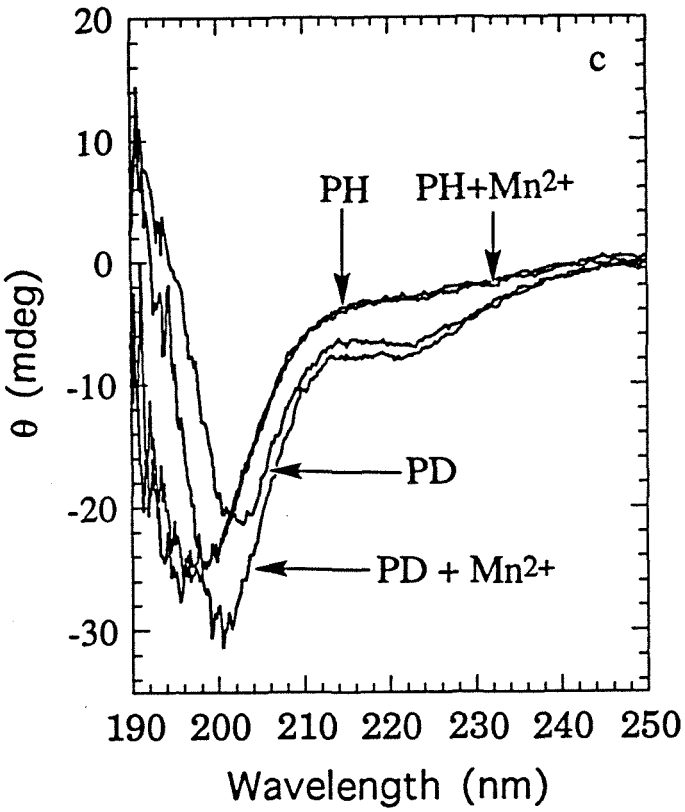
where

θ_{222} is the ellipticity (mdeg) measured from the CD spectrum, c is the peptide concentration (mM), n is the number of amino acids in the peptide, and l is the path length (cm) of the CD cell.[18]

Figure 3.7. Circular dichroism of Sk-PD, Sk-PH, PD and PH in 10 mM Tris-HCl, pH 7.0. The spectra of the metal-peptide complexes were recorded over 350-200 nm range in buffer at r.t., with the addition of 5 mM $MnCl_2$ at r.t. (Mn^{2+}), in buffer at 55° C (Δ) or with the addition of 5 mM $MnCl_2$ at 55° C (Mn^{2+}, Δ). (**a**) Sk-PD (23.0 μM). (**b**) Sk-PH (79.5 μM). (**c**) CD spectra of PD (182.0 μM) in buffer at r.t. (PD) or with the addition of 5 mM $MnCl_2$ (PD+Mn^{2+}); PH (112.3 μM) in buffer at r.t. (PH) or with the addition of 5 mM $MnCl_2$ (PH+Mn^{2+}). The spectra of the free peptides were recorded over the 250-190 nm range. A 0.1 cm path length cell was used in these experiments.


a
Mn2+
Δ
Mn2+, Δ
θ (mdeg)
10
5
0
-5
-10
-15
-20
-25
-30
200
220
240
260
280
300
Wavelength (nm)


b
Mn2+
Δ
Mn2+, Δ
θ (mdeg)
10
5
0
-5
-10
-15
-20
-25
-30
200
220
240
260
280
300
Wavelength (nm)


20
10
0
-10
-20
-30
θ (mdeg)
c
PH
PH+Mn2+
PD
PD + Mn2+
190 200 210 220 230 240 250
Wavelength (nm)

peptide enhances the helicity of the peptide (perhaps by virtue of capping the N-terminus, despite the positive charge of the skeleton).

A related series of spectra taken with Sk-PG and Sk-PE (Sk-PG-E10Q) at 5 μM concentration indicated only modest helicity that increased slightly upon addition of Mn^{2+} and heat (Figure 3.8). No differences were detected between the two metal-peptide complexes. This result is consistent with Sk-PG showing weaker cleavage at the 5'-CCA-3' site compared to Sk-PD, even though both contain E^{10}, while Sk-PE does not cleave the site at all.

It should also be noted that these CD spectra were taken in the absence of DNA. It is possible that minor structural differences in the native CD of the different metal-peptide complexes become accentuated in the presence of cognate DNA sites, resulting in differences in the targeting of the 5'-CCA-3' site. High absorptivity of the skeletal complex in the region of interest (250-190 nm) made the experiment in the presence of DNA difficult to accomplish.

3.3.6. Differences Between Alanine and Glutamate at Position 10. The experiments with Sk-PH (E10D mutant) clearly indicate that the 5'-CCA-3' recognition is observed only if E^{10} is present. However, the CD studies indicate that Sk-PD is helical while Sk-PH is not. Thus, there also appears to be a glutamate dependent conformational switch. To determine the role of helicity in recognition, Sk-PJ (Sk-PD-E10A) was synthesized next. The glutamate to alanine mutation was predicted to maintain helicity (or possibly even enhance it) while losing the ability to make base specific contacts with DNA through hydrogen bonding.

Figure 3.9 shows the CD spectra of Sk-PD, Sk-PH and Sk-PJ. Sk-PJ is highly helical, more so than even Sk-PD. Thus our control for DNA-recognition by a peptide in a helical conformation is valid. Figure 3.10 shows a direct comparison of Sk-PD, Sk-PH and Sk-PJ in their ability to recognize the 5'-CCA-3' sequence. Again only Sk-PD shows

Figure 3.8. Circular dichroism of Sk-PE and Sk-PG. 5 μM samples of the two metal-peptide complexes were made in 10 mM sodium phosphate, pH 7.0. Each spectrum was recorded using a 1 cm pathlength cell over the range 350-200 nm. The spectra were taken for each metal-peptide complex in buffer at room temperature or in the presence of 5 mM $MnCl_2$ at 55° C. Note a small change in the spectrum upon addition of Mn^{2+}. But this same change is evident for Sk-PG, which does target the 5'-CCA-3' sequence under these conditions, and Sk-PE, which does not.

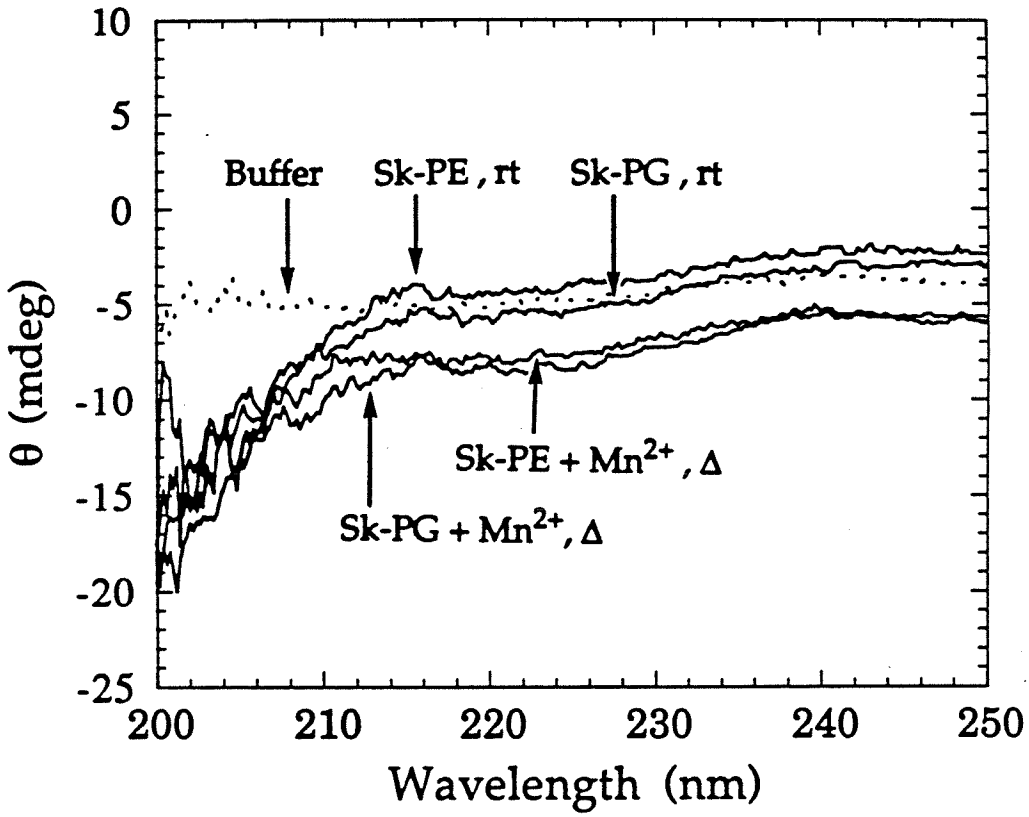

10
5
0
-5
-10
-15
-20
-25
θ (mdeg)
200
210
220
230
240
250
Wavelength (nm)
Buffer
Sk-PE , rt
Sk-PG , rt
Sk-PE + Mn2+ , Δ
Sk-PG + Mn2+, Δ

Figure 3.9. Comparison of the circular dichroism of Sk-PD, Sk-PH, Sk-PJ and Sk-PK in 10 mM Tris-HCl, pH 7.0. The spectra of the metal-peptide complexes (24.8 μM) were recorded in a 0.1 cm path length cell (200 μL total volume) over 350-200 nm range in buffer at r.t. and were the average of four scans. The calculated helicity of the metal-peptide complexes is: Sk-PH: 11.2%; Sk-PK: 26.2%; Sk-PD: 52.8% and Sk-PJ: 99.3%. Note the absence of a CD signal above 250 nm indicative of the metal-peptide complex being racemic at the metal-center and the absence of an induced CD. (See also Chapter 2, section 2.3.5.)

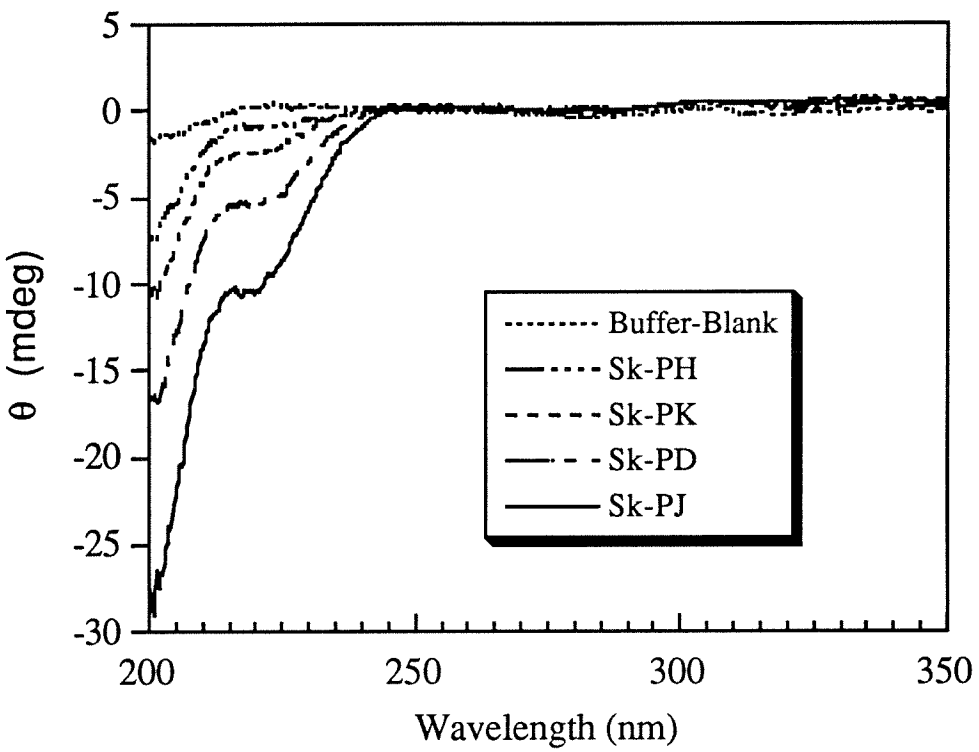
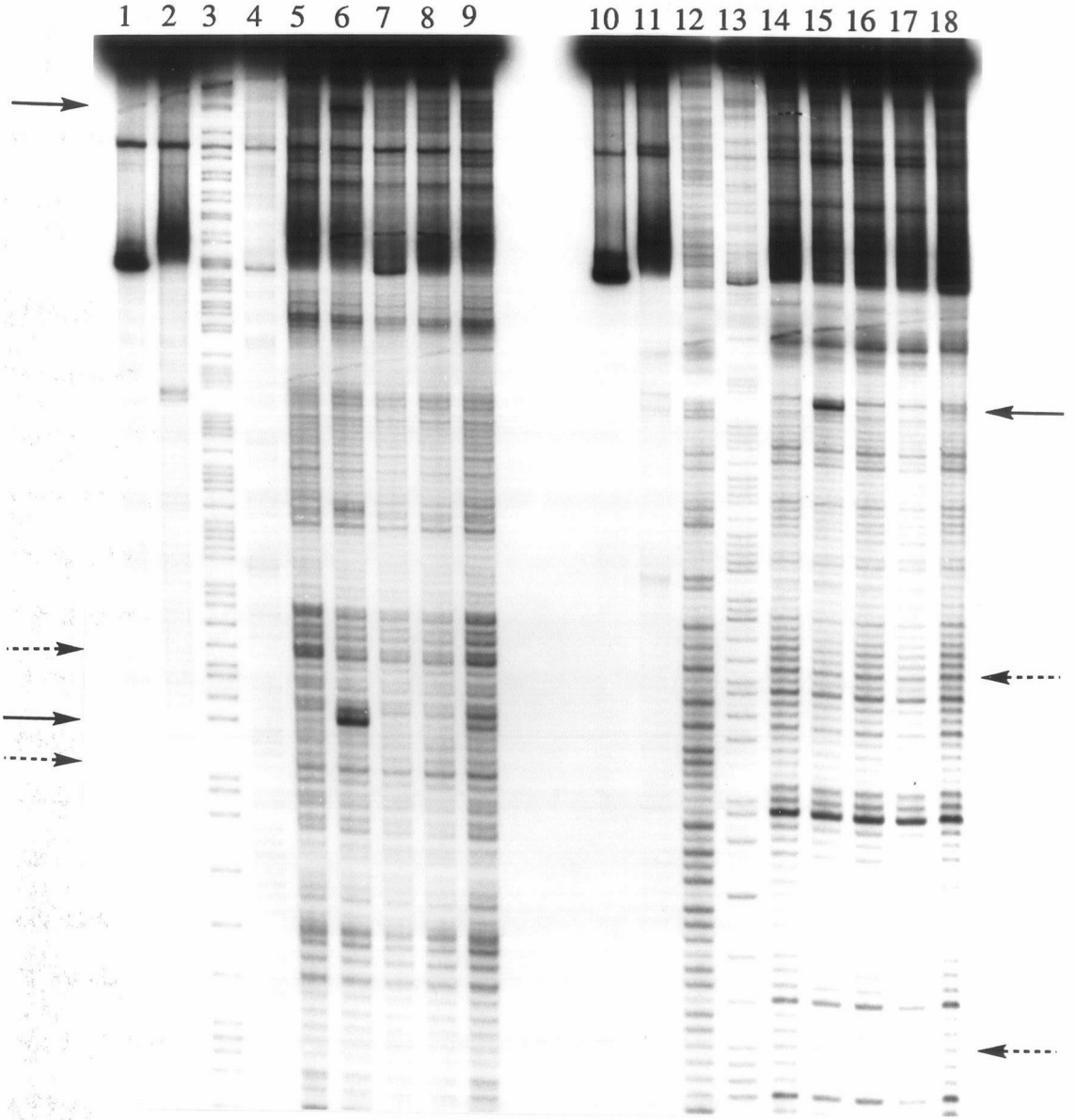

5
0
-5
-10
-15
-20
-25
-30
θ (mdeg)
200
250
300
350
Wavelength (nm)
Buffer-Blank
Sk-PH
Sk-PK
Sk-PD
Sk-PJ

Figure 3.10. DNA photocleavage by Sk-PD, Sk-PH, Sk-PJ and Sk-PK. Autoradiograms showing photocleavage reactions on the 180 base-pair fragment [5'-^{32}P]- (lanes 1-9) or [3'-^{32}P]- (lanes 10-18) end-labeled at the *EcoRI* site. The reaction samples were incubated in the presence of 5 mM $MnCl_2$ for 5 min at 55° C and irradiated (313 nm) for 10 min at 55° C. Lanes 1 and 10: DNA control. Lanes 2 and 11: fragment irradiated in the absence of rhodium complex. Lanes 3 and 12: Maxam-Gilbert A+G sequencing reaction. Lanes 4 and 13: Maxam-Gilbert C+T sequencing reaction. Lanes 5-9 and 14-18 fragment irradiated in the presence of 1 μM of Sk, Sk-PD, Sk-PH, Sk-PJ and Sk-PK, respectively. 5'-CCAN-3' sites have been indicated with solid arrows, while 5'-CAAN-3' or 5'-NTTG-3' sites have been indicated by dashed arrows. The cleavage sites are the bases in bold. Note: In spite of being highly helical, Sk-PJ does not show any 5'-CCA-3' recognition (compare lanes 8 and 17 with 6 and 15 respectively of Sk-PD). Also, a comparison of the photocleavage by Sk-PD and Sk-PK (lanes 6 and 9) indicates a modest enhancement in the cleavage by Sk-PK at 5'-CAAN-3' sites compared to that at 5'-CCAN-3' sites. Significant cleavage is also seen at 5'-CCA-3' sites by Sk-PK (not seen in the lanes containing Sk, Sk-PH or Sk-PJ) indicating the strong inherent influence of E^{10} in recognition.


1 2 3 4 5 6 7 8 9
10 11 12 13 14 15 16 17 18
5'-CAAN-3' or 5'-NTTG-3'
5'-CCAN-3'

the observed 5'-CCA-3' recognition. Sk-PH is neither helical nor does it show the site selectivity. In contrast, Sk-PJ is highly helical but does not show the 5'-CCA-3' recognition either. Sk-PJ, like Sk-PH, shows a background DNA site-selectivity identical to the skeletal complex.

3.3.7. Effect of Introducing Glutamines into the Peptide Sequence for DNA Recognition. The studies described above have been directed towards understanding the origins of the DNA recognition first observed in Sk-PA. Alanine substitution was carried out on the various potential DNA-contacting residues, and the specificity was localized to a single glutamate at position 10. Mutation of none of the other residues to alanine appeared to affect recognition. Also, the necessity for helicity was established by a correlation between peptide helicity and intensity of cleavage at the 5'-CCA-3' site. We envisioned that changing residues lying on the same face of a putative helix as E^{10} to residues containing other side-chain functional groups might alter the DNA specificity of the metal-peptide complexes. N^3 and S^7 of Sk-PA appeared likely candidates for this functional group substitution.

Since glutamines are frequently found making hydrogen bonding contacts with adenosine bases in a number of protein-DNA crystal structures,[15, 16] we chose to incorporate them in the peptide sequence for recognition. Sk-PK was synthesized with glutamines at positions 3 and 7 in the context of a Sk-PD framework (N3Q/A7Q). To compensate for the predicted reduction in helicity necessitated by these mutations (alanine has the highest helical propensity amongst all the residues),[19] two other mutations, V4A/I6A, were also made.**

As shown in Figure 3.9, the CD spectra indicate that Sk-PK has reduced helicity (26%) compared to Sk-PD (53%), but the helicity is higher than that for Sk-PH (11%).

** V^4 and I^6 have low α-helix propensities compared to alanine, and being hydrophobic lead to reduced aqueous solubility.

The DNA recognition properties of Sk-PK are shown in Figure 3.10. Compared to Sk-PD, the DNA recognition by Sk-PK has been perturbed towards 5'-CAA-3' sites on a 180 bp restriction fragment. Weak cleavage is also observed at 5'-CCA-3' sites consequent to the presence of E^{10} in the peptide sequence.

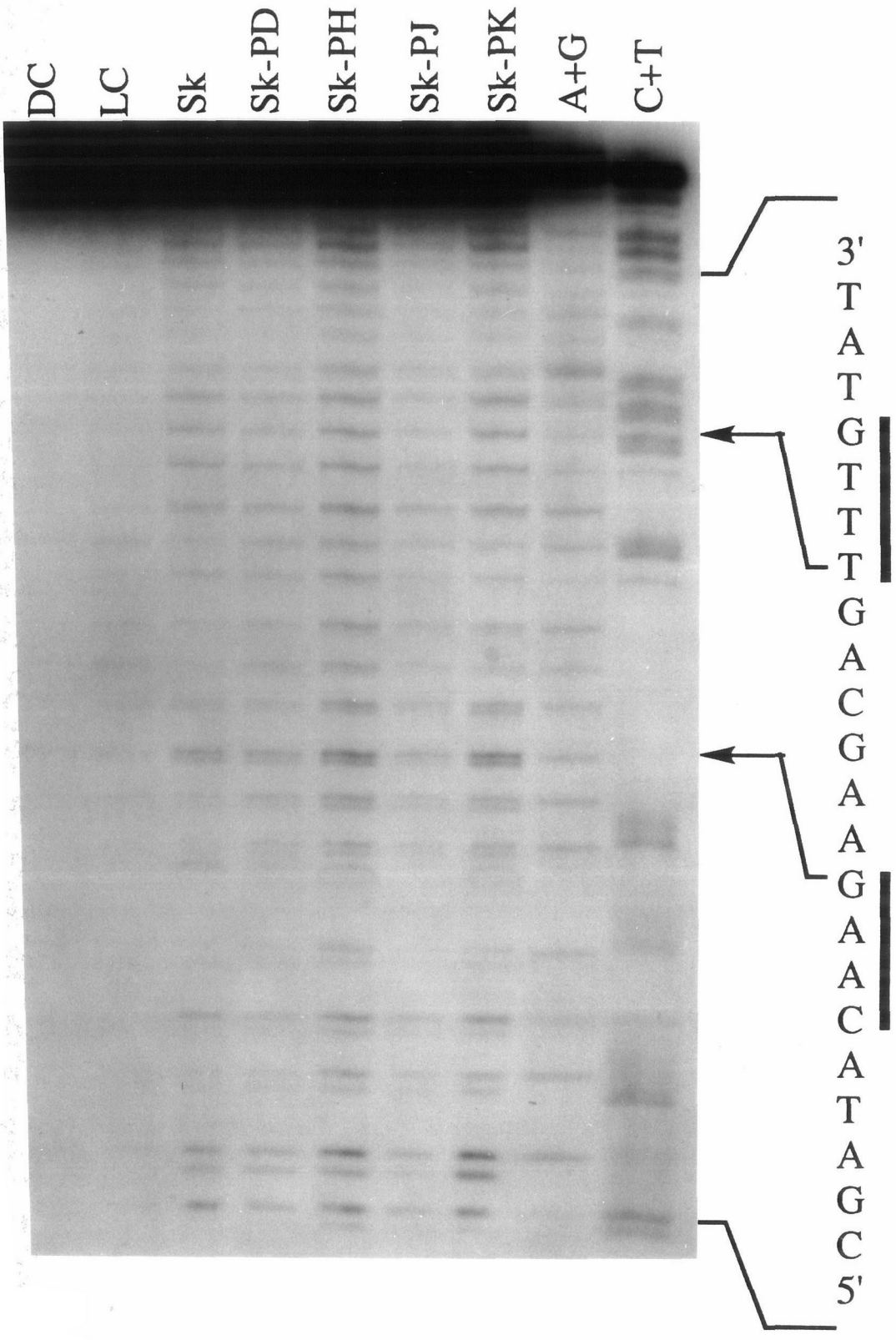
Photocleavage of two 30-mer oligonucleotide duplexes that contained two 5'-CAA-3' sites but no 5'-CCA-3' site was carried out next (Figure 3.11). The rationale being that if the glutamate recognition of 5'-CCA-3' was dominating all else, then by eliminating that site, the metal-peptide complexes may target other secondary sites. None of the complexes, Sk, Sk-PD, Sk-PH, Sk-PJ and Sk-PK, cleaved any site strongly. However, modest differences in cleavage intensity around the 5'-CAA-3' sites were observed between Sk-PD and Sk-PK. This discrimination is to be contrasted with Sk-PH which shows uniformly higher cleavage and Sk-PJ which shows uniformly lower cleavage at all sites while retaining a skeletal complex-like background DNA selectivity. Interestingly, in the absence of 5'-CCA-3' sites, secondary cleavage sites were not selected by Sk-PD.

3.4. Discussion

3.4.1. DNA Binding by the Metal-Peptide Complexes. Based upon the comparison to other phi complexes of rhodium,[4-6] the metal-peptide complexes appear also to bind DNA by intercalation of one of the phi ligands in the major groove. Like $[Rh(phi)_2(bpy)]^{3+}$, HPLC analysis of DNA cleavage products yields base propenoic acids and free bases with a similar total photoefficiency.[5] The similar pattern of photocleavage observed for the metal-peptide complexes compared to the skeletal complex at sites other than the 5'-CCA-3' sequences also lends support to this notion.

Binding affinity for a site is a function both of the metal complex and of the peptide moieties. The non-specific binding affinity of these metal-peptide complexes is estimated to be $\geq 10^6\ M^{-1}$ and largely derived from intercalation of the metal complex,

Figure 3.11. Photocleavage of a [5'-^{32}P]-end-labeled 30 bp oligonucleotide duplex with Sk, Sk-PD, Sk-PH, Sk-PJ and Sk-PK. DC: untreated oligonucleotide duplex. LC: oligonucleotide duplex irradiated in the absence of rhodium complex. A+G: Maxam-Gilbert A+G sequencing reaction. C+T: Maxam-Gilbert C+T sequencing reaction. Lanes containing photocleavage products have been labeled with the appropriate metal-peptide complex. 5'-CAAN-3' sites have been indicated by solid lines. Note the modest enhancement in cleavage at these sites (arrows) by Sk-PK compared to Sk-PD. Cleavage in the Sk-PH lane appears to be greater on the whole at all sites.


DC
LC
Sk
Sk-PD
Sk-PH
Sk-PJ
Sk-PK
A+G
C+T
3'
T
A
T
G
T
T
T
T
G
A
C
G
A
A
G
G
A
A
C
A
T
A
G
C
5'

while the contribution of the peptide alone to binding is estimated to be 10^3 M^{-1}.*** Indeed free peptide at 100 times the concentration does not compete with metal-peptide binding to its specific site. We may, therefore, consider the side-chain functionalities on the peptide as augmenting the DNA binding properties of the skeletal complex so as to achieve site-selectivity. This means of targeting a site resembles that of the larger DNA-binding proteins, where non-specific electrostatic and hydrogen bonding interactions contribute substantial binding affinity in the major groove, and site-selectivity is derived primarily from functional group interactions of a peptide recognition domain.

3.4.2. Site-Specificity of the Metal-Peptide Complexes. The specificity in DNA recognition by the family of peptide complexes is governed by several factors. First and foremost, the specificity is seen to be exquisitely sensitive to the presence of a glutamate at position 10. Conservative changes such as methylation of the carboxylate side-chain, or substitution with glutamine (isostere) or aspartate (lower homologue) are not tolerated. The high overall cleavage by Sk-PG indicates that increasing the net positive charge on the peptide may raise the overall binding affinity of the metal-peptide complex but does not appear to increase the binding specificity. In fact the relative cleavage at 5'-CCA-3' for Sk-PG is less than that for Sk-PD concomitant with a substantially lower helicity.

Peptide conformation certainly appears to affect recognition, as judged by the sensitivity of site recognition to the reaction conditions as well as sequence. Specificity is seen to be enhanced with the addition of reagents such as trifluoroethanol or in the presence of divalent cations (Mn^{2+}, Mg^{2+}) at elevated temperature (55° C). Also, CD data indicate a correlation between helicity and extent of the 5'-CCA-3' recognition. That *helicity is necessary but not sufficient for the observed 5'-CCA-3' selectivity* is

*** Cleavage by Sk-PD at the 5'-CCA-3' site is approximately 6.5 fold stronger than that by Sk at 1.0 μM metal concentration, which translates into a $\Delta\Delta G$ of $\geq$ 1 kcal/mole contribution of the peptide to specific binding over the skeletal complex.

exemplified by the comparison between Sk-PD and Sk-PJ. Sk-PJ is highly helical but does not recognize the 5'-CCA-3' sequence. Sk-PD is helical, contains the glutamate at position 10 and exhibits the 5'-CCA-3' selectivity.

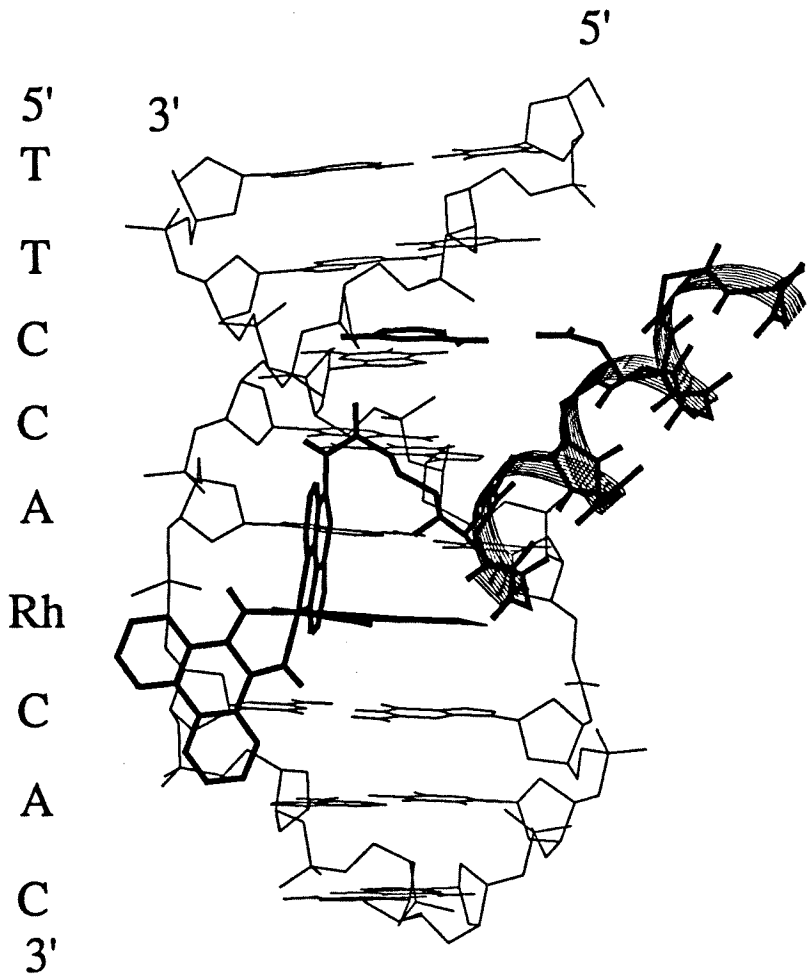
How can a single carboxylate be responsible for recognition of a three-base pair (5'-CCA-3') sequence? The absolute requirement of the glutamate for recognition gives rise to two possibilities : (i) the glutamate is required indirectly in folding the metal-peptide complexes into a unique conformation under the appropriate conditions; or (ii) the glutamate participates directly in interaction with the DNA base(s) in a structure shared by all the complexes when bound to DNA. The results presented here suggest that the glutamate may be serving both these roles. The differences in intensity of specific cleavage amongst the different metal-peptide complexes suggests that the glutamate plays a strong direct role in DNA recognition within the context of a structural role ascribed by the CD data. That glutamate is essential in both respects may be the source of the glutamate switch.

The construction of the metal-peptide family was based upon the α_3-helix of the phage P_{22} repressor protein. The solution structure of the protein has been recently solved,[10] but the exact nature of its interactions with DNA can only be inferred based upon analogy to the co-crystal structure of the closely related phage 434 repressor protein with its operator sequence.[20] The putative operator sequence for protein $P_{22}R$ is 5'-ANTNAAG-3'. The fact that these metal-peptide complexes do not preferentially target this sequence may arise because the protein serves to position the α-helix in the major groove of DNA through contacts with a different face of the peptide than in the metal-peptide complex, or that the shape of the protein complements the shape of the DNA around the operator sequence so as to orient the recognition helix differently. Furthermore, in the metal-peptide complexes, which obviously lack stabilizing interactions with the rest of the protein, the peptide conformation may be directed instead by the metal center and by solvent. The solution structure of the DNA binding domain of

the P_{22} repressor indicates that I^{35} and W^{38} (I^{6} and W^{9} respectively in our system) form the core of the hydrophobic interior of the protein and that E^{39} (E^{10} in our system) is solvent inaccessible (and presumably also inaccessible to DNA) as a result of hydrogen bonding to Nε of R^{11} / R^{14} in the interior of the protein.[10] Thus, intra-protein interactions involving these three side chains rigidly set the orientation of α_3 in the P_{22} repressor. In our system, the corresponding Ile and Trp side chains may reorient themselves to point into the major-groove and away from solvent while the glutamate side chain of E^{10} is no longer tied down and is available to make functional group interactions with the DNA.

3.4.3. Model for 5'-CCA-3' Recognition by the Metal-Peptide Complexes. As illustrated schematically in Figure 3.12, we can model the 5'-CCA-3' recognition by the metal-peptide complexes using a combination of base-specific contacts and shape-selection. We propose that the peptide folds across the intercalated metal complex (likely as a helix), necessitating an enlarged major groove; there is evidence that 5'-pyr-pyr-pur-3' sequences are somewhat opened in the major groove.[21] The folding of the peptide then presents the glutamate near the carboxy terminus in position for specific interaction with the 5'-cytosine (Figure 3.12). There is precedence for glutamate side chains being involved in the recognition of cytosines.[22] Photocleavage experiments also show that the metal-peptide complex is bound asymmetrically in the intercalation site and is canted substantially towards one strand. This strand asymmetry provides room in the major groove for the peptide to interact with DNA, as the model illustrates. The added cleavage observed at the 3'-base flanking the 5'-CCA-3' site can be modeled by positioning the linker in a more extended conformation. Further structural studies are, however, needed to validate this model.

Figure 3.12. Model for DNA recognition by metal-peptide complexes. The model shows the rhodium complex intercalated in the major groove of DNA at the 5'-AC-3' base step and canted towards the 5'-strand. Also shown is the glutaryl linker and the peptide backbone in a likely α-helical conformation. All the side chains on the peptide have been omitted for clarity except for the essential glutamic acid residue. The glutamate side-chain is shown in a position to make a hydrogen bond with the 4-amino group on the cytosine. The model was built by docking the metal-peptide complex to an 8-mer DNA duplex (5'-TTCCA-CAC-3') containing an intercalation site between A_5 and C_6 using Macromodel. Care was taken to avoid all steric clashes between the various non-bonded atoms in the model. The additional cleavage band resulting from intercalation between C_6 and A_7 in the figure can also be modelled to account for the same 5'-CCA-3' recognition if the linker were to adopt a slightly more extended form away from the metal-complex.


5'
3'
5'
T
T
C
C
A
Rh
C
A
C
3'

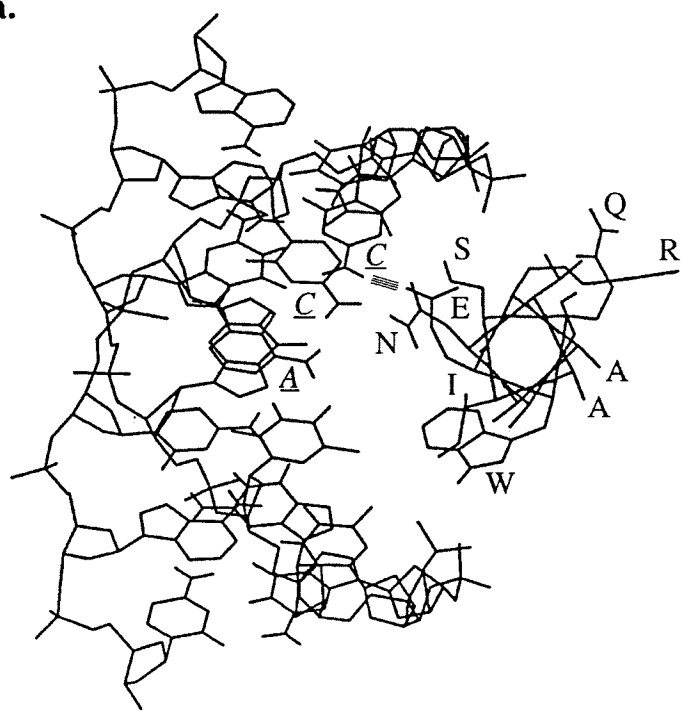
3.4.4. Predictions Based on the Model. Given that these metal-peptide complexes evidently bind to DNA while presenting a different face of the α-helix than in the native protein, can we make predictions for altering the DNA site-selectivities of designed mutants? Shown in Figure 3.13 (a) is a view of the peptide helix docked into the major groove of B-DNA contacting a 5'-CCA-3' site. The E^{10}:5'-C contact is also shown. In this model the side chains of N^3 and S^7 are also buried in the major groove, but are too short to contact any of the bases in the floor of the groove. Instead, by making a N3Q/S7Q double mutation, we can model the tethered peptide complex to recognize a 5'-CAA-3' sequence (Figure 3.13 (b) and 3.14). The glutamine side chains are now in hydrogen bonding distance from the two adenosine bases and have the potential to each make a pair of donor-acceptor hydrogen bonds.

This prediction is borne out by Sk-PK which has the modeled N3Q/S7Q double mutation. The complex shows a perturbed DNA recognition towards the predicted 5'-CAAN-3' sequence. The relatively modest change in specificity may be explained by two factors. First, Sk-PK is not very helical compared to Sk-PD (Figure 3.9). The model predicts recognition of 5'-CAAN-3' sites only if the peptide adopts a helical conformation. Second, weak but nonetheless competing 5'-CCA-3' recognition is seen, as a result of the presence of E^{10}.

Further predictions may be made by introducing residues such as Arg or His at specific positions and distances away from the glutamate for recognition of guanosine bases with the caveat that the peptides are helical. This mode of recognition using two or three residues in an α-helix to recognize three base pair sites on one side of the major groove is reminiscent of zinc-finger DNA binding proteins.[23] The glutamate may serve as a common second point of attachment for pinning the metal-peptide complexes to DNA, the intercalated skeletal complex being the first. Indeed the underlying 5'-CCA-3' recognition seen in Sk-PK is suggestive of this notion. An important consideration in further design is to ascertain the position dependence of the glutamate in DNA

Figure 3.13. Models showing the peptide docked into the major groove. The relative dispositions of the metal-peptide complex and DNA are as in Figure 3.12. The skeletal complex and the linker have been deleted for clarity. The views are along the helical axis of the peptide. Putative hydrogen bonding contacts have been represented by hatched lines. The amino acid residues comprising the peptides have been indicated by their one letter codes while the nucleic acid bases have been indicated in italics and underlined. **(a)** Model of Sk-PD showing the single hydrogen bonding contact between E^{10} and 5'-C. N^3 and S^7 side chains are not long enough to contact any bases in the major groove in this model. **(b)** Model of Sk-PK with the N3Q/S7Q double mutation. The longer glutaminyl side chains have the potential to make hydrogen bonding contacts with adenosines of a predicted 5'-CAA-3' site.

a.



b.

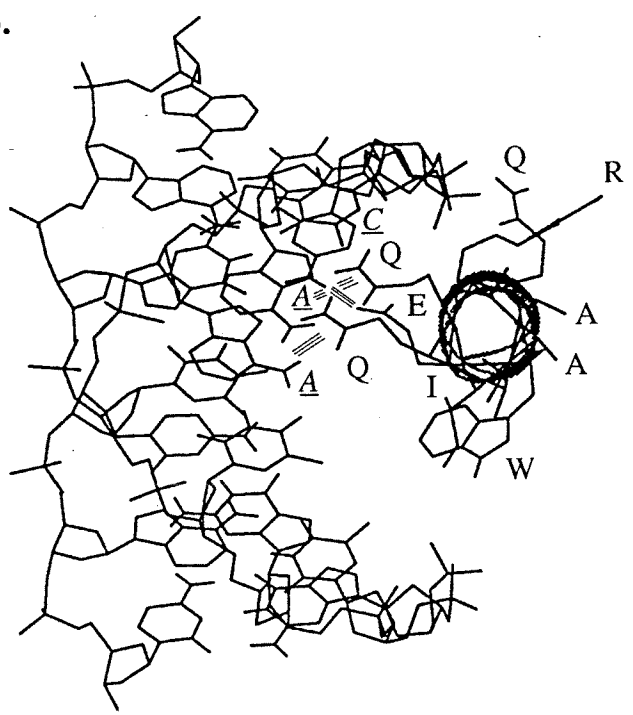
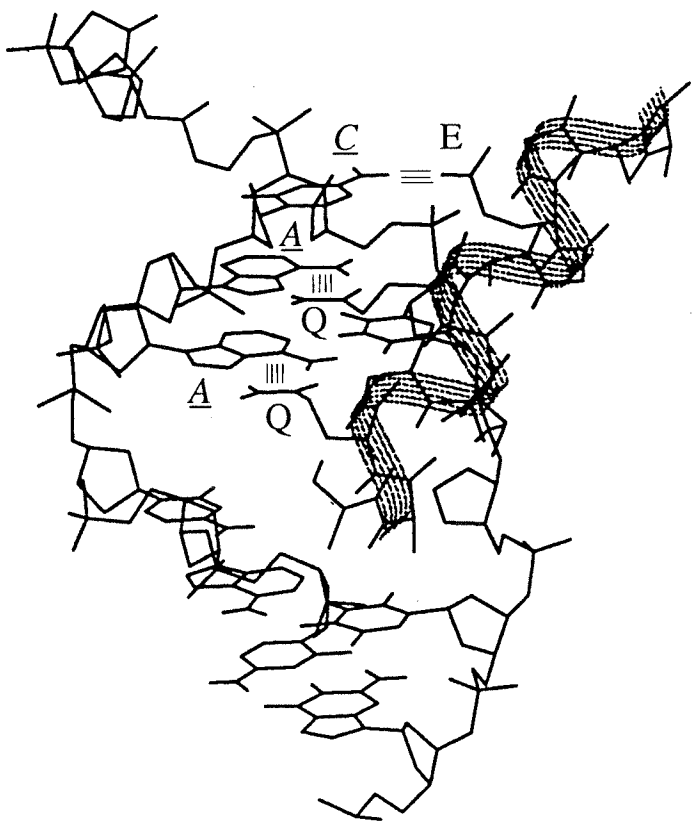


Figure 3.14. Model of Sk-PK docked into the major groove of B-DNA showing the three potential hydrogen bonding contacts between E^{10}, Q^{7} and Q^{3} and a 5'-CAA-3' site. The N-terminus of the peptide is at the bottom as in Figure 3.12 with the skeletal complex intercalated between the 5'-A.N-3' base step. The skeletal complex, linker and all other side chains have been omitted for clarity.


C
E
A
Q
A
Q

recognition. Perhaps by simply altering the location of the single glutamate residue along the helix, new DNA recognition scaffolds may emerge and may be built on by introducing other functional side chains.

3.4.5. Glutamate Switch in DNA Recognition. The recognition characteristics of this family of metal-peptide complexes provide an example of a glutamate switch in site-specific DNA recognition; if glutamate is present, 5'-CCA-3' recognition is achieved. The sensitivity of recognition to the presence of a unique residue has been documented in other natural proteins. A single glutamic acid residue was shown to be significant for transcriptional activation by the lambda repressor, a mutant containing Asp at this position activated transcription only two fold above the basal level.[24a] Similarly, a double substitution of VE/DQ in the HIV type I integrase enzyme rendered the integrase protein inactive for all its functions.[24b] Mutant enzymes D190E and D192E, of the rat DNA polymerase β, in which aspartic acid residues at positions 190 and 192, respectively, were replaced by glutamic acid residues, showed only about 0.1% activity of the wild type enzymes.[24c] Hence *this small metal-peptide complex may serve as a particularly useful mimic for site-specific DNA-binding proteins.*

In general, these metal-peptide complexes appear to share an array of characteristics with DNA-binding proteins. Both may be viewed as consisting of a DNA recognition element that augments high affinity DNA binding by the rest of the molecule through sequence-specific contacts. Both show sequence-specificity that is highly sensitive to functional group placement. Indeed the specificity and affinity of these metal-peptide complexes exceeds that of most single zinc-finger domains of transcription factors.[23] What distinguishes these metal-peptide complexes is their small molecular size and, perhaps, the simplicity of the recognition switch.

Therefore, these metal-peptide complexes may be viewed as small molecular models for larger DNA-binding proteins. The metallointercalator may be used to anchor

functionality in the DNA major groove and the peptide functionality may be used sensitively to control site recognition. The construction of these molecules represents a new strategy to create an array of monomeric metal-peptide complexes with differing specificities for double helical DNA.

3.5. References

1. (a) Pyle, A. M.; Barton, J. K. *Prog. Inorg. Chem.* **1990**, *38*, 413. (b) Dervan, P. B. *Science* **1986**, *232*, 464. (c) Hecht, S. M. *Acc. Chem. Res.* **1986**, *19*, 383. (d) Zein, N.; Sinha, A. M.; McGahren, W. J.; Ellestad, G. A.; *Science* **1988**, *240*, 1198. (e) Nicholaou, K. C.; Dai, W. M.; Tsay, S. C.; Estevez, V. A.; Wrasidlo, W. *Science* **1992**, *256*, 1172.
2. (a) Cuenod, B.; Schepartz, A. *Science* **1993**, *259*, 510. (b) Talanian, R. V.; McKnight, C. J.; Kim, P. S. *Science* **1990**, *249*, 769. (c) Talanian, R. V.; McKnight, C. J.; Rutkowski, R.; Kim, P. S. *Biochem.* **1992**, *31*, 6871. (d) Morii, T.; Simomura, M.; Morimoto, S.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 1150. (e) Dervan, P. B. *Methods Enz.* **1991**, *208*, 497.
3. (a) Chow, C. S.; Barton, J. K. *Methods Enz.* **1992**, *212*, 219. (b) Sitlani, A.; Barton, J. K. *Handbook of Metal-Ligand Interactions of Biological Fluids*, Marcel-Dekker, Inc., New York, **1994** in press.
4. (a) Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520. (b) David, S. S.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 2984. (c) Krotz, A. H.; Kuo, L. Y.; Shields, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 3877.
5. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K. *J. Am. Chem. Soc.* **1992**, *114*, 2303.
6. (a) Sitlani, A.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12589. (b) Krotz, A. H.; Hudson, B. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12577.

7. (a) O'Neil, K. T.; Hoess, R. H.; DeGrado, W. F. *Science* **1990**, *249*, 774. (b) Ueno, M.; Murakami, A.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1993**, *115*, 12575.

8. (a) Ebright, R. H.; Ebright, Y. W.; Pendergrast, P. S.; Gunasekera, A. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 2882. (b) Sigman, D. S.; Chen, C. B. *Annu. Rev. Biochem.* **1990**, *59*, 207.

9. (a) Sluka, J. P.; Horvath, S. J.; Bruist, M. F.; Simon, M. I.; Dervan, P. B. *Science* **1987**, *238*, 1129. (b) Mack, D. P.; Shin, J. A.; Griffin, J. H.; Simon, M. I.; Dervan, P. B. *Biochemistry* **1990**, *29*, 6561. (c) Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847. (d) Graham, K.; Dervan, P. B. *J. Biol. Chem.* **1990**, *265*, 16534. (e) Sluka, J. P.; Griffin, J. H.; Mack, D. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 6369.

10. (a) Sevilla-Sierra, P.; Otting, G.; Wüthrich, K. *J. Mol. Biol.* **1994**, *235*, 1003. (b) Poteete, A. R.; Ptashne, M. *J. Mol. Biol.* **1982**, *157*, 21.

11. Caruthers, M. H.; Barone, A. D.; Beaucage, S. L.; Dodds, D. R.; Fisher, E. F.; McBride, L. J.; Matteucci, M.; Stabinsky, Z.; Tang, J-Y. *Methods Enz.* **1987**, *154*, 287.

12. Zimmermann, K. *Post Doctoral Research Report* (**1993**), Caltech.

13. Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning*; Cold Spring Harbor Laboratory, **1982**.

14. Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 560.

15. Pabo, C. O.; Sauer, R. T. *Ann. Rev. Biochem* **1992**, *61*, 1053.

16. Steitz, T. A. *Q. Rev. Biophys.* **1990**, *23*, 205.

17. Chen, Y. H.; Yang, J. T.; Martinez, H. M. *Biochemistry* **1972**, *11*, 4120.

18. Lehrman, S. R.; Tuls, J. L.; Lund, M. *Biochemistry* **1990**, *29*, 5590.

19. (a) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526. (b) Padmanabhan, S.; Marqusee, S.; Ridgeway, T.; Laue, T. M.; Baldwin, R. L. *Nature* **1990**, *344*,

268. (c) O' Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646. (d) Presta, L. G.; Rose, G. D. *Science* **1988**, *240*, 1632. (e) Padmanabhan, S.; York, E. J.; Gera, L.; Stewart, J. M.; Baldwin, R. L. *Biochemistry* **1994**, *33*, 8604. (f) Chakrabartty, A.; Kortemme, T.; Baldwin, R. L. *Prot. Science* **1994**, *3*, 843.

20. (a) Anderson, J. E.; Ptashne, M.; Harrison, S. C. *Nature*, **1987**, *326*, 846. (b) Wharton, R. P.; Ptashne, M. *Nature*, **1985**, *316*, 601.

21. (a) Larsen, T. A.; Kopka, M. L.; Dickerson, R. E. *Biochemistry* **1991**, *30*, 4443. (b) Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432.

22. (a) Schultz, S. C.; Shields, G. C.; Steitz, T. A. *Science* **1991**, *253*, 1001. (b) Raskin, C. A.; Diaz, G.; Joho, K.; McAllister, W. T. *J. Mol. Biol.* **1992**, *228*, 506.

23. (a) Berg, J. M. *Annu. Rev. Biophys. Biophys.Chem.* **1990**, *19*, 405. (b) Lee, M. S.; Gottesfeld, J. M.; Wright, P. E. *FEBS Lett.* **1991**, *279*, 289.

24. (a) Bushman, F. D.; Shang, C.; Ptashne, M. *Cell* **1989**, *58*, 1163. (b) LaFemina, R. L.; Schneider, C. L.; Robbins, H. L.; Callahan, P. L.; LeGrow, K.; Roth, E.; Schleif, W. A.; Emini. E. A. *J. Virol.* **1992**, *66*, 7414. (c) Date, T.; Yamamoto, S.; Tanihara, K.; Nishimoto, Y.; Matsukage, A. *Biochemistry* **1991**, *30*, 5286.

Chapter 4:

DNA Recognition by Metal-Peptide Complexes Bearing the Recognition Helix of the 434 Repressor Protein: Operator Recognition by an Artificial Repressor

4.1. Introduction

DNA binding proteins play an important regulatory role in controlling replication and transcription of genomic material. In a number of prokaryotic and eukaryotic transcription factors, the DNA recognition element is an α-helix in the context of different structural motifs such as helix-turn-helix, basic region-leucine zipper (bZIP), zinc-finger, homeodomain, helix-loop-helix and the like. Indeed, in a number of cases, the DNA binding domain is distinct from the transcripton activation domain and the two may be separated for studying DNA recognition by these proteins. Solution studies, footprinting, interference experiments and NMR and X-ray crystallographic analyses of co-crystals have led to a better understanding of DNA-protein interactions.[1-5]

With the exception of the recent discovery of a monomeric DNA binding domain in the Skn-1 protein in *Caenorhabditis elegans* embryos,[6] all the DNA binding proteins studied hitherto act as dimers or linked multimers. The requirement for dimerization is also borne out from results of studies on minimalist DNA binding protein domains. Early work on the Hin recombinase protein, a member of the helix-turn-helix motif DNA binding proteins, indicated that a 31-residue piece containing only the helix-turn-helix domain was not sufficient to impart sequence-specific DNA binding.[7] Instead, a larger 52-residue piece was necessary to specifically interact with the DNA site. More recently, the basic region (BR) α-helix of the bZIP protein GCN4 and the helix-loop-helix protein MyoD has been the focus of studies on dimerization of the two recognition helices via a linker.[8] The nature of the linker has varied in the literature and includes both covalent as

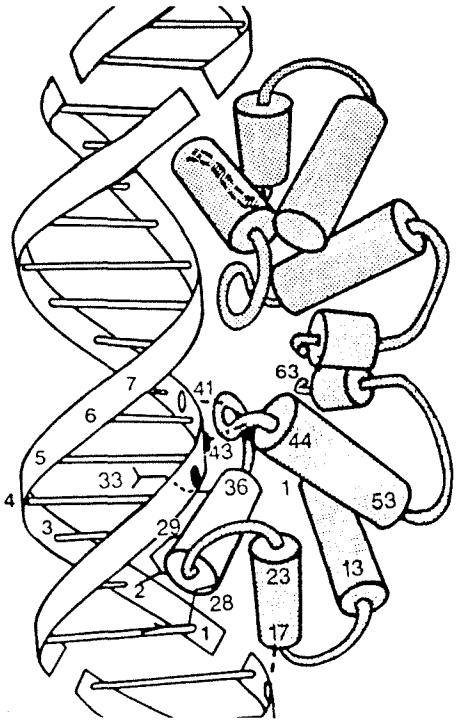
well as non-covalent modes of dimerization. As monomers, these α-helices or protein domains lack the ability to bind tightly and therefore lose their specificity.

Functionally, DNA binding proteins may be broken down into two elements; the DNA recognition domain and the protein scaffolding associated with the recognition domain. The protein scaffolding may be viewed as being required for: (a) delivering the recognition element (e.g., an α-helix) in a distinct orientation to the major groove; (b) providing a high degree of non-specific binding affinity through extensive phosphate contacts and shape complementarity to supplement the specific interactions between the recognition element and the operator site; and (c) making other protein-protein contacts for transcription regulation.

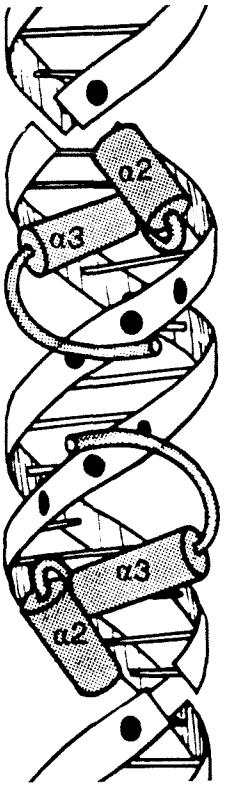
Described in Chapters 2 and 3 is our design of metallointercalator mimics of DNA binding proteins. In a manner similar to DNA binding proteins, the parent metal complex, $[Rh(phi)_2(phen')]^{3+}$, is designed to serve as a carrier to deliver peptides to the major groove of DNA and provide a high level of non-specific binding affinity ($K_a \geq 10^6$ M^{-1}). The appended peptides may serve as recognition elements and direct site-specific binding of the metal-peptide complex. Parallel work on the recognition helix of the phage 434 repressor protein (434R) is described in this chapter. 434R is a well studied protein, both biochemically[9] and structurally[10] and the specifics of the interactions with its DNA operator sites are known in atomic detail. 434R is a helix-turn-helix protein. The X-ray crystal structure of the DNA recognition domain (N-terminal residues, R1-69) complexed to a synthetic 20 bp operator site has been solved to 2.5 Å resolution.[10] R 1-69 has five α-helices, is a monomer in solution and binds to it's DNA operator site as a dimer (Figure 1a). The dimer interface consists of a patch of hydrophobic residues, a salt bridge from Arg^{41} to Glu^{47} and a hydrogen bond between two Arg residues. Helix 3 (α_3) is seen to lie in the major groove (Figure 1b). Mutational studies and helix swap experiments with the closely related phage 434 Cro protein and phage P22R protein show that α_3 is solely responsible for making base-specific contacts with the operator site.[11]

Figure 4.1. (a) Schematic showing the 434R 1-69 dimer bound to a DNA operator site. Helices are shown as cylinders; numbers indicate key residues as well as first and last residues of α_1-α_4. α_2 (residues 17-23) and α_3 (residues 28-36) form the helix-turn-helix motif. **(b)** Schematic of the helix-turn-helix motif bound to DNA and the phosphate contacts made by 434R1-69. α_3 lies in the DNA major groove and serves to make base specific contacts with the operator site. **(c)** Diagram of α_3 positioned in the DNA major groove showing the base specific contacts that lead to recognition of the operator site by 434R. Figure adapted from reference 10 d.

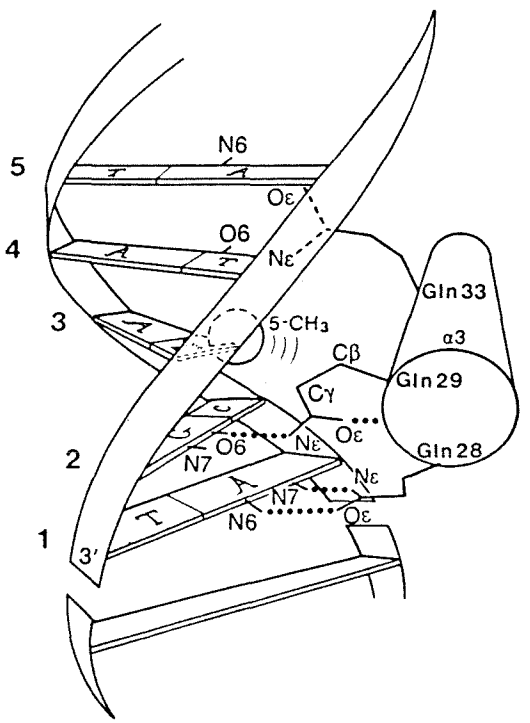
a.


7
6
5
4
3
2
1
41
63
43
44
33
36
1
53
29
23
13
28
17

b.


α2
α3
α3
α2

c.


5
4
3
2
1
3'
N6
Oε
O6
Nε
5-CH3
Cβ
Cγ
Gln33
α3
Gln29
Gln28
O6
N7
Nε
Oε
Nε
N7
N6
Oε

Some non-specific interactions are provided by salt bridges between Arg^{10} and Arg^{43} (not on α_3) with the phosphate backbone and through hydrophobic interactions with side chains of other residues. All the amino acid side chain:nucleic acid base contacts are summarized in Figure 1c. Gln^{28}, Gln^{29} and Gln^{33} make base specific contacts with the operator site through a combination of hydrogen bonds and van der Waals interactions. The consensus half operator site for 434R is 5'-ACAATAT-3'.[9, 10] However, only the 5'-ACAA-3' sequence is strictly conserved in 11 of the 12 naturally occurring operator half sites and is directly contacted by the 3 Gln residues.[9]

We have investigated the DNA binding properties of the recognition helix (α_3) of 434R with the aim of utilizing isolated peptides in DNA recognition by directing side-chain functionalities to the major groove. We wanted to address the question whether the helix-turn-helix domain or the recognition helix itself can target the operator DNA. Furthermore, are all the recognition elements to specify a DNA sequence contained in the peptide helix once it sits snugly in the DNA major groove? As with the peptides described in Chapter 3, a family of metal-peptide complexes has been constructed and characterized by their structural, photochemical, and DNA recognition properties. The metal-peptide complexes described here share the remarkable property of reproducing operator site recognition. Single amino acid mutations in the peptide sequence introduce small but distinct variations in the affinities for the DNA sequences selected. This ability to reproduce operator site binding in a metallo-intercalator provides a route to the synthesis of artificial repressors. The utility of the approach stems from the fact that these metal-peptide complexes bind DNA specifically as monomers.

4.2. Experimental

4.2.1. Materials. Sonicated calf thymus DNA was purchased from Pharmacia, plasmid pUC 18 was purchased from Boehringer-Mannheim and all enzymes utilized were from commercial sources. [α-^{32}P]dATP and [γ-^{32}P]ATP were obtained from NEN-Dupont

and [α-^{32}P]ddATP from Amersham, Inc. Oligonucleotides were synthesized by the phosphoramidite method,[12] using 1.0 μmol columns on an ABI 391 DNA-RNA synthesizer. d_3-2,2,2-trifluoroethanol (d_3-TFE, 99.94%) was purchased from Cambridge Isotope Labs. All other chemicals and reagents were from Aldrich. $[Rh(phi)_2(bpy')]^{3+}$-AANVAISQWERAA-$CONH_2$ [bpy'=4-(4-carboxybutyl), 4'-methyl 2, 2'-bipyridine] was graciously provided by Dr. Kaspar Zimmerman.

4.2.2. Design Considerations in Choosing Peptide Sequences for Study. The wild type sequence of α_3-434R (α_3, Table 4.1) was the first peptide chosen for study and serves as the basis for subsequent design. The peptide is only 11 residues in length and is not expected to be highly helical by itself in solution in the absence of the stabilizing interactions with the rest of the protein.

A new series of peptide sequences (α_A and α_C) was designed by considering the P_α and σ values of the different residues.[13] The T1L, S4A and G11A substitutions were prompted by the fact that Thr and Ser have low helix propagation values and Gly is a helix terminator. In contrast, Leu is a helix initiator while Ala has the highest helix propagation value amongst all residues.[13] The four Ala residues were introduced at the C-terminus to promote helicity and increase the length of the putative helix to 4 turns. The Gln residues were left unchanged because Gln^2, Gln^3, and Gln^7 make the base-specific operator-site contacts in 434R (Q^{28}, Q^{29} and Q^{33} respectively). These residues have been left as such in the subsequent sequences also except when designed to alter the DNA recognition properties of the peptide. α_C contains a Q7R mutation that is expected to alter the operator site recognition by introducing a guanidinium group for potentially contacting a guanosine base.

The next series of peptides, α_E, α_F, α_G, α_H, and α_I, were synthesized as amides to further promote helicity. Lysine residues were introduced into the sequence to increase the overall affinity of the metal-peptide complex for DNA. The Lys residues serve to

Table 4.1. Family of Metal-Peptide Complexes Based on α_3-434R and the Charge on the Peptide.

Metal Complex	Metal-Peptide Sequence[a, b]	Net Charge on Peptide (pH 7.0)[c]
Sk	[Sk- COO^-]$^{2+}$	nil
Sk-α_3	[Sk- T**Q**Q**S**IE**Q**LQNG-COOH]$^+$	-2
Sk-α_A	[Sk- L**Q**Q**A**IE**Q**LQNAAAA-COOH]$^+$	-2
Sk-α_A-OMe[d]	[Sk- L**Q**Q**A**IE<u>**Q**</u>LQNAAAA-COOMe]$^{3+}$	0
Sk-α_C	[Sk- L**Q**Q**A**IE**R**LQNAAAA-COOH]$^{2+}$	-1
Sk-α_C-9[e]	[Sk- L**Q**Q**A**IE<u>**R**</u>LQNAAAA-COOH]$^+$	-2
Sk-α_C-9-OMe[d, e]	[Sk- L**Q**Q**A**IE<u>**R**</u>LQNAAAA-COOMe]$^{3+}$	0
Sk-α_E	[Sk-AL**Q**Q**S**IA**Q**LQNKAAA-$CONH_2$]$^{4+}$	+1
Sk-α_F	[Sk- T**Q**Q**S**KK**Q**LQNKAAA-$CONH_2$]$^{6+}$	+3
Sk-α_G	[Sk- T**Q**Q**S**KK**R**LQNKAAA-$CONH_2$]$^{7+}$	+4
Sk-α_H	[Sk- T**Q**R**S**KK**Q**LQNKAAA-$CONH_2$]$^{7+}$	+4
Sk-α_I	[Sk- T**Q**A**S**KK**Q**LQNKAAA-$CONH_2$]$^{6+}$	+3

[a] The skeletal complex has a free carboxylate. The carboxy terminus of the peptides is shown as an acid, amide or methyl ester as the case may be. The overall charge shown on each metal-peptide complex is that expected at pH 7.0 assuming that the imine-nitrogen atoms on the phi ligands are fully protonated. [b] The Gln residues in bold make base specific contacts with the DNA operator site in 434R. [c] Overall charge on the peptide after accounting for the N-terminus being blocked as an amide by Sk. [d] The underlined glutamate has its side chain carboxylate protected as a methyl ester. [e] The underlined Arg has a tosyl protecting group on the side chain.

make the peptides more hydrophilic. The increased positive charge on the peptide and the lower hydrophobicity was also designed to prevent association with the skeletal complex. α_F, α_G, α_H, and α_I form an independent family of peptides and are different from one another by a single amino acid change at one of the three Gln residues implicated in DNA recognition. α_F is the wild-type analogue in this family of peptides.

4.2.3. Synthesis of the Metal-Peptide Complexes. The details of the synthesis and characterization of the metal-peptide complexes studied in this chapter (Table 4.1) have been described in Chapter 2.

4.2.4. Circular Dichroism Studies. Samples of the free peptides were prepared in 1 mM or 8.2 mM sodium phosphate buffer (pH 7.0). The solutions of the metal-peptide complexes were prepared in water. CD spectra were recorded in a 1.0 or 0.1 cm path length cell at room temperature. A 200 μL solution of the desired compound was sequentially diluted by the addition of 2,2,2-trifluoroethanol (TFE) and the CD spectra recorded at each dilution. The mole % of TFE added and the effective concentration of the peptide was calculated for each dilution to determine the % helicity of the peptide. A UV-Visible spectrum of the metal-peptide complexes was taken after the final TFE titration point to ensure that the calculated and observed final concentration was the same.

4.2.5. NMR Spectra of Sk and Sk-α_F in the Presence of TFE. Sk (3.1 mg) and Sk-α_F (4.5 mg) were lyophilized twice in D_2O and finally taken up in 400 μL of D_2O. The approximate initial concentration of the two complexes was calculated to be: Sk 8.33 mM and Sk-α_F 3.72 mM. 1D ^1H NMR spectra were recorded with solvent (D_2O) presaturation and referenced to the water signal at 4.65 ppm before presaturation. d_3-TFE was titrated in (up to 21.5 mole% for Sk-α_F and 14.6 mole% for Sk) and the spectra retaken. The concentration of Sk changes to 5.12 mM and that of Sk-α_F to 1.86 mM

during the course of the titration. The spectra were plotted as stacked plots and referenced to a singlet at 2.62 ppm that is common to all the spectra to avoid the uncertainty of the water peak shifting with a change in pH during the TFE titration.

4.2.6. Photocleavage of DNA Restriction Fragments. Plasmid pUC-18 was digested with *EcoR I* restriction endonuclease. The digested plasmid was 3'-end-labeled by treatment with Klenow fragment of the DNA polymerase I and [α-^{32}P]dATP, dTTP, dCTP and dGTP.[14] A separate batch of the digested plasmid was treated with calf intestinal alkaline phosphatase and 5'-end labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP.[14] After labeling, the DNA was digested with *Pvu II* to yield a 180 and a 140 base pair fragment. The 180 base pair fragment was isolated by 8% nondenaturing preparative polyacrylamide gel electrophoresis followed by electroelution. In a separate set of labeling reactions, the plasmid was first cut with *Pvu II* and the blunt 3'-end was labeled with [α-^{32}P]ddATP using terminal deoxynucleotidyl transferase. The blunt 5'-end was labeled as before with T4 polynucleotide kinase and [γ-^{32}P]ATP. In this manner the 180-mer restriction fragment was sequentially end-labeled at all four ends for photocleavage and MPE-footprinting assays.

Cleavage reactions were carried out in 20 μL total volume contained in 1.7 mL presiliconized eppendorf tubes. The reaction mixtures contained the appropriate amounts of calf thymus DNA, labeled restriction fragment (~ 40,000 cpm per sample) and rhodium complex in 50 mM sodium cacodylate, pH 7.0. Reaction mixtures were incubated and then irradiated at 313 nm as indicated. Controls without metal complex were also irradiated in parallel to test for light damage. Reaction samples may contain added reagents such as 5 mM $MnCl_2$, 8 mole% TFE, 10 vol% CH_3CN or $MgCl_2$ as indicated.

All samples were ethanol precipitated after irradiation by addition of 10 μL of 7.5 M NH_4OAc and 130 μL of ethanol. The precipitated DNA was dried and resuspended in

3 μL of 80% formamide loading buffer. The samples along with Maxam-Gilbert G+A and C+T sequencing reactions[14, 15] were loaded onto an 8% denaturing polyacrylamide gel and electrophoresed at 1500 V for approximately 120 min until the bromophenol blue dye ran off the bottom of the gel. The gel was transferred to paper and dried prior to autoradiography.

4.2.7. Photocleavage Experiment Under Differing Incubation Conditions. Reaction samples containing ca. 40,000 cpm of 3'-end-labeled 180-mer and 1, 5, 10 or 50 nM Sk-α_F were made in 50 mM sodium cacodylate, pH 7.0 buffer. Cold calf thymus DNA (CT-DNA) was added as a carrier under the following incubation conditions:

(a) The samples were incubated at room temperature (rt) for 28h in the presence of a 1:1 ratio of rhodiun complex to bp CT-DNA.

(b) The samples were incubated at (rt) for 28h with only the labeled fragment. CT-DNA was added to the samples (1 rhodium:50 bp CT-DNA) and the samples allowed to equilibrate for a further 5 min prior to irradiation.

(c) The samples were incubated at rt for 28h in the presence of a 1:50 bp ratio of rhodium complex to CT-DNA.

(d) The reaction samples were made in 10 mM Tris·HCl (pH7.5), 10 mM $CaCl_2$ and 10 mM $MgCl_2$ in lieu of the sodium cacodylate buffer. The samples were incubated at rt for 28h in the presence of a 1:50 bp ratio of rhodium complex to CT-DNA.

Following incubation, the samples were irradiated using 313 nm light at rt for 5 min except the 1 nM rhodium samples which were irradiated for 10 min. All the reaction samples were ethanol precipitated after irradiation, taken up in 3 μL loading buffer and electrophoresed on an 8% denaturing polyacrylamide gel at 1500 V until the bromophenol blue dye had run off the bottom of the gel. The gel was transfered to paper and dried prior to autoradiography.

4.2.8. Protocol for MPE-Fe Footprinting. (Adapted from reference 16.) Reaction samples (20 μL) consisting of the rhodium complex (5 μM Δ-Sk or 1 μM Δ-Sk-α_F), 50 mM sodium cacodylate (pH 7.0) buffer and 40,000 cpm end labeled fragment were allowed to equilibrate in the dark at rt for 18 h. Carrier CT-DNA was added (final rhodium:DNA bp ratio was 1:10) 5 min prior to the MPE-Fe reactions. A freshly prepared solution of MPE-Fe made by mixing the appropriate amounts of MPE and ferrous ammonium sulfate was used in the reactions. 2 μL of 5 μM MPE-Fe (0.5 μM final concentration) was added to the reaction samples. The samples were briefly vortexed, spun down in a centrifuge and allowed to equilibrate for 5 min. 2 μL of 5 mM sodium ascorbate (0.5 mM final concentration) was then added to the reaction samples and the reaction allowed to proceed for 10 min. The reaction was quenched by addition of excess CT-DNA and ethanol precipitation. The footprinting reactions were analyzed on an 8% denaturing polyacrylamide gel as described above for the photocleavage reactions.

4.2.9. HPLC Analysis of the Photochemical Properties of the Metal-Peptide Complexes. (pH 7.0): Samples (4 μM, 20 μL total volume) of Sk, Sk-α_A and Sk-α_E were made in 5 mM $MnCl_2$ and 50 mM sodium cacodylate, pH 7.0. The control samples were incubated at 55° C for 20 minutes. The irradiation samples were incubated at 55° C for 5 min and then irradiated with 313 nm light at 55° C for 15 min. The samples were diluted to 100 μL by the addition of 80 μL of 0.1% TFA in water, vortexed and briefly spun down in a centrifuge. The entire sample (100 μL) was injected and analyzed by HPLC.

A Vydac C_{18} reverse phase column was used for the analysis of the irradiation product. The quantity of starting rhodium complex left intact was monitored at the characteristic phi-transition at 373 nm (0.005 AUFS). 0.1% TFA in water (Solvent A) and 0.1% TFA in CH_3CN (Solvent B) were used for eluting the complexes at a flow rate

of 4.5 mL/min under the following gradient: 0-4 min (15% B); 4-19 min (15-40% B); 19-21 min (40% B); 21-26 min (40-15% B); 26-30 min (15% B).

(pH 6.0 and 8.0): Samples of Sk, Sk-α_A, Sk-α_A-OMe and Sk-α_E were made as above without the addition of $MnCl_2$ using 50 mM sodium cacodylate buffer at the required pH. The samples were incubated and irradiated with 313 nm light at rt for 5 min each. The control samples were incubated at rt for 10 min. HPLC analysis was as above.

4.2.10. Photocleavage of Oligonucleotide Duplexes. The two strands of a 51 base pair oligonucleotide duplex with the sequence 5'-CGCGATACAAGATGCAAGATACGA-GATAGACGATAGAAGATAGGAGATCGC-3' (51T) and it's complement (51B) were 5'- or 3'-end labeled individually using T4 polynucleotide kinase and [γ-^{32}P]ATP or terminal deoxynucleotidyl transferase and [α-^{32}P]ddATP respectively. The duplex contains six related 4 bp sites (represented in bold letters above) to test for operator site discrimination by the metal-peptide complexes. The cold single strands corresponding to the 51 base-pair duplex and the labeled strand were annealed by heating up to 90° C followed by slow cooling to room temperature. Reaction samples contained rhodium complex and the oligonucleotide duplex (including ~100,000 cpm of labeled oligonucleotide) in a 1:51 bp ratio in 50 mM sodium cacodylate, pH 7.0. The absolute concentration of the rhodium complex was 200, 100, 50 or 10 nm. Samples were incubated and irradiated at 313 nm as indicated. Approximately 20,000 cpm/sample were transferred into fresh eppendorf tubes, dried and then resuspended in 3 μL loading buffer.

Piperidine Treatment: 3 μL aliquot of an irradiation sample was transfered to another eppendorf tube, diluted with 7 μL water and reacted with 25 μL of 10% piperidine (in water) at 90° C for 30 min. The reaction mixture was lyophilized in a speed vac thrice (with addition of 20 μL of water between each time) and finally taken up in 3 μL of loading buffer.

The samples were loaded onto a 20% denaturing polyacrylamide gel and electrophoresed at 2000-2400 V until the bromophenol blue dye was ca. 25 cm from the wells. The gel was then transferred to a film and stored at -70° C during autoradiography.

4.2.11. Gel Electrophoretic Mobility Shift Assay. (Adapted from reference 17.) Incubation mixtures (20 μL total volume) containing 200 nM duplex (51 bp), 0-3 μM rhodium complex and 25 mM sodium cacodylate, pH 7.0 were allowed to equilibrate at 4° C for 3h. 2 μL of a 10x non-denaturing loading buffer (50% glycerol in 1 mM EDTA and 10 mM tris, pH 8.0) without any dye, was added to the samples. The samples were loaded onto a running 5% preparative non-denaturing polyacrylamide (29:1) gel in 0.25x TBE. The samples were electrophoresed in 0.25x TBE at 4° C (in the cold room) at a constant voltage of 200 V for ca. 1h. Bromophenol blue and xylenecyanol dyes were run in adjacent lanes as markers. The bromophenol blue dye was 6-8 cm from the wells when the electrophoresis was stopped and the gel transferred to paper and dried for autoradiography.

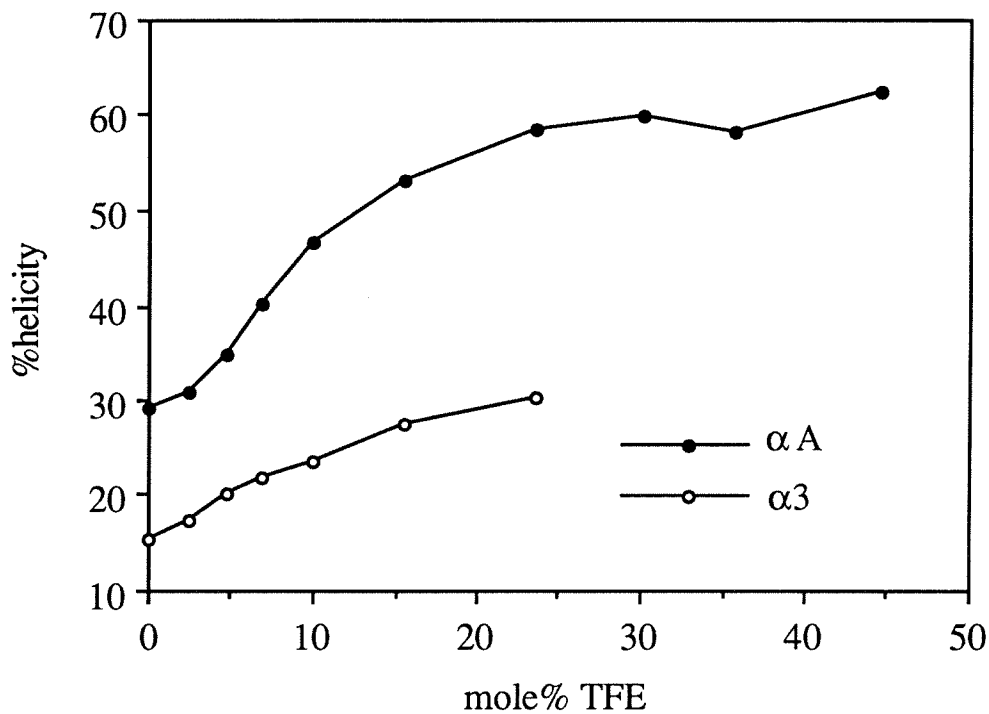
4.3. Results

4.3.1. Circular Dichroism of the Metal-Peptide Complexes. The peptide, α_3, is an α-helix in 434R. Figure 4.2 shows that α_3 in isolation is not helical and has a low helical propensity even upon the addition of 23.6 mole% TFE.* Earlier studies in the literature with the BR-peptides of bzip proteins indicated that a great deal of helix induction is observed when the peptides are bound to their cognate DNA sites.[19] We can envisage a similar effect in our system. However, the inherent helicity and the helical propensity in

---

* α-helicity of peptides in aqueous TFE is often indicative of their potential for α-helix formation. Up to 90% of maximal helicity is induced in peptides upon the addition of 10 mole% TFE[18].

Figure 4.2. CD data showing %helicity of the free peptides, α_3 and α_A, as a function of increasing TFE concentration. The spectra were recorded on a Jasco J 600 spectrometer using a 0.1 cm pathlength cell and averaged over 8 scans. 210 μM α_3 and 171 μM α_A were used in the experiment. %helicity was calculated from the observed θ_{222} values as described in Chapter 3. Note the greater inherent helicity and the helical propensity for α_A compared to α_3.


70
60
50
40
30
20
10
%helicity
0
10
20
30
40
50
mole% TFE
α A
α3

α_3 is too low to begin with so that other conformations of the peptide may be energetically more accessible.

Figure 4.2 also shows the observed helical content of α_A as a function of increasing amounts of TFE. The inherent helicity, in the absence of TFE, is clearly greater than that for α_3. More significantly, α_A has a strong propensity to be helical with a maximal helicity of ~ 63% as opposed to α_3 that has only ~ 30%. An increase in helicity upon addition of TFE is also observed for the metal-peptide complex (Figure 4.3a). Furthermore, the helical propensity of Sk-α_A is slightly greater than that for α_A (Figure 4.3b).

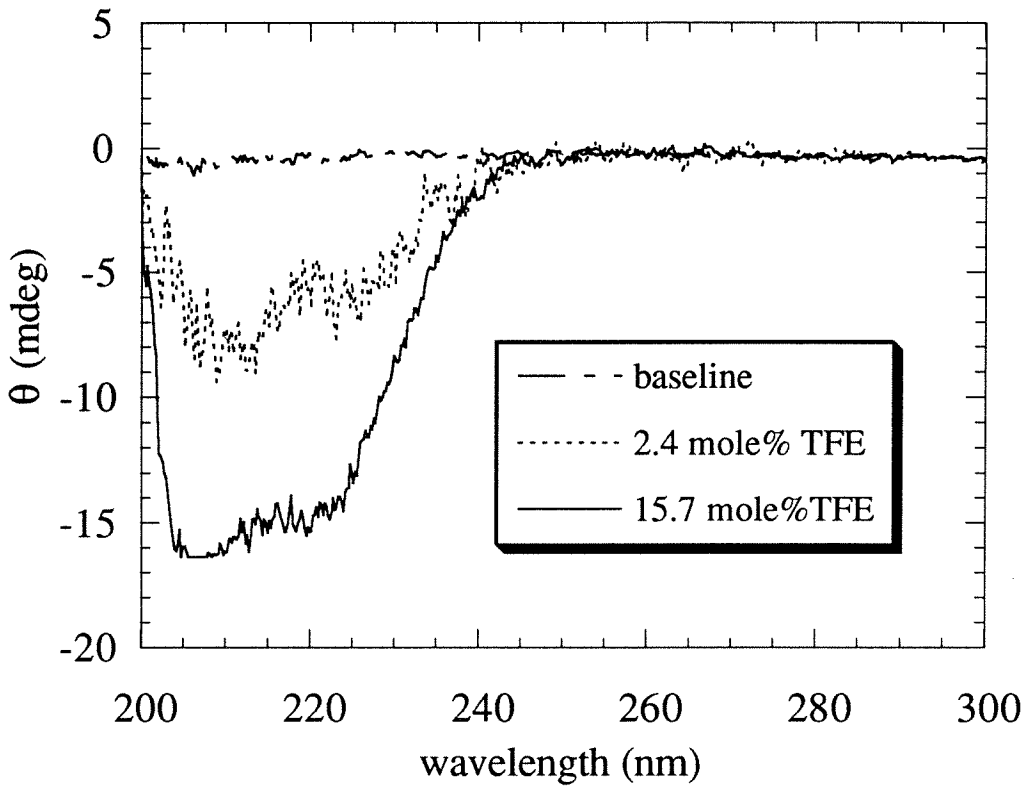
Similar CD studies were carried out with Sk-α_F, Sk-α_G, Sk-α_H, and Sk-α_I. The results are summarized in Table 4.2 and compared to Sk-α_A. Although the metal-peptide complexes have low inherent helicity, in the presence of TFE and likely DNA, they have the potential to be significantly helical.

4.3.2. NMR Spectra of Sk-α_F as a Function of TFE Concentration. The folding transition in Sk-α_F induced upon adding increasing amounts of TFE was monitored by 500 MHz ^{1}H-NMR spectroscopy. A control study was carried out with Sk to differentiate between TFE induced changes in chemical shifts of Sk protons from those induced by a folding transition in the peptide. We wanted to ascertain whether the peptide associates with/wraps around the rhodium center; or if the two moieties, i.e., Sk and the peptide, are independent in solution. A preliminary assignment of the different peptide residue resonances in the 1-D spectrum of Sk-α_F is shown in Figure 4.4a. The resonances were assigned on the basis of random coil ^{1}H chemical shift values for the common amino acid residues.[20]

The aromatic resonances and the aliphatic signals from the linker are nearly identical in Sk and Sk-α_F. Only the C6H (singlet) resonance of phen' is slightly perturbed (Figure 4.4b). The point of attachment of the linker is at the 5 position of phen

Figure 4.3. **(a)** Top: CD spectra of Sk-α_A recorded with increasing TFE concentration in the sample. A 195 μM solution of Sk-α_A was made in water and the spectra (average of 8 scans) were recorded on a Jasco J500 spectrometer using a 0.1 cm path length cell. The baseline is a spectrum of water. Only the spectra of samples containing 2.4 and 15.7 mole% TFE are shown for clarity. Note the increase in helicity upon addition of TFE. The %helicity levels off at 15.7 mole% TFE.

(b) Bottom: CD data from the TFE titration of Sk-α_A and α_A shown as a plot of the %helicity as a function of TFE concentration. Sk-α_A has a slightly higher helical propensity compared to α_A.


5
0
-5
-10
-15
-20
θ (mdeg)
200
220
240
260
280
300
wavelength (nm)
baseline
2.4 mole% TFE
15.7 mole%TFE

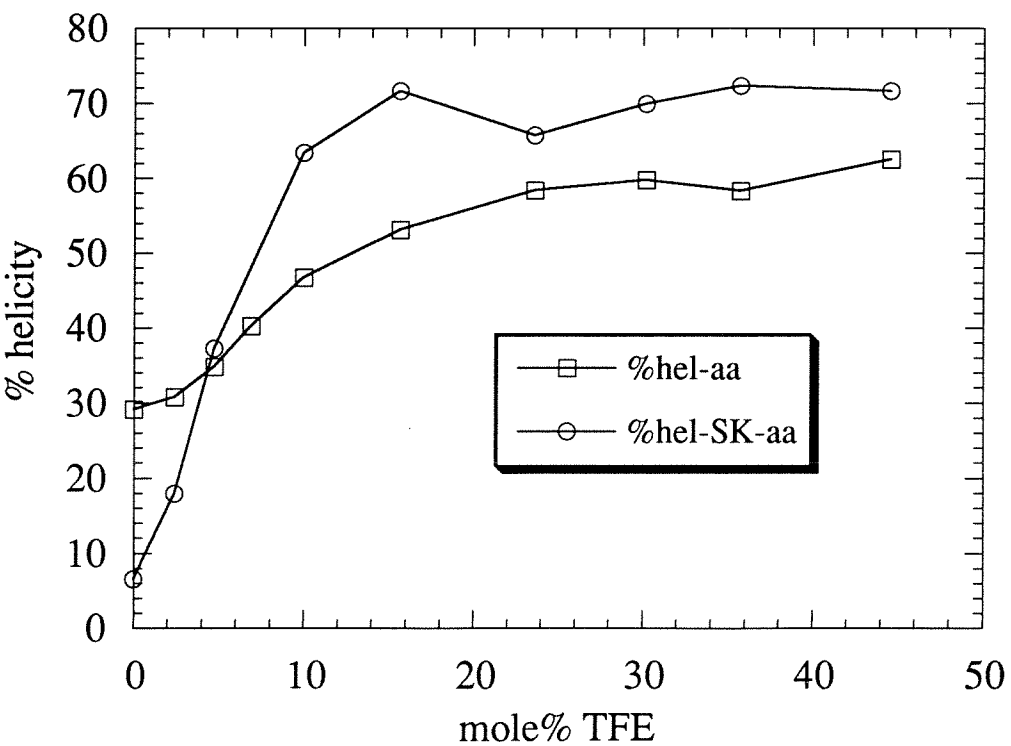

80
70
60
50
40
30
20
10
0
% helicity
0
10
20
30
40
50
mole% TFE
%hel-aa
%hel-SK-aa

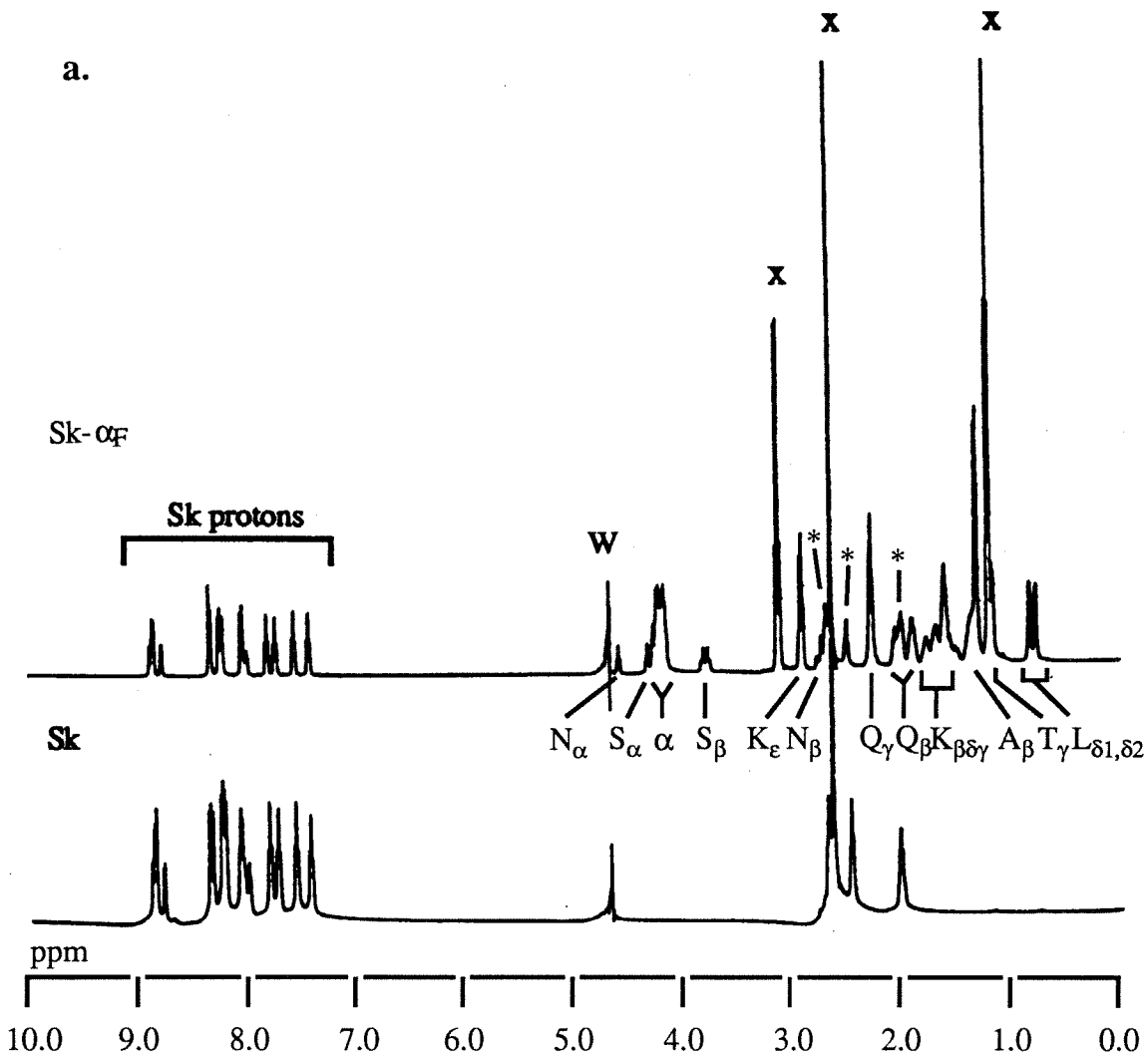
Table 4.2. Comparison of the Helical Propensities of Sk-α_F, Sk-α_G, Sk-α_H, and Sk-α_I with Sk-α_A as a Function of Mole % TFE.

Sample [a]	0.0 mole% TFE	4.7 mole%TFE	10.0 mole% TFE
Sk-α_A	6.5 [b]	37.2	63.4
Sk-α_F	12.6	16.4	44.4
Sk-α_G	8.2	9.9	40.4
Sk-α_H	5.4	14.1	45.9
Sk-α_I	12.5	14.5	53.9

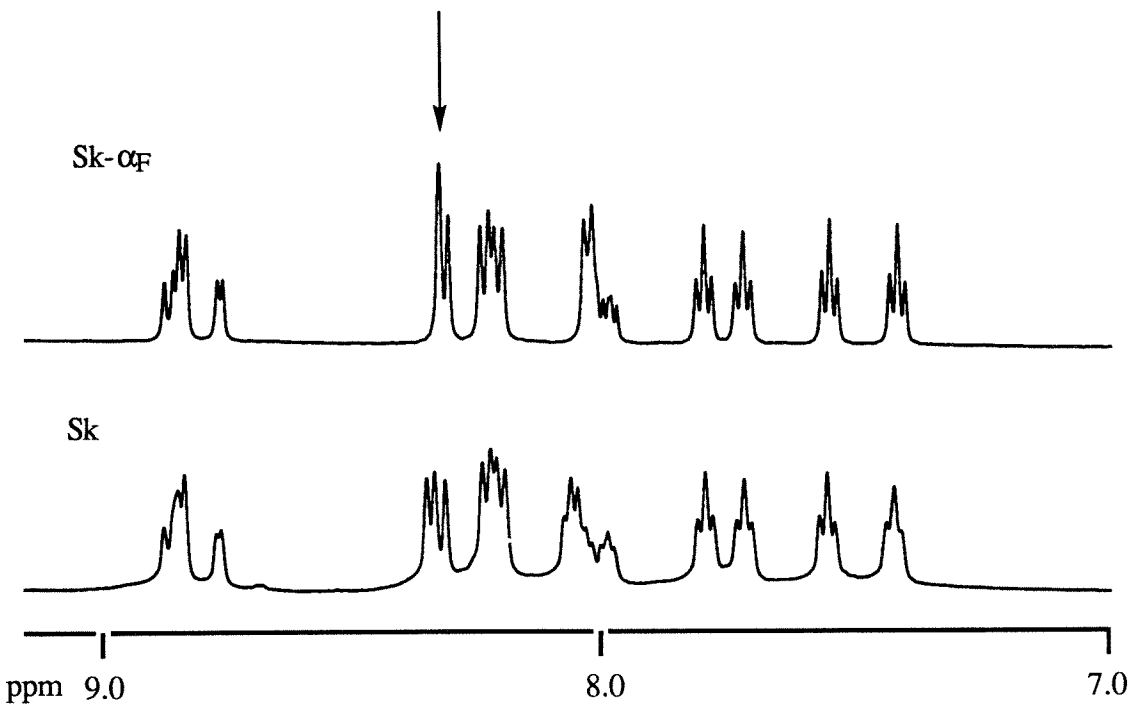
[a] CD spectra were recorded in a 0.1 cm path length cell. The initial concentrations of the different complexes in water were: Sk-α_A (196 μM); Sk-α_F (25 μM); Sk-α_G (25 μM); Sk-α_H (25 μM); Sk-α_I (25 μM). [b] The numbers represent %helicity calculated for each sample as in Chapter 3.

Figure 4.4. **(a)** 500 MHz ^{1}H-NMR spectra of Sk and Sk-α_F in D_2O. A preliminary assignment of the peptide side chain resonances is shown. The asterisks denote the positions of the linker protons (on phen') in the spectrum of Sk-α_F. W: denotes the residual water peak after solvent presaturation; x: denotes peaks that are neither from Sk nor from the peptide.

(b) An expansion of the aromatic region of the spectra of Sk and Sk-α_F. The spectra of the two complexes are identical except for the C6H resonance (arrow). The near identity of the spectra indicates that Sk and peptide are independent in solution and not associated in some way.


a.
x
x
x
Sk-αF
Sk protons
W
*
*
*
Sk
Nα Sα α Sβ Kε Nβ Qγ Qβ Kβδγ Aβ Tγ Lδ1,δ2
ppm
10.0 9.0 8.0 7.0 6.0 5.0 4.0 3.0 2.0 1.0 0.0

b.



and upon linkage of the peptide, the environment around the adjacent C6H is likely to be perturbed the most. The near identity of the Sk component of the spectra in the two complexes indicates that the peptide and the rhodium center are distinct. TFE induces downfield shifts (0.1-0.2 ppm) in both the aromatic and aliphatic protons of Sk (Figure 4.5).

The spectrum of Sk-α_F shows large changes in both the aromatic (Sk) and aliphatic (peptide) regions on titrating in TFE (Figure 4.6). The different resonances tend to broaden with increasing TFE. The α-protons exhibit the greatest differences indicating a TFE induced structural transition. Information from CD studies above links this change to a random coil to α-helix transition.

4.3.3. DNA Recognition by the Metal-Peptide Complexes. The parent metal-peptide complex, Sk-α_3, recognizes 5'-ACAA-3' sites on restriction fragments as shown in Figure 4.7. These sites are the putative operator half sites for the 434R protein. Photocleavage is observed at the 5'-base and internal to the operator site. We then investigated the DNA recognition properties of a family of metal-peptide complexes (Figure 4.8). The Sk-α_n complexes cleave DNA upon photoactivation and show an enhancement in cleavage at 5'-ACAA-3' sites over the skeletal complex. The predominant enhancement in cleavage is seen at the 5'-base to the 5'-ACAA-3' sequences. 5'-ACA-3' sites are also targeted with cleavage observed at the central cytosine. This 5'-ACAA-3' selectivity is reduced when photocleavage is carried out in the presence of 5 mM $MnCl_2$. Minor variations in the intensity of cleavage are observed between the complexes. Sk-α_A cleaves more strongly at the 5'-ACAA-3' sites than does Sk-α_3. Also, comparison between Sk-α_C and Sk-α_C-9 (Sk-α_C-R-Tos) indicates a preference for Sk-α_C. These slight differences in cleavage intensity may be explained by differences in the helicity amongst the peptides and the presence of the bulky tosyl-protecting group on Arg in Sk-α_C-9. The differences are heightened in the presence of 5 mM $MnCl_2$ which

Figure 4.5. 500 MHz ^{1}H-NMR spectra of Sk with increasing amounts of TFE. Note the slight downfield shifts in the resonances upon addition of TFE.

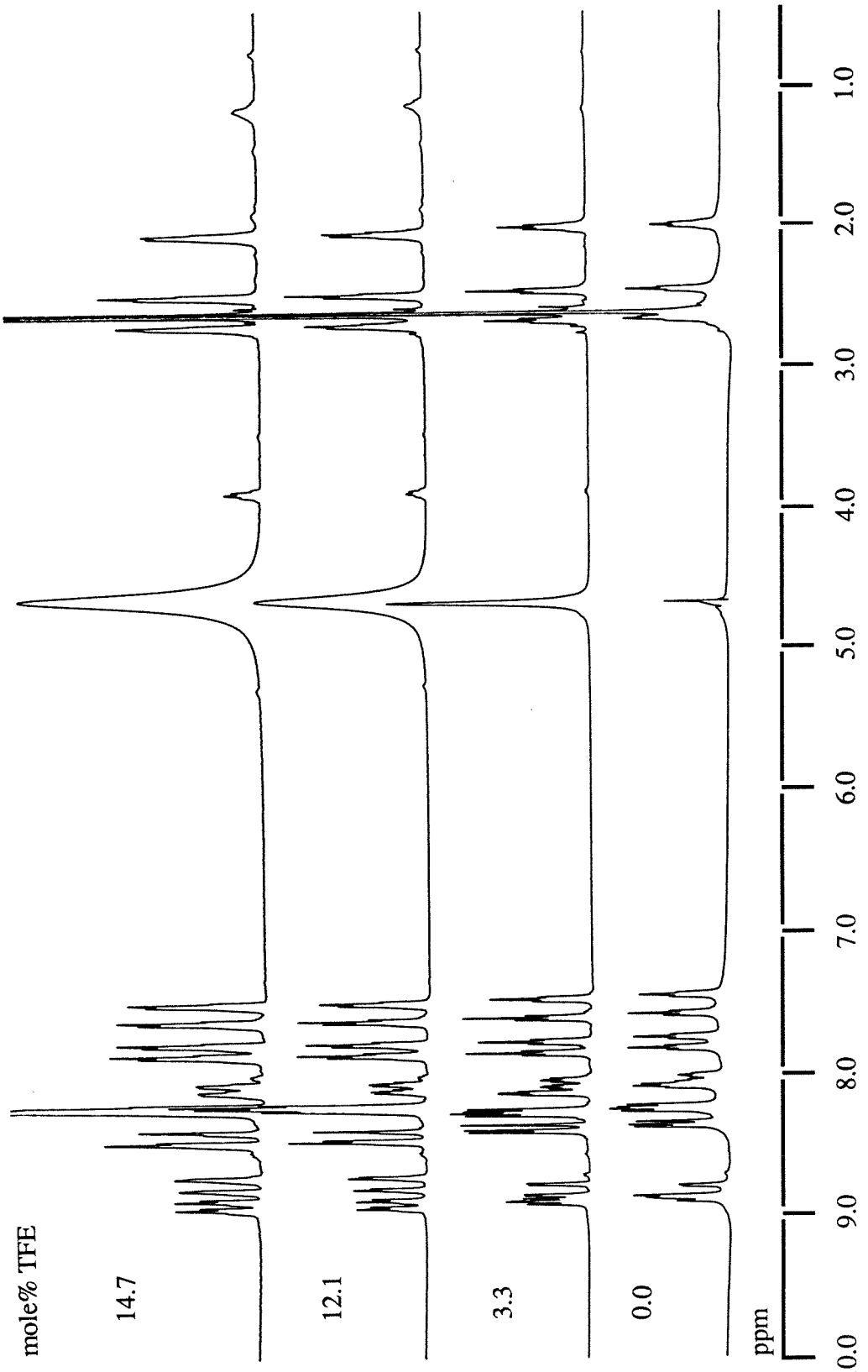

mole% TFE
14.7
12.1
3.3
0.0
ppm
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0

Figure 4.6. 500 MHz ^{1}H-NMR spectra of Sk-α_F with increasing amounts of TFE. Note the large shifts in the aromatic and aliphatic resonances with increasing concentration of TFE. The α-protons (3.8-4.6 ppm) show the greatest changes suggestive of a conformational change in the peptide.

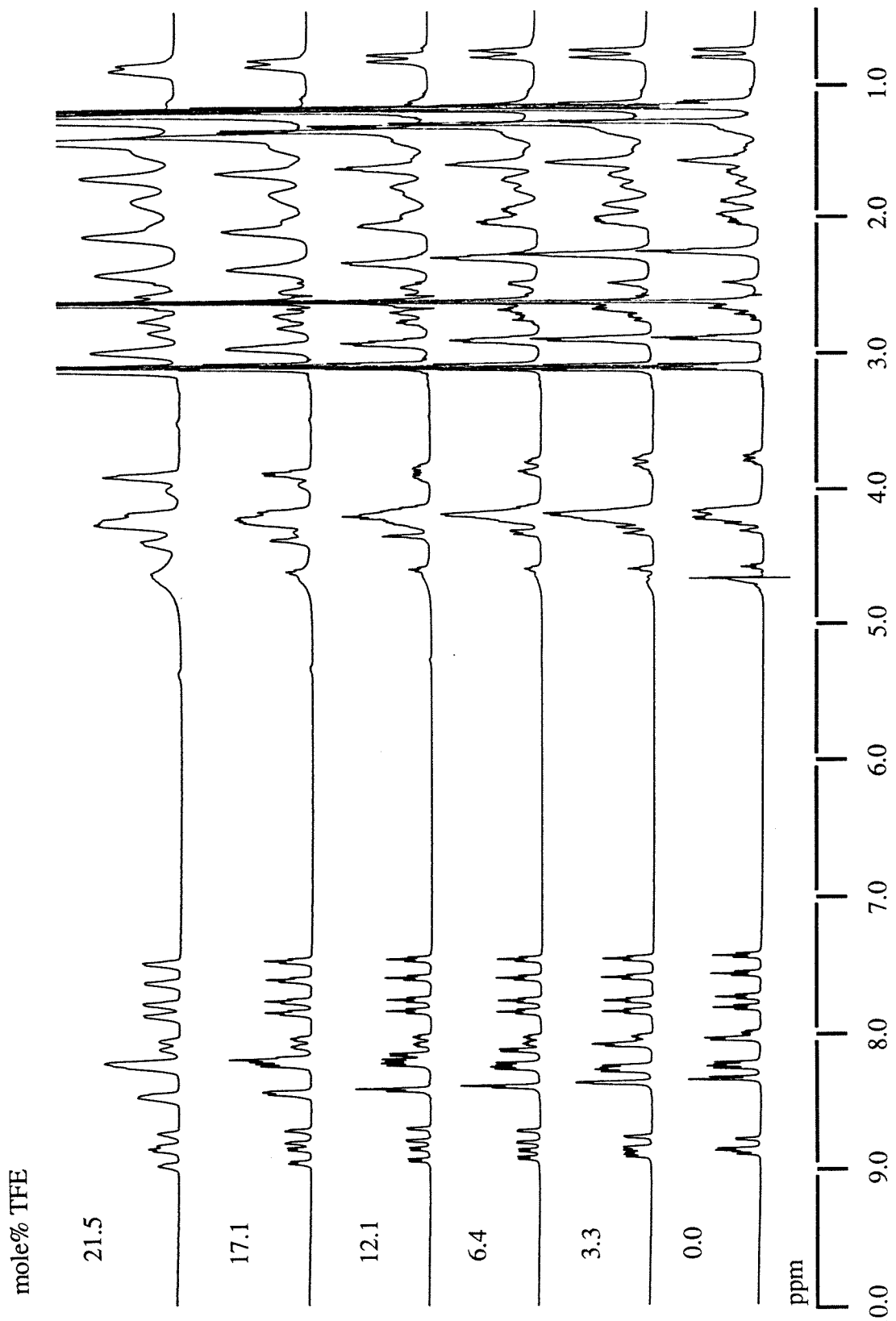

mole% TFE
21.5
17.1
12.1
6.4
3.3
0.0
ppm
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0

Figure 4.7. Photocleavage on a 5'-end labeled restriction fragment by Sk-α_3. Reaction samples contained 100 μM rhodium complex, 100 μM bp CT-DNA in 50 mM sodium cacodylate, pH 7.0. The samples were incubated for 2 h and irradiated (313 nm) for 5 min at rt. C+T: Maxam-Gilbert C+T sequencing reaction; LC: DNA fragment irradiated in the absence of rhodium complex; M (skel): fragment irradiated in the presence of Sk; M+Peptide: fragment irradiated in the presence of Sk-α_3. The 434R operator sites have been shown at the bottom and bases cleaved by Sk-α_3 are marked by arrows. The bold arrows represent the N-->C directionality of α_3 in 434R when bound to the operator half site.

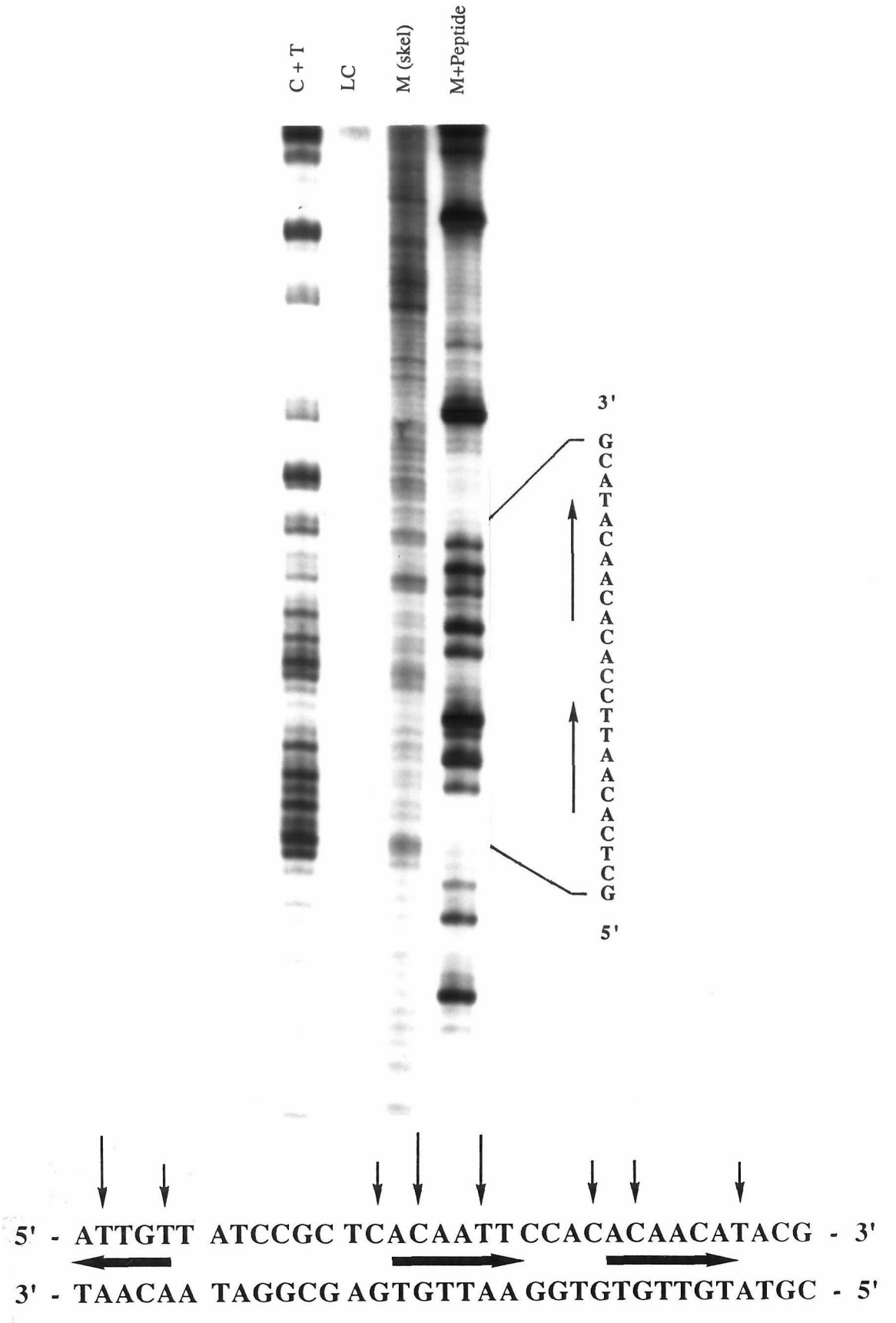
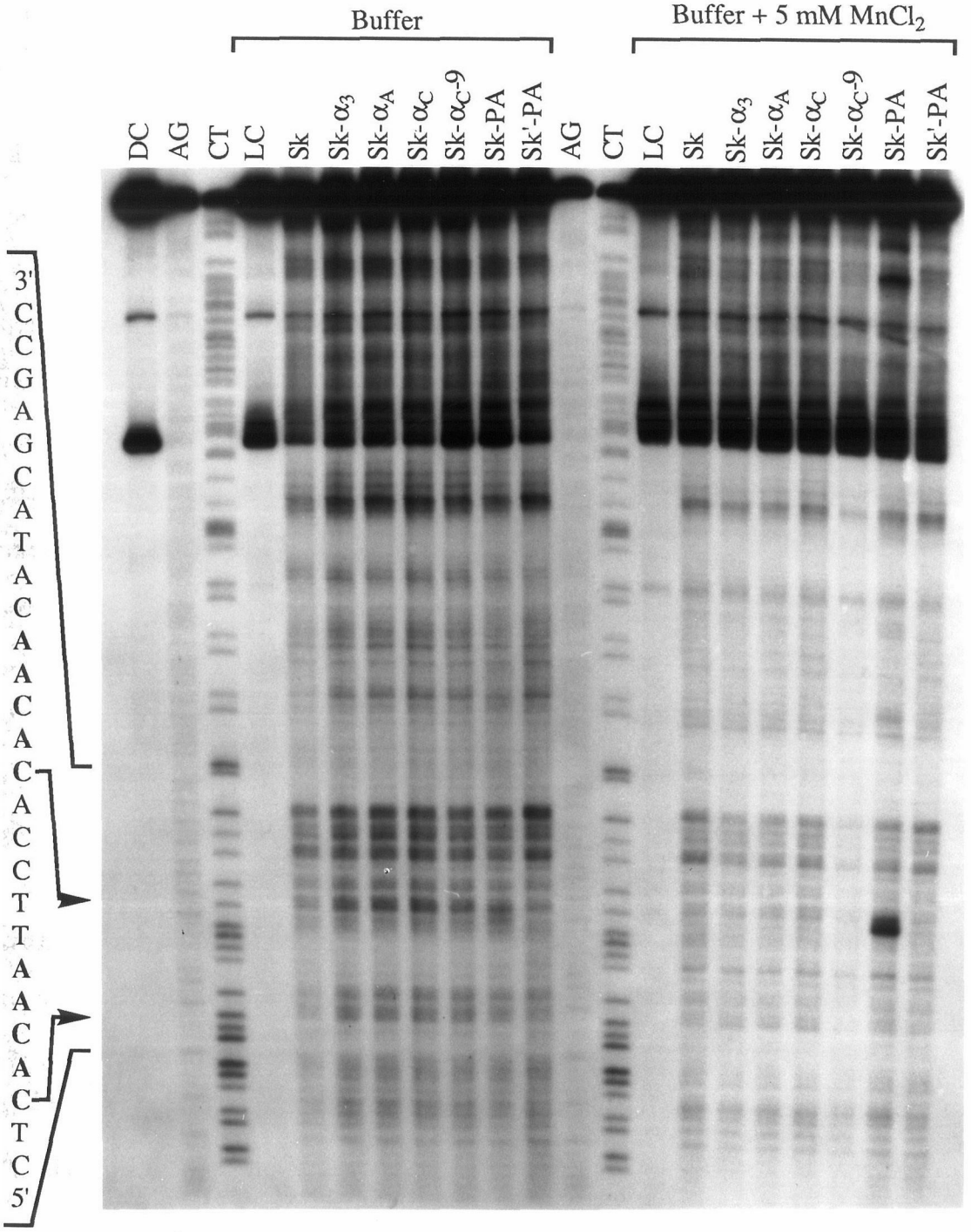

C + T
LC
M (skel)
M+Peptide
3'
GCATACAACACACCTTAACACACTCG
5'
5' - ATTGTT ATCCGC TCACAATT CCACACAACATACG - 3'
3' - TAACAA TAGGCG AGTGTTAA GGTGTGTTGTATGC - 5'

Figure 4.8. Photocleavage of the 5'-end labeled (*EcoR I**) 180 mer restriction fragment from pUC-18 by a family of Sk-α_n complexes. Reaction samples contained 1 μM rhodium complex and 50 μM bp CT-DNA. Samples were incubated for 5 min and irradiated (313 nm) for 15 min at the appropriate temperature. Cleavage lanes on the left are reactions at rt in 50 mM sodium cacodylate buffer, pH7.0. Cleavage lanes on the right are reactions at 55° C in 50 mM sodium cacodylate buffer, pH 7.0 and 5 mM $MnCl_2$. DC: untreated fragment; LC: fragment irradiated in the absence of rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions respectively. The 5'-NACAA-3' sites are shown in bold. Note the enhanced cleavage at these sites by the Sk-α_n family of metal-peptide complexes compared to Sk, Sk-PA or Sk'-PA. The effect of $MnCl_2$ at the elevated temperature is only dramatic for Sk-PA and is discussed in Chapter 3. $MnCl_2$ tends to reduce photocleavage by the Sk-α_n family of metal-peptide complexes.


Buffer
Buffer + 5 mM MnCl2
DC
AG
CT
LC
Sk
Sk-α3
Sk-αA
Sk-αC
Sk-αC-9
Sk-PA
Sk'-PA
AG
CT
LC
Sk
Sk-α3
Sk-αA
Sk-αC
Sk-αC-9
Sk-PA
Sk'-PA
3'
C
C
G
A
G
C
A
T
A
C
A
A
C
A
C
C
A
C
C
T
T
A
A
C
A
C
A
C
T
C
5'

may serve to screen out weaker binding sites. In particular, Sk-α_C-9 cleaves less at all sites while Sk-α_C retains some 5'-ACAA-3' selectivity possibly because of the increased binding affinity from the added positive charge on Arg.

Also shown in Figure 4.8 is a comparison of photocleavage with Sk-PA (Sk-AANVAISQWERAA-CONH2) and Sk'-PA. Sk' is a modified $[Rh(phi)_2(bpy)]^{3+}$ complex in which the peptide is linked to a 2,2'-bipyridyl ligand via a carboxybutyl linker. This linker is conformationally more flexible than the glutaryl linker on phen'. These complexes do not show a preference for the 434R operator sequences over the skeletal complex in contrast to the metal-peptide complexes containing the α_3-434R peptide sequence. Sk-PA shows its characteristic 5'-CCA-3' selectivity in the presence of 5 mM $MnCl_2$ (detailed in Chapter 3) while Sk'-PA has the Sk-like selectivity under these conditions also.

A screen of different reaction conditions for photocleavage by the α_3-434R family of peptides was carried out (data not shown) to optimize conditions for maximal specificity. The conditions included photocleavage in the presence of 8 mole% TFE, 5 mM $MnCl_2$ or 50 mM $MgCl_2$. A reaction temperature screen with photocleavage by Sk-α_A at 4° C, 24° C, 37° C and 55° C in the presence or absence of 8 mole% TFE was also carried out. Furthermore, photocleavage under varying rhodium:DNA ratios ranging from 1:1 to 1:250 bp DNA and rhodium concentration from 0.05 μM to 10 μM was investigated. None of the differing conditions led to any significant enhancements in cleavage specificity, again in contrast to the metal-peptide complexes discussed in Chapter 3.

4.3.4. Effect of Net Charge of the Peptide Moieties on DNA Recognition. Table 4.1 lists the net charge on the peptide and the overall charge on the metal-peptide complex at pH 7.0. The peptides on Sk-α_3, Sk-α_A, Sk-α_C, and Sk-α_C-9 have a net negative charge. We considered the possibility that this negative charge may interfere with binding of the

peptide to the DNA through electrostatic repulsion from the phosphate backbone. The peptide may possibly remain in solution away from the DNA while the carrier rhodium complex is intercalated in the DNA helix. Such a mode of binding, in which the influence of the peptide in determining the binding site is minimal, can account for the high background skeletal complex-like cleavage seen for these metal-peptide complexes compared to that seen for the Sk-P_n family discussed in Chapter 3. We methylated the two free carboxylates (Glu^6 and the C-terminus) in Sk-α_A and Sk-α_C-9 and studied their DNA recognition. Methylation of the carboxylates neutralizes the charge on the peptides. In addition, Sk-α_E was synthesized with a Glu to Ala substitution, a Lys residue and a C-terminal amide thereby having a net +1 charge on the peptide.

Making the peptides charge neutral or positively charged helps in the affinity and possibly specificity. A comparison of Sk-α_A and Sk-α_A-OMe (Figure 4.9) indicates an increase in affinity and selectivity for cleavage at 5'-NACAA-3' sites by Sk-α_A-OMe over Sk and Sk-α_A. Sk-α_E shows similar behaviour and cleavage specificity is comparable to that seen with Sk-α_A-OMe. (Compare, e.g., lanes 2 and 4 in Figure 4.10.) Again, conditions of 5 mM $MnCl_2$ at 55° C lower the overall cleavage by all the metal complexes. Methylation of Sk-α_C-9 assists in DNA selectivity moderately but the additional protecting groups (R-tos and E-OMe) introduce steric clashes and lower the photocleavage (Figure 4.10 lanes 3, 8, 13 or 18) compared to Sk-α_E or Sk-α_A-OMe.

4.3.5. Effect of Net Charge of the Peptide Moieties on the Photochemical Behaviour of the Metal-Peptide Complexes. A series of analytical HPLC experiments were performed on the metal-peptide complexes to study their stability under photoirradiation conditions and get an estimate for the rates of photoanation of the phi ligands. The stability was monitored as a function of the pH of the irradiation solution. The skeletal complex is stable upon irradiation at pH 7.0 or 8.0. A 10% loss in the intensity of the signal is observed under the irradiation conditions (Figure 4.11). In contrast, the tethered

Figure 4.9. Comparison of photocleavage of the 5'-end labeled (*EcoR I**) 180 mer restriction fragment from pUC-18 by Sk, Sk-α_A and Sk-α_A-OMe showing the effect of methylation of the carboxylates on the peptide. Reaction samples contained 1 μM rhodium complex, 50 μM bp CT-DNA and 5 mM $MnCl_2$ in 50 mM sodium cacodylate buffer, pH 7.0. Samples were incubated for 5 min and irradiated (313 nm) for 15 min at 55° C. DC: untreated fragment; LC: fragment irradiated in the absence of rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions respectively. The 5'-**N**ACAA-3' sites are shown in bold. Note the enhanced cleavage at these sites by Sk-α_A-OMe compared to Sk or Sk-α_A. The bold arrow points to a retarded band migrating slower than the uncut fragment in the photocleavage lanes containing the metal-peptide complexes but not Sk.

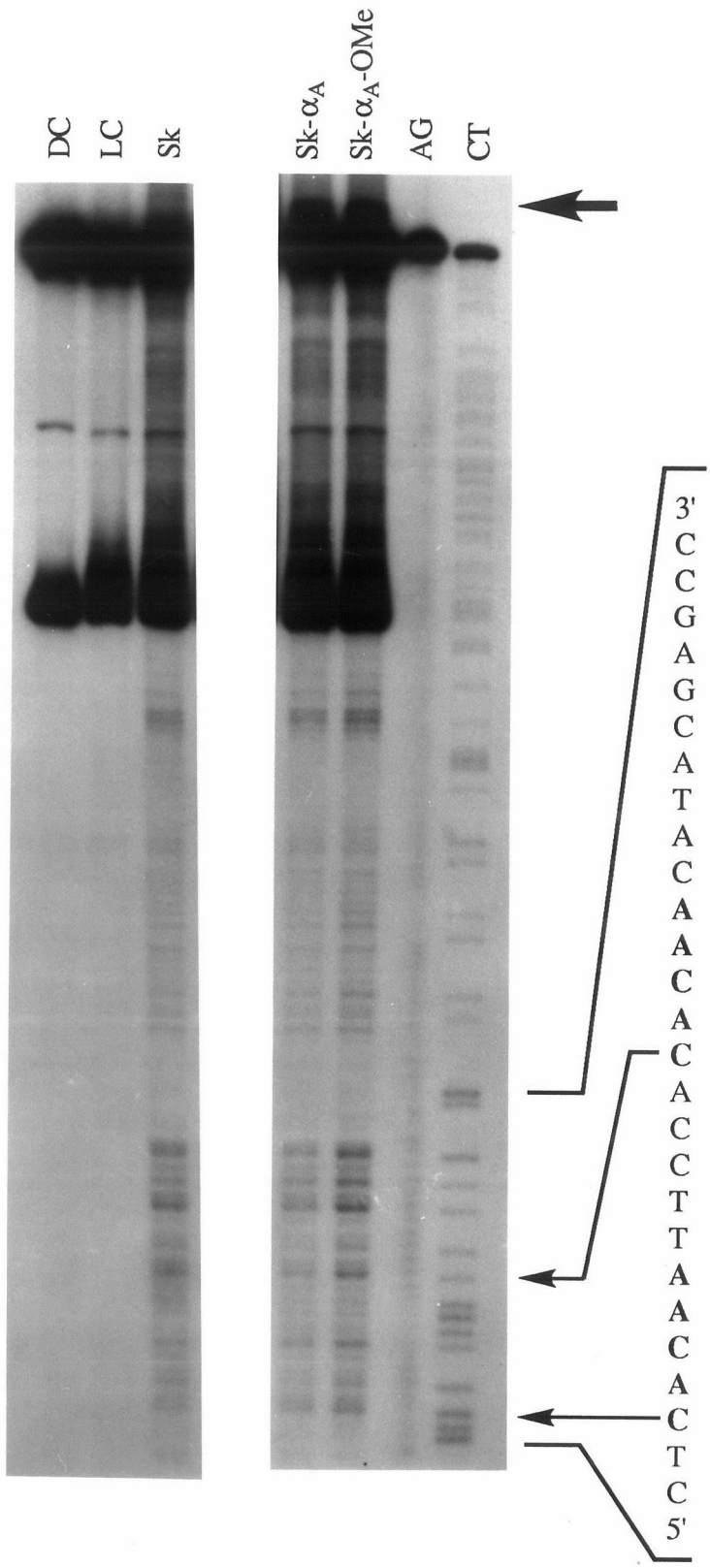

DC
LC
Sk
Sk-α_A
Sk-α_A-OMe
AG
CT
3'
C C G A G C A T A C A A C A C A C C T T A A C A C T C
5'

Figure 4.10. Comparison of photocleavage of the 5'-end labeled (*EcoR I**) 180 mer restriction fragment from pUC-18 by a Sk-α_n family of metal-peptide complexes containing neutral or positively charged peptides. Reaction samples contained 1 or 5 μM rhodium complex and 1:50 bp ratio of rhodium to CT-DNA in 50 mM sodium cacodylate buffer, pH 7.0. Samples were incubated for 1 min and irradiated (313 nm) for 5 min at rt. A parallel set of samples containing 5 mM $MnCl_2$ were incubated and irradiated at 55° C for 5 min each. DC: untreated fragment; LC: fragment irradiated in the absence of rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions respectively. Sk: lanes 5, 10, 11 and 16; Sk-α_A-OMe: lanes 4, 9, 12, 17; Sk-α_C-9-OMe: lanes 3, 8, 13, 18; Sk-α_E: lanes 1-2, 6-7, 14-15, 19. The solid bars indicate positions of the 5'-ACAA-3' sequences. (*) denotes base 5'- to a putative 434R operator site and shows enhanced cleavage compared to Sk. (#) denotes 5'-TGT/ACA-3' sequences that also show enhanced cleavage compared to Sk. The bold arrow points to a retarded band migrating slower than the uncut fragment in the photocleavage lanes containing the metal-peptide complexes but not Sk.

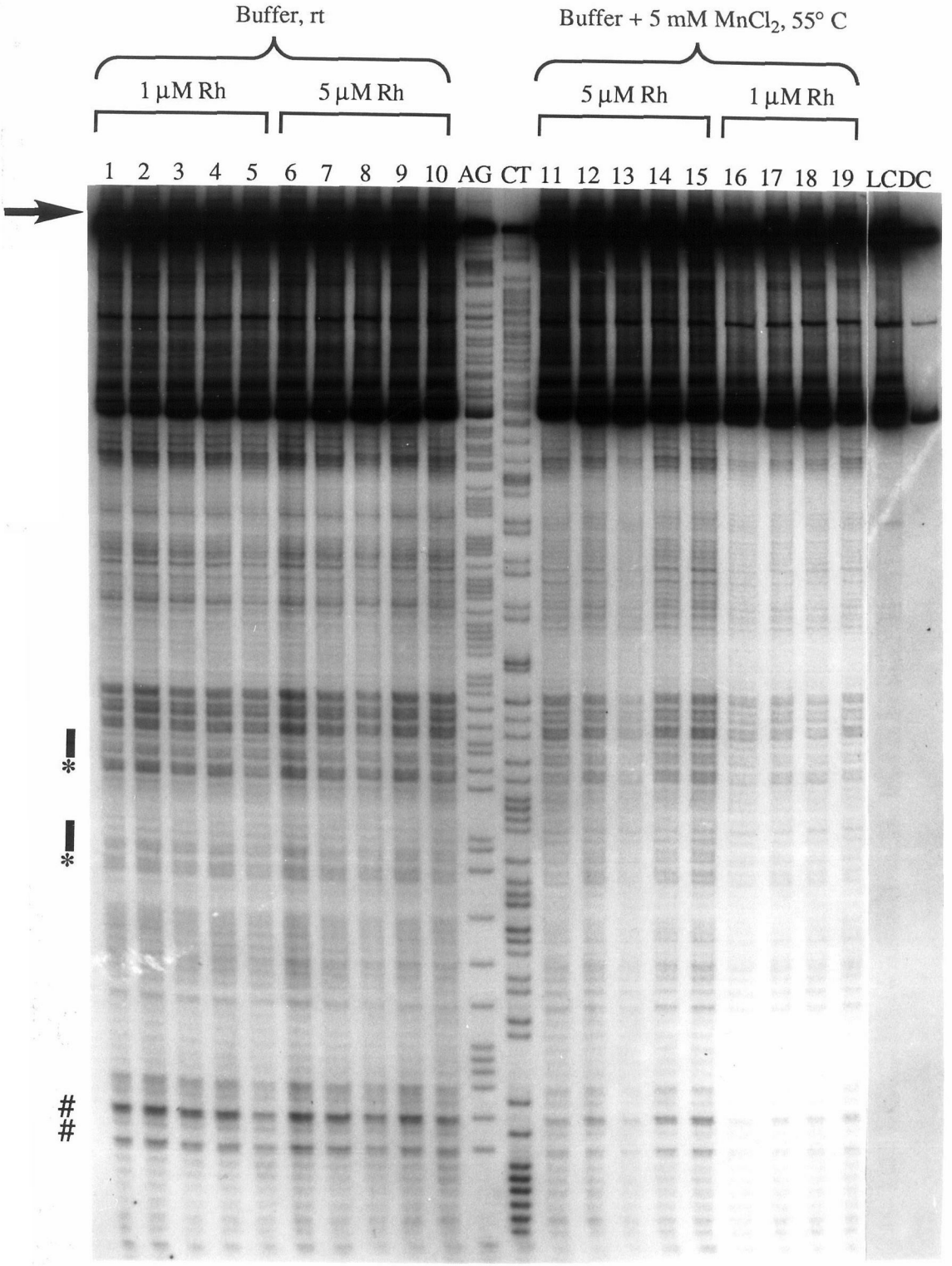
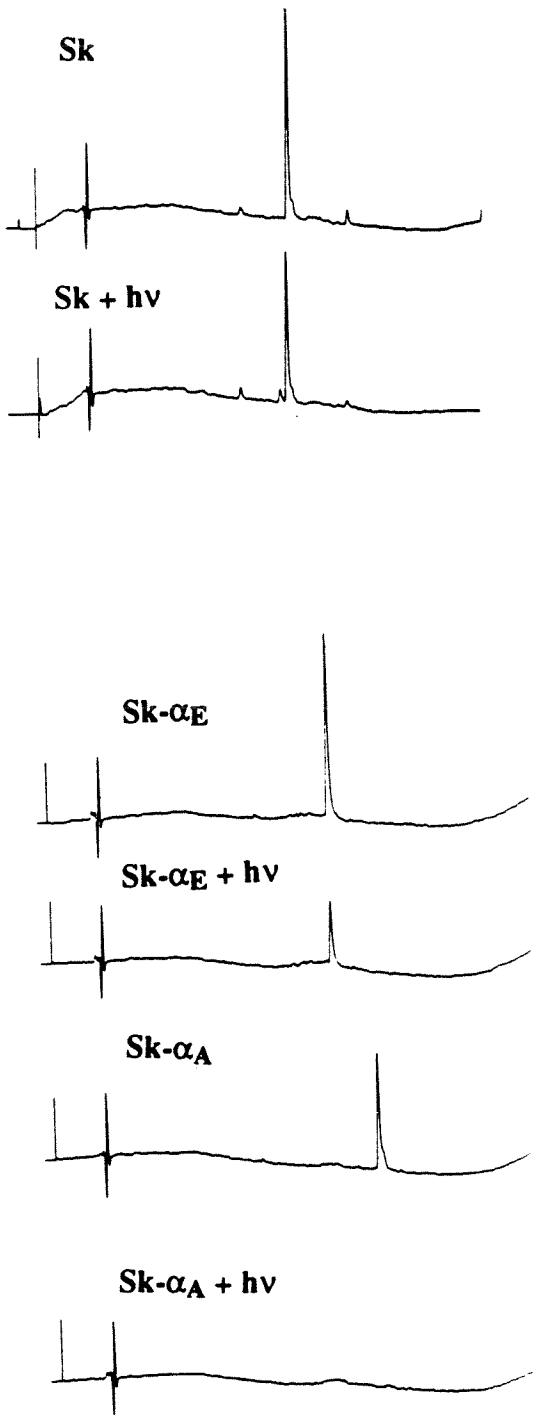

Buffer, rt
Buffer + 5 mM $MnCl_2$, 55° C
1 μM Rh
5 μM Rh
5 μM Rh
1 μM Rh
1 2 3 4 5 6 7 8 9 10 AG CT 11 12 13 14 15 16 17 18 19 LCDC
*
*
#
#

Figure 4.11. HPLC profiles of samples of Sk, Sk-α_A and Sk-α_E before and after irradiation with 313 nm light in sodium cacodylate buffer, pH 7.0. The peaks correspond to intact metal-peptide complexes. Note the complete loss in signal for Sk-α_A upon irradiation compared to a 65% loss for Sk-α_E and a 10% loss for Sk.


Sk
Sk + hν
Sk-αE
Sk-αE + hν
Sk-αA
Sk-αA + hν

peptide complexes show dramatic differences depending on the charge of the attached peptide.

Sk-α_A, having a net -2 charge on the peptide, is completely degraded after 5 min of irradiation at pH 6.0, 7.0 or 8.0 (Figure 4.11 and 4.12). Sk-α_A-OMe is substantially more stable under the conditions showing only ca. 40% loss in signal intensity at pH 6.0 and 8.0 (Figure 4.12). Sk-α_A-OMe shares the same peptide sequence with Sk-α_A with the exception that the two carboxylates on Sk-α_A have been methylated, thereby neutralizing the charge on the peptide. Sk-α_E with a net charge of +1 on the peptide is likewise stable and shows ca. 40% loss in signal intensity at pH 6.0 or 8.0 upon irradiation for 5 min (Figure 4.12). A loss of 65% of the signal is observed for Sk-α_E after irradiation for 15 min at pH 7.0 (Figure 4.11). It should be pointed out here that Sk, Sk-α_3 and Sk-α_A are stable in sodium cacodylate buffer, pH 7.0 over 14 h in the absence of photoirradiation. The complexes are stable for longer periods of time under acidic conditions (pH<5.0).

4.3.6. Formation of an Unusual DNA Product. An additional observation of note is the formation of a severely retarded band under a variety of reaction conditions in the photocleavage lanes containing the Sk-α_n family of metal-peptide complexes (marked by a bold arrow in Figure 4.9 and 4.10). The band is not observed in reactions with Sk. The presence of 1 mM $MgCl_2$ or 8 mole% TFE or 10 vol% CH_3CN does not appear to perturb the formation of the product as demonstrated in the reaction with Sk-α_E (Figure 4.13). The retarded band migrates slower than the full fragment on 8% denaturing polyacrylamide gels. The band is present in reactions with both 3'- and 5'- end-labeled DNA fragments and is only seen upon irradiation. Metal-peptide complexes mixed with DNA in the dark (without photoirradiation) do not show the formation of the retarded band on 8% denaturing polyacrylamide gels. The amount of the retarded band grows with concentration of the metal-peptide complex and correlates with the total

Figure 4.12. HPLC profiles of samples of Sk, Sk-α_A, Sk-α_A-OMe, and Sk-α_E before and after irradiation with 313 nm light in sodium cacodylate buffer at pH 6.0 and 8.0. Again, only Sk-α_A shows complete loss of signal after 5 min of irradiation.

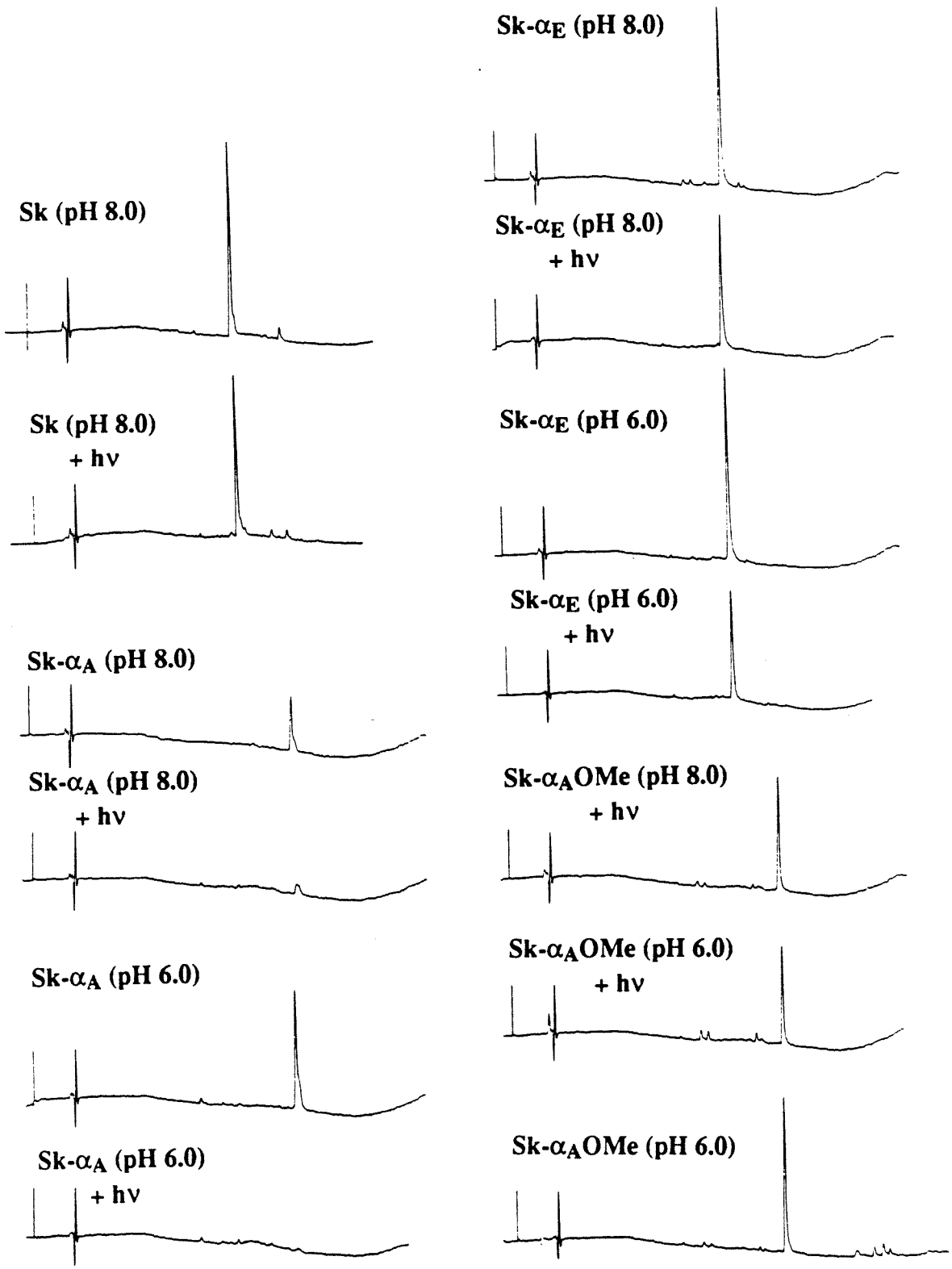
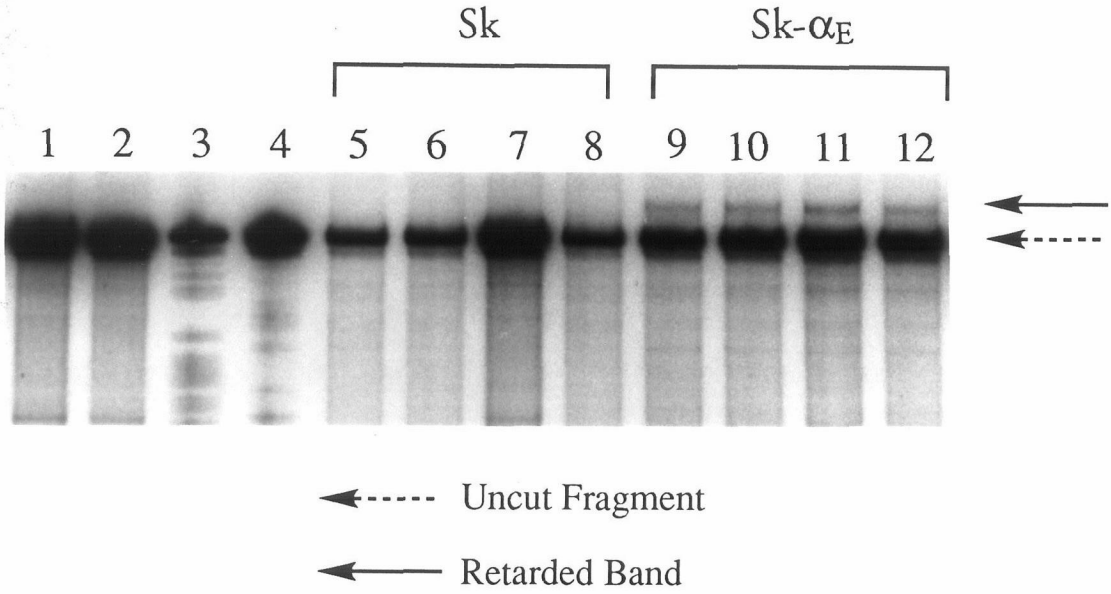

Sk (pH 8.0)
Sk (pH 8.0) + hν
Sk-αA (pH 8.0)
Sk-αA (pH 8.0) + hν
Sk-αA (pH 6.0)
Sk-αA (pH 6.0) + hν
Sk-αE (pH 8.0)
Sk-αE (pH 8.0) + hν
Sk-αE (pH 6.0)
Sk-αE (pH 6.0) + hν
Sk-αAOMe (pH 8.0) + hν
Sk-αAOMe (pH 6.0) + hν
Sk-αAOMe (pH 6.0)

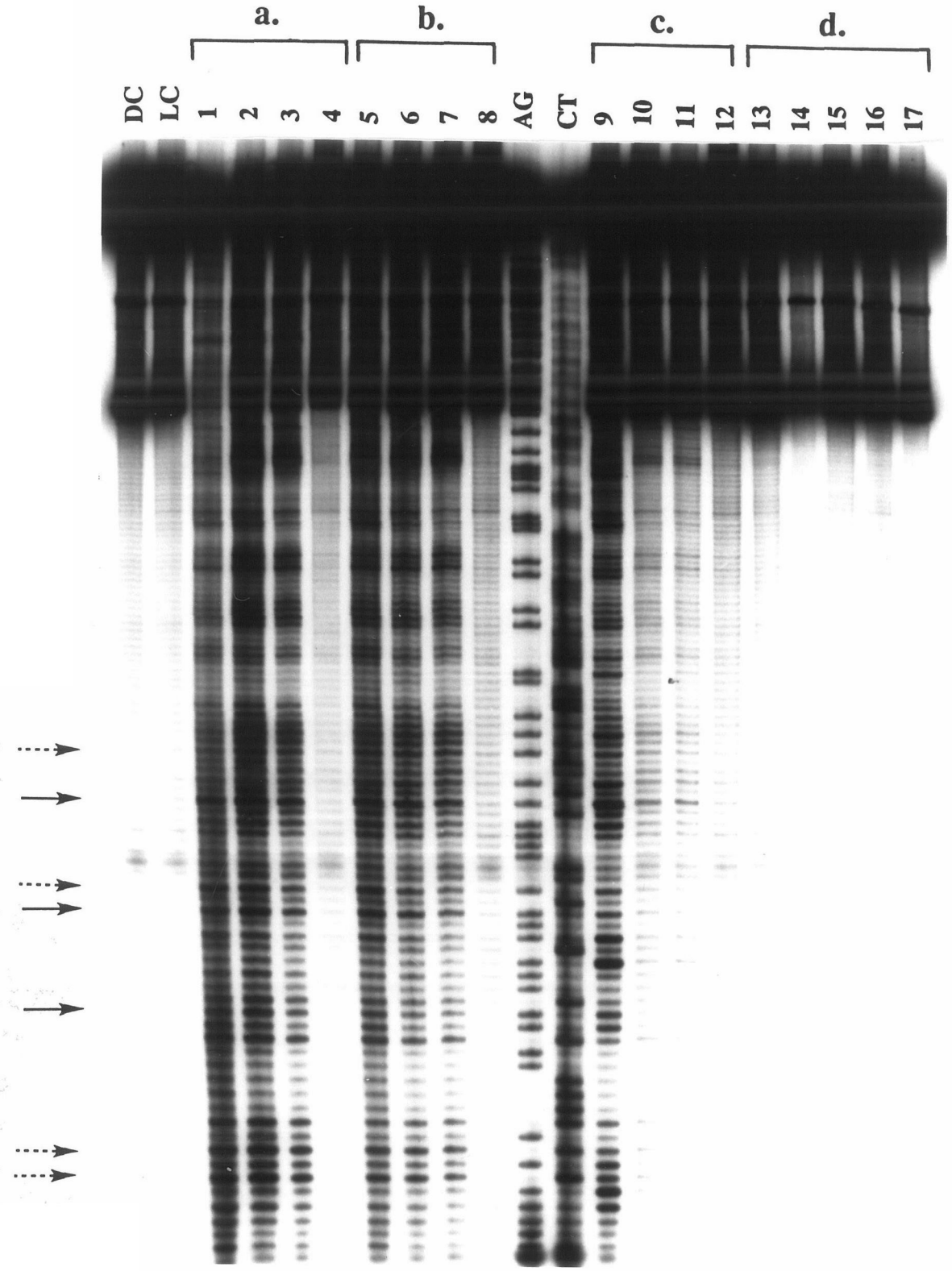
Figure 4.13. Enlarged autoradiogram of the top portion of an 8% denaturing polyacrylamide gel showing the uncut fragment (continuous arrow) and the formation of a retarded band (dashed arrow) following irradiation of the 3'-(*EcoR I**) end labeled 180-mer DNA fragment in the presence of Sk and Sk-α_E. Reaction samples contained 1 μM of rhodium complex and 1 μM bp CT-DNA (1:1 ratio of rhodium: bp CT-DNA) in 50 mM sodium cacodylate buffer, pH 7.0. The samples also contained added reagents as indicated. The samples were incubated for 2 min at 55° C followed by 5 min at rt and then irradiated for 2 min (313 nm) at rt. DC: untreated fragment; LC: fragment irradiated without the rhodium complex. AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions respectively; lanes 5-8: photocleavage by Sk; lanes 9-12: photocleavage by Sk-α_E. The photocleavage reactions were in buffer alone (lanes 5 and 9) or in the presence of added 1 mM $MgCl_2$ (lanes 6 and 10); 8 mole% TFE (lanes 7 and 11); 10 vol.% CH_3CN (lanes 8 and 12). Note the formation of the retarded band in only the lanes containing Sk-α_E. The added reagents have little or no effect on the formation of the photoproduct.


Sk
Sk-αE
1 2 3 4 5 6 7 8 9 10 11 12
Uncut Fragment
Retarded Band

photocleavage seen in the reaction. These results suggest that the retarded band may be a DNA photoadduct of the metal-peptide complex or its fragment. Pre-irradiation of the metal-peptide complex followed by incubation and irradiation in the presence of DNA enhances the intensity of the retarded band slightly. Extraction of the reaction samples with MeOH or repeated ethanol precipitation does not eliminate the retarded band thereby minimizing the possibility of a non-covalent tightly bound metal-peptide-DNA species. Furthermore, the photoproduct is not affected by piperidine treatment. Piperidine treatment of DNA fragments can cleave DNA at abasic sites or at sites damaged by light. Indeed, we observe that other light induced retarded bands in the gels migrate faster than the full fragment after piperidine treatment (see Figure 4.22).

4.3.7. Effect of Incubation Conditions on DNA Photocleavage. Photocleavage on a 3'-end labeled DNA restriction fragment under differing incubation conditions was carried out with Sk-α_F to estimate the rate of dissociation (k_{off}) of bound metal-peptide complex from DNA (Figure 4.14). Strong photocleavage is seen at bases 5' to 5'-ACAA-3' and at 5'-ACA/TGT-3' sequences. The rhodium-peptide complex (50, 10, 5 or 1 nM) was preincubated with labeled fragment overnight. Excess cold carrier DNA was added to the samples as a competitor 5 min prior to irradiation. The 5 min incubation time is not sufficient to reduce the photocleavage intensity significantly suggesting a $k_{off} < 0.2$ min^{-1} (Figure 4.14a, b). A substantial reduction in the photocleavage intensity is observed when the rhodium-peptide complex is incubated with the hot fragment and excess cold carrier DNA simultaneously (Figure 4.14c). Similarly, no cleavage is observed when the reaction samples contain a 1:50 bp metal to cold carrier DNA ratio in 10 mM Tris·HCl (pH 7.5),10 mM $CaCl_2$, 10 mM $MgCl_2$ buffer (Figure 4.14d). Presumably, the relatively high salt (compared to 50 nM rhodium) and a low rhodium to DNA ratio (1:50 bp) competes out binding to the DNA under these conditions. The specificity does not change with incubation time.

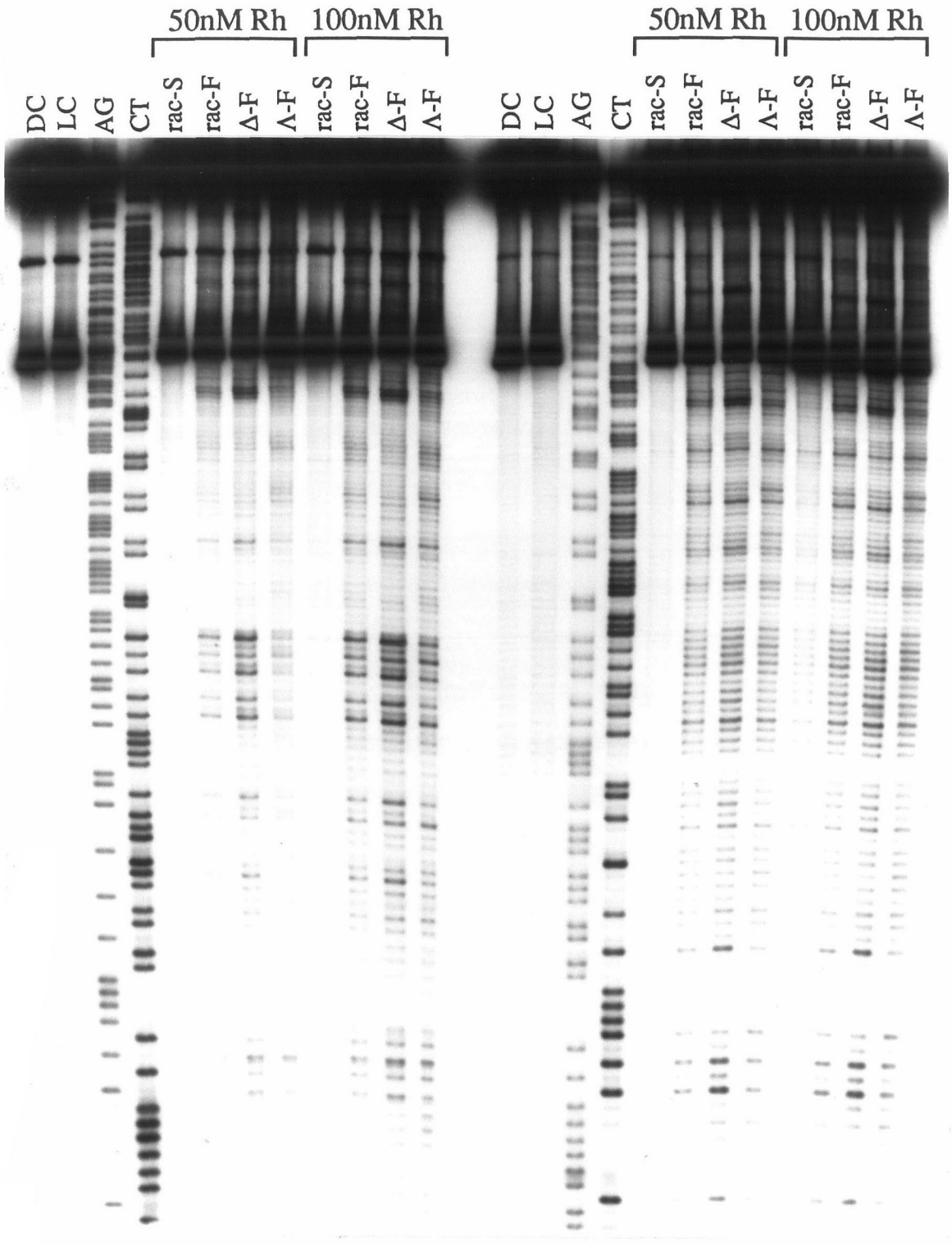
Figure 4.14. Photocleavage of the 3'-(*EcoR I**) end labeled 180-mer restriction fragment by Sk-α_F under differing incubation conditions. DC: untreated fragment; LC: fragment irradiated in the absence of the rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions; lanes 1, 5, 9 and 13: samples containing 50 nM Sk-α_F; lanes 2, 6, 10 and 14: samples containing 10 nM Sk-α_F; lanes 3, 7, 11 and 16: samples containing 5 nM Sk-α_F; lanes 4, 8, 12 and 16: samples containing 1 nM Sk-α_F; lane 17: sample containing the Tris·HCl, Ca^{2+}, Mg^{2+} buffer irradiated in the absence of rhodium complex. The reaction conditions are detailed in section 4.2.7. The solid arrow represents cleavage at 5'-NACAA/TTGTN-3' sequences and, the dashed arrows represent cleavage at 5'-ACA/TGT-3' sequences.


a.
b.
c.
d.
DC
LC
1
2
3
4
5
6
7
8
AG
CT
9
10
11
12
13
14
15
16
17

4.3.8. DNA Recognition by Diastereomers of Sk-α_F. We next studied the DNA recognition by the diastereomers of Sk-α_F. The two diastereomers (Λ- and Δ- at the rhodium center) were resolved on a cation exchange column using $K_3[Co(l\text{-cysu})_3]$ as the chiral resolving agent.[22] Our initial modeling studies and enantioselectivity observed with $[Rh(phi)_2(bpy)]^{3+}$ and $[Rh(phen)_2(phi)]^{3+}$ suggest that the delta diastereomer would be best oriented to deliver the attached peptide along the DNA major groove.[23] The predicted diastereoselectivity is observed in photocleavage gels using the diastereomers of Sk-α_F (Figure 4.15). Δ-Sk-α_F shows a 2 to 3-fold enhancement in photocleavage over Λ-Sk-α_F at the strongest sites, namely 5'-NACAA-3', 5'-ACA/TGT-3' and 5'-ATG/CAT-3'. DNA sites showing preferential Λ-diastereoselectivity, though weak by comparison to Δ-sites, tend to lie at the ends of purine/pyrimidine stretches. At these low concentrations, racemic Sk did not show appreciable cleavage.

Photocleavage by the diastereomers of Sk-α_F and the enantiomers of Sk are compared in Figure 4.16. A 10-fold lower concentration of Sk-α_F was used in the experiment to try and offset the increased non-specific affinity compared to Sk resulting from the greater positive charge on Sk-α_F. The delta enantiomers of both complexes cleave strongly at 5'-ACA/TGT-3' sequences. However, the strongest enhancement in cleavage by Δ-Sk-α_F is seen at 5'-CACAA-3' (two sites); 5'-TGTGTGAA-3'; 5'-TAATGAGT/ACTCATTA-3' and 5'-ATAAAG-3'. At these sequences cleavage by Δ-Sk-α_F exceeds that by Δ-Sk even at a 10-fold lower concentration. The differences in the lambda enantiomers of the two complexes follows the opposite trend. Λ-Sk shows some site selectivity for purine/pyrimidine tracts but site selectivity and cleavage by Λ-Sk-α_F is minimal at these low (50, 10 nM) concentrations. Presumably, the greater steric bulk of the attached peptide oriented into the DNA phosphate backbone lowers the overall binding constant of the lambda diastereomer of Sk-α_F. Compared to Sk, the peptide introduces ca. 10-fold increase in affinity for the strong sites by the delta diastereomer and ca. 2-fold increase in diastereoselectivity above that observed for Sk.

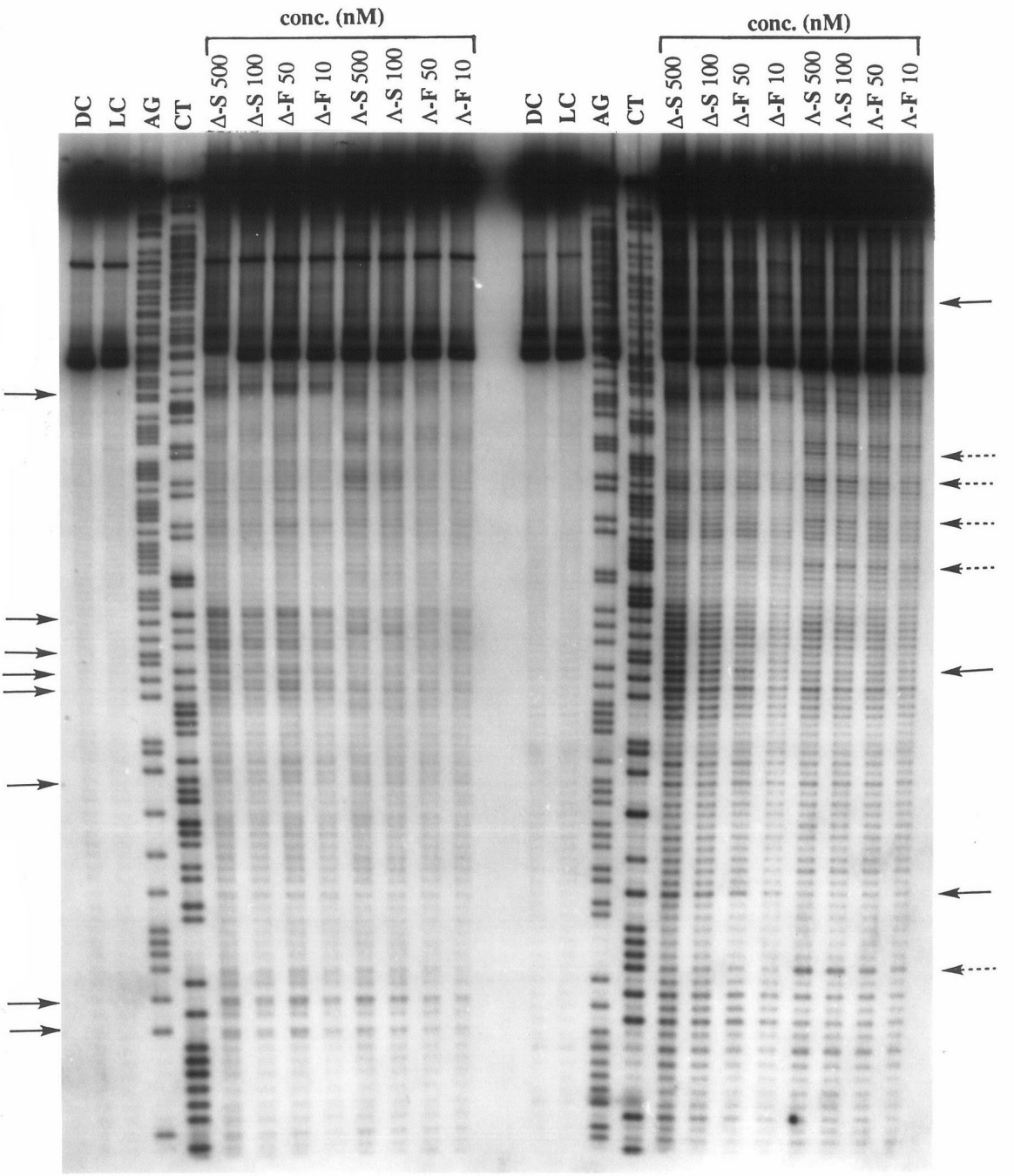
Figure 4.15. Photocleavage of the 5'- and 3'- (*EcoR I**) end labeled 180-mer fragment by racemic Sk and the diastereomers of Sk-α_F. Reaction samples contained 50 or 100 nM rhodium complex and a 1:50 bp ratio of rhodium to CT-DNA in 50 mM sodium cacodylate, pH 7.0. The samples were incubated overnight in the dark at rt and irradiated (313 nm) for 5 min at rt. DC: untreated fragment; LC: fragment irradiated in the absence of the rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions; rac-S: racemic Sk; rac-F: racemic Sk-α_F; Δ-F: Δ-Sk-α_F; Λ-F: Λ-Sk-α_F. Note the high level of diastereoselectivity observed for Sk-α_F. Sites showing strong delta-diastereoselectivity are represented at the bottom. The cleavage sites are italicized and in bold. The 5'-NACAA-3' sites show strong preference for the delta diastereomer over lambda. Sites showing preference (albeit weak) for the lambda diastereomer may be seen at the ends of purine/pyrimidine stretches.



5'*-T*G**T*GTGAA.. .*C**A**C*AAT.. ..*C**A**C*AA*C*ATA*C*GAG.. ..AA*T*GAG-3'

3'*-A*C**A**C*ACTT.....AA*C*AA.....*G*TGTT.....T*T*ACTC.....G*C*AAC.....-5'

Figure 4.16. Photocleavage of the 5'- and 3'- (*EcoR I**) end labeled 180-mer fragment by the enantiomers of Sk and the diastereomers of Sk-α_F to show the enhancement in site-selectivity and diastereoselectivity introduced by the attached peptide. Reaction samples contained 500, 100 nM Sk or 50, 10 nM Sk-α_F and a 1:50 bp ratio of rhodium to CT-DNA in 50 mM sodium cacodylate, pH 7.0. The samples were incubated at rt for 12 h and irradiated at rt for 5 min (313 nm). DC: untreated fragment; LC: fragment irradiated in the absence of the rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions; Δ-S: Δ-Sk; Λ-S: Λ-Sk; Δ-F: Δ-Sk-α_F; Λ-F: Λ-Sk-α_F. The numbers refer to the concentration (in nM) of the appropriate rhodium complex in the reaction sample. The solid arrows denote sites preferred by the delta diastereomer and the dashed arrow sites preferred by the lambda diastereomer. Note that while the "delta sites" are common to both Sk and Sk-α_F, the "lambda sites" are predominantly seen with Sk, and Λ-Sk-α_F shows little or no diastereoselectivity.


5'-180 mer
3'-180 mer
conc. (nM)
conc. (nM)
DC
LC
AG
CT
Δ-S 500
Δ-S 100
Δ-F 50
Δ-F 10
Λ-S 500
Λ-S 100
Λ-F 50
Λ-F 10
DC
LC
AG
CT
Δ-S 500
Δ-S 100
Δ-F 50
Δ-F 10
Λ-S 500
Λ-S 100
Λ-F 50
Λ-F 10

MPE-Fe footprinting of the diastereomers was carried out to ascertain the binding sites and corroborate the photocleavage data. MPE-Fe cleavage of the DNA fragment in the presence of Δ-Sk-α_F produces distinct footprints that correlate well with the sites of strong photocleavage (Figure 4.17 and 4.18). The footprints extend in the 3'- direction from the cleavage site and the size (7-10 bp) is larger than the 4 bp coverage estimated for Sk by modeling. This result implies that the peptide is intimately associated with DNA and not simply increasing the affinity of the skeletal complex by introducing positive charges in the vicinity of the DNA phosphate backbone. Samples containing Δ-Sk or Λ-Sk exhibit hypersensitivity to MPE-Fe cleavage at all sites at 500 nM rhodium concentration. Under higher Sk concentration (5 μM) weak footprints are observed with Δ-Sk that cover 4-5 bp. Significantly, the reaction samples containing Λ-Sk-α_F show hypersensitivity to MPE-Fe cleavage too. The fragment contains no sites that are preferentially occupied by Λ-Sk-α_F and, consequently, the photocleavage and MPE-Fe cleavage results are similar to Sk. The MPE-Fe and photocleavage data for Δ-Sk-α_F was quantitated on a phosphorimager (Figure 4.19) and is summarized as a histogram in Figure 4.20.

4.3.9. DNA Recognition by a Family of Metal-Peptide Complexes Differing by a Single Amino Acid Residue. Photocleavage by a family of metal-peptide complexes, Sk-α_F, Sk-α_G, Sk-α_H and Sk-α_I, was investigated to understand the effect of introducing a single Q/R or Q/A substitution at select positions along the peptide sequence. All four complexes recognize the 5'-NACAA-3' sequences and show enhanced cleavage at these sites over Sk (Figure 4.21). Differences in the relative photocleavage intensities at different sites for the four metal-peptide complexes are observed. In particular, cleavage by Sk-α_G and Sk-α_H at 5'-ATAAAG-3' and by Sk-α_G at 5'-CTGGGG-3' sites is enhanced compared to the other complexes. All the four metal-peptide complexes in this family show strong photocleavage at a 5'-TAATGAGT/ACTCATTA-3' site on the

Figure 4.17. Autoradiogram of an 8% denaturing polyacrylamide gel showing photocleavage and protection from MPE-Fe cleavage of the 5'-(*EcoR I**) and 3'-(*EcoR I**) end labeled 180 bp restriction fragment by the diastereomers of Sk and Sk-α_F. DC: untreated fragment; LC: fragment irradiated in the absence of the rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions; Δ-S: Δ-Sk; Λ-S: Λ-Sk; Δ-F: Δ-Sk-α_F; Λ-F: Λ-Sk-α_F. The MPE-Fe cleavage lanes have been labeled by "M." The numbers indicate concentration of the rhodium complexes in nM. A 1:10 bp rhodium to bp DNA ratio was used in the reactions. Regions of DNA protected from MPE-Fe cleavage are indicated by brackets. Note the faint "footprints" observed with Δ-Sk-α_F. The regions of protection correlate well with the sites of photocleavage. Δ-Sk, Λ-Sk and Λ-Sk-α_F show hypersensitivity to MPE-Fe cleavage.

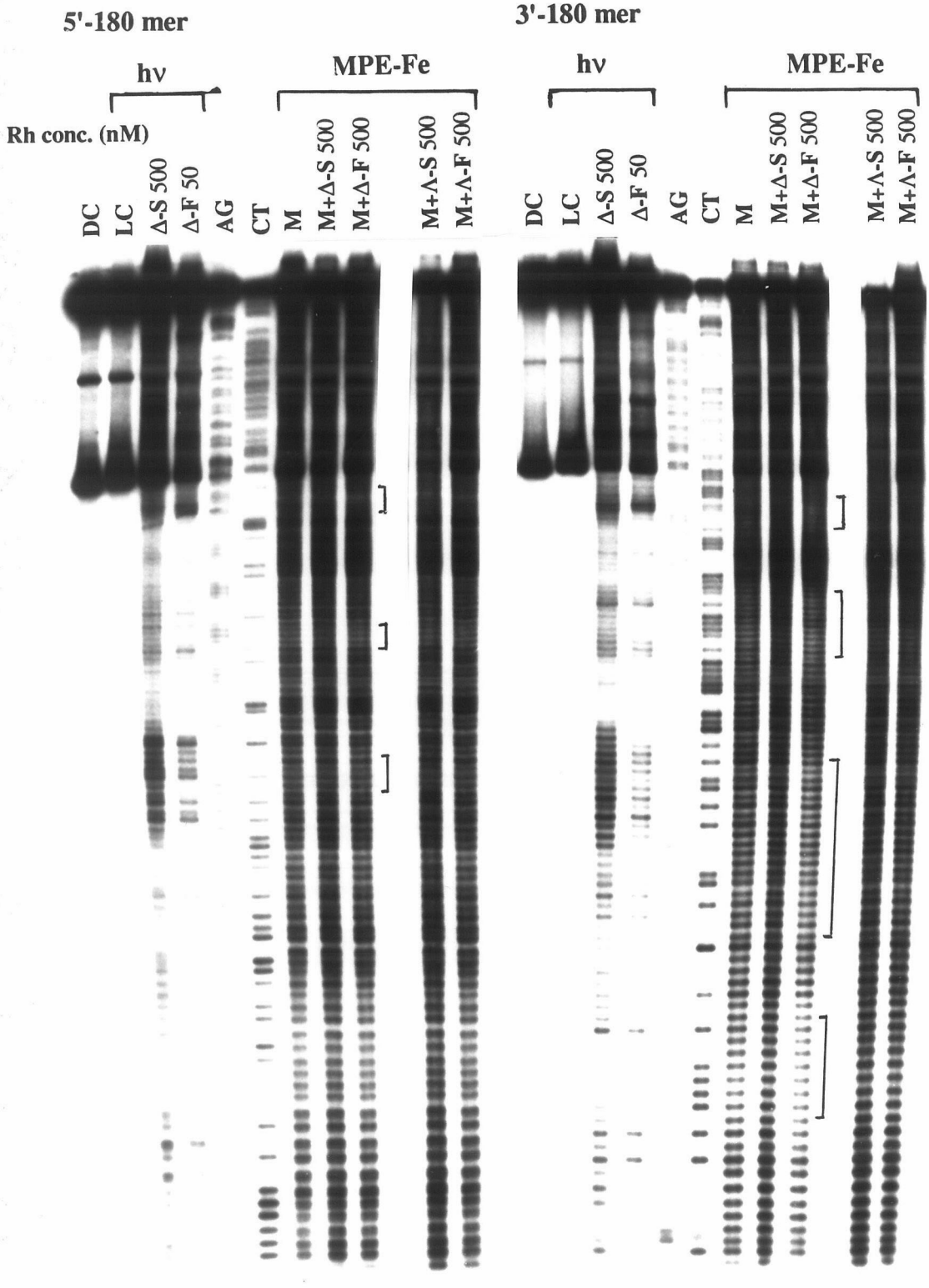
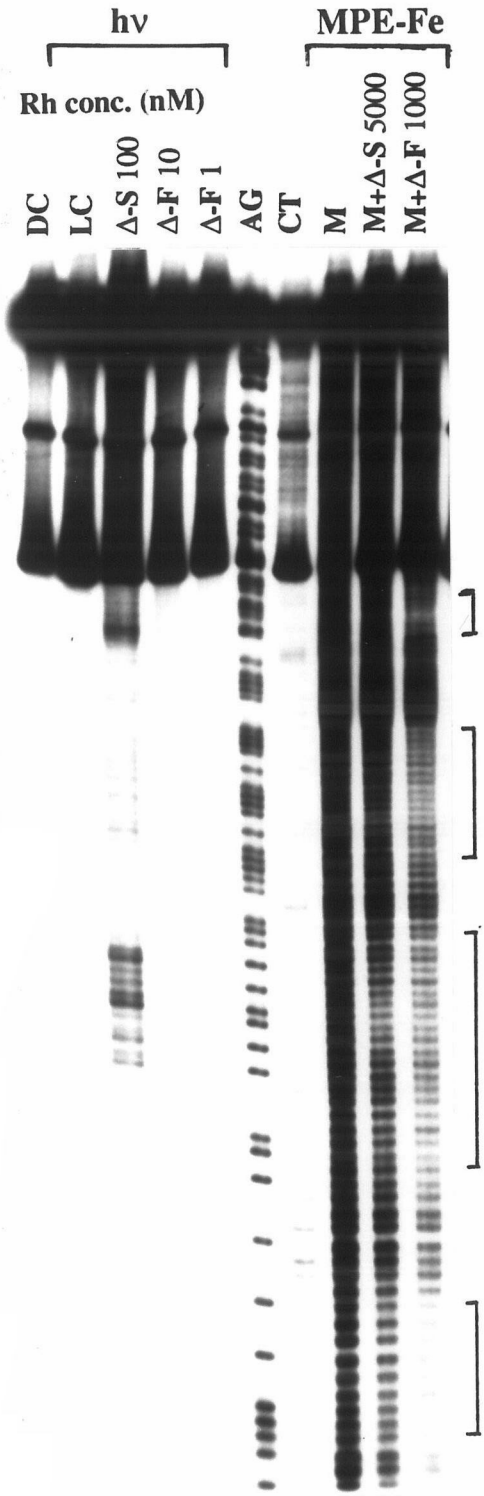

5'-180 mer
3'-180 mer
hν
MPE-Fe
Rh conc. (nM)
DC
LC
Δ-S 500
Δ-F 50
AG
CT
M
M+Δ-S 500
M+Δ-F 500
M+Λ-S 500
M+Λ-F 500

Figure 4.18. Autoradiogram of an 8% denaturing polyacrylamide gel showing photocleavage and protection from MPE-Fe cleavage of the 5'-(*EcoR I**) and 3'-(*Pvu II**) end labeled 180 bp restriction fragment by Δ-Sk and Δ-Sk-α_F. These reactions represent cleavage on the same DNA strand labeled at the two ends. DC: untreated fragment; LC: fragment irradiated in the absence of the rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions; Δ-S: Δ-Sk; Δ-F: Δ-Sk-α_F. The MPE-Fe cleavage lanes have been labeled by "M." The numbers indicate concentration of the rhodium complexes in nM. A 1:10 bp rhodium to bp DNA ratio was used in the reactions. Regions of DNA protected from MPE-Fe cleavage are indicated by brackets. Note: distinct "footprints" covering 7-10 bp are observed with Δ-Sk-α_F. Weak "footprints" covering ca. 4 bp may be observed for Δ-Sk at the high concentrations used here.

5'-180 mer (*EcoRI**)



3'-180 mer (*PvuII**)

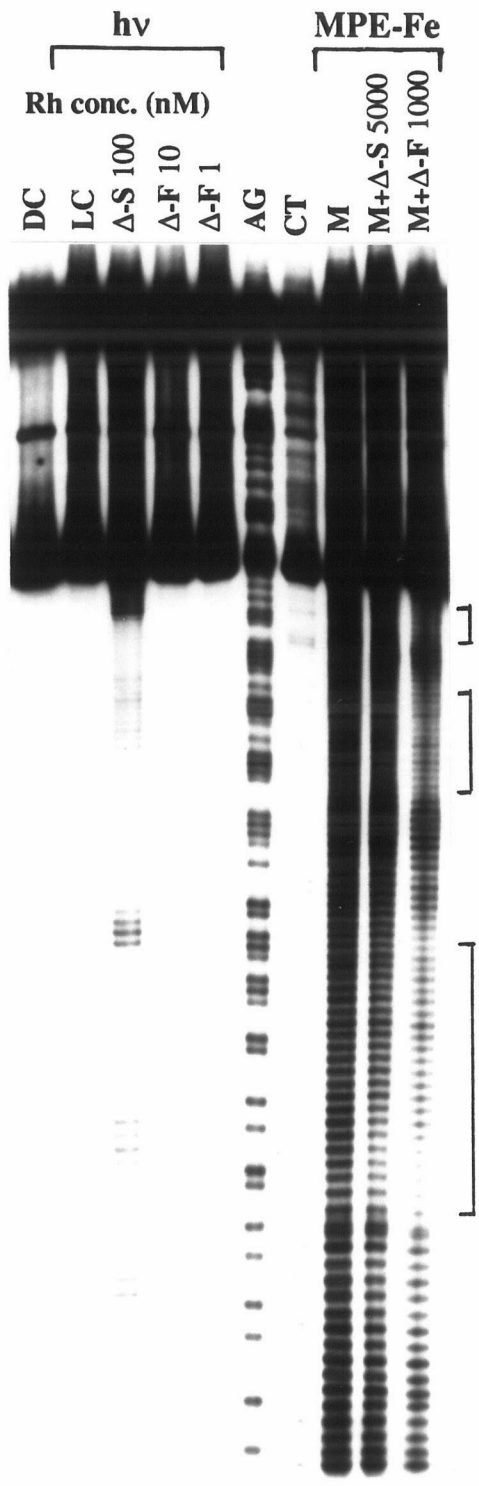


Figure 4.19. Phosphorimager quantitation data from the 5'-end labeled gels shown in Figures 4.17 and 4.18. **(a)** MPE-Fe cleavage data in the absence of rhodium complex or in the presence of Δ-Sk and Δ-Sk-α_F. Note: Only Δ-Sk-α_F shows regions of protection while Δ-Sk shows uniform hypersensitivity.

(b) Photocleavage data for Δ-Sk-α_F in the corresponding region. Note: The correlation between strong photocleavage sites and regions protected from MPE-Fe cleavage. The "footprint" appears extended in the 3' direction relative to the photocleavage site.

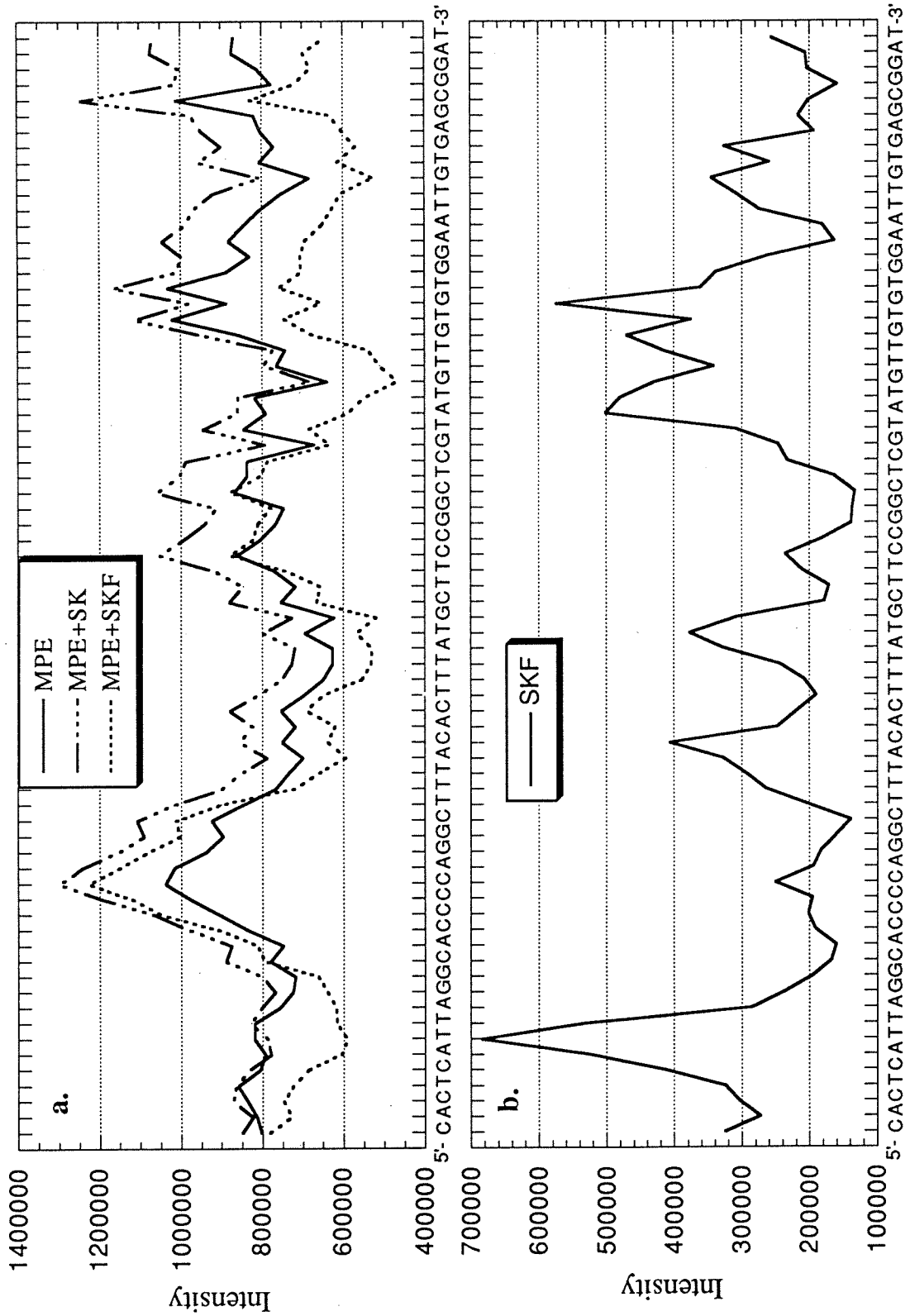
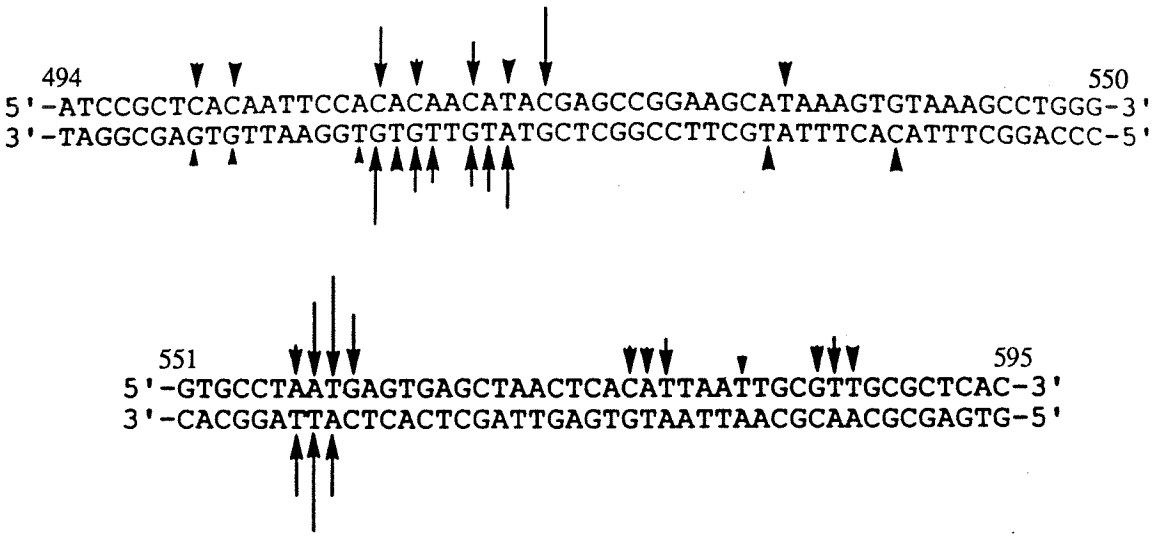

a.
MPE
MPE+SK
MPE+SKF
Intensity
1400000
1200000
1000000
800000
600000
400000
5'- CACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGAT-3'
b.
SKF
Intensity
700000
600000
500000
400000
300000
200000
100000
5'- CACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGAT-3'
Nucleotide Sequence

Figure 4.20. Histograms summarizing the photocleavage and MPE-Fe cleavage protection data for the reaction of Δ-Sk-α_F with differently labeled 180 bp (*EcoR I/Pvu II*) fragments from plasmid pUC 18. **(a)** Photocleavage data. The arrows represent cleavage sites and the lengths correlate with relative intensity at the site. **(b)** MPE-Fe cleavage data. The bars represent relative extent of protection from MPE-Fe cleavage at that site compared to cleavage in the absence of rhodium complex. The open triangles represent sites where greater MPE-Fe cleavage is seen in the presence of the rhodium complex (hypersensitive sites).

a.


494
550
5'-ATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG-3'
3'-TAGGCGAGTGTTAAGGTGTGTTGTATGCTCGGCCTTCGTATTTCACATTTCGGACCC-5'
551
595
5'-GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC-3'
3'-CACGGATTACTCACTCGATTGAGTGTAATTAACGCAACGCGAGTG-5'

b.

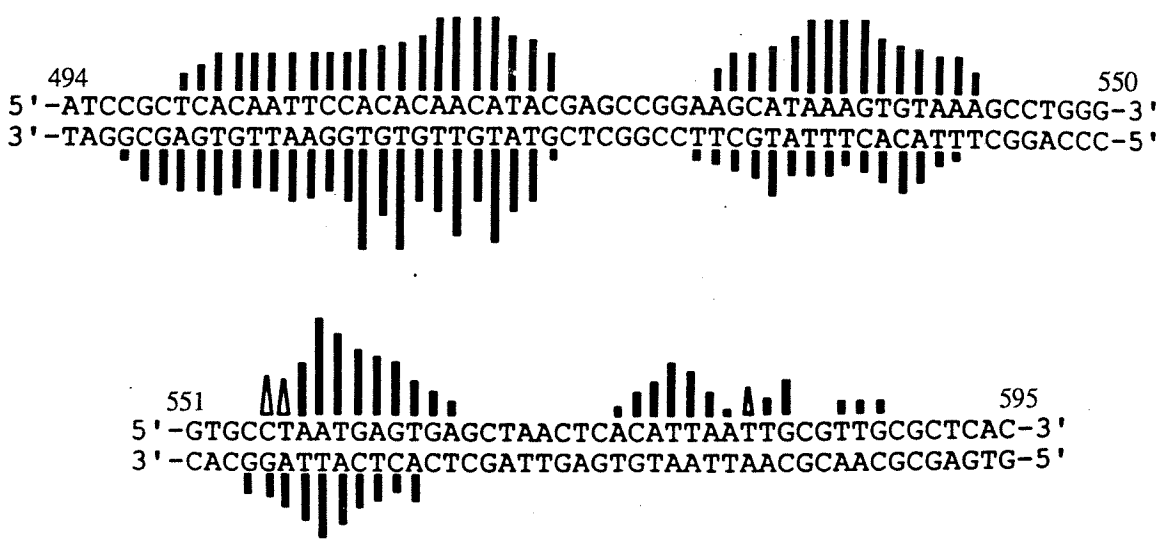
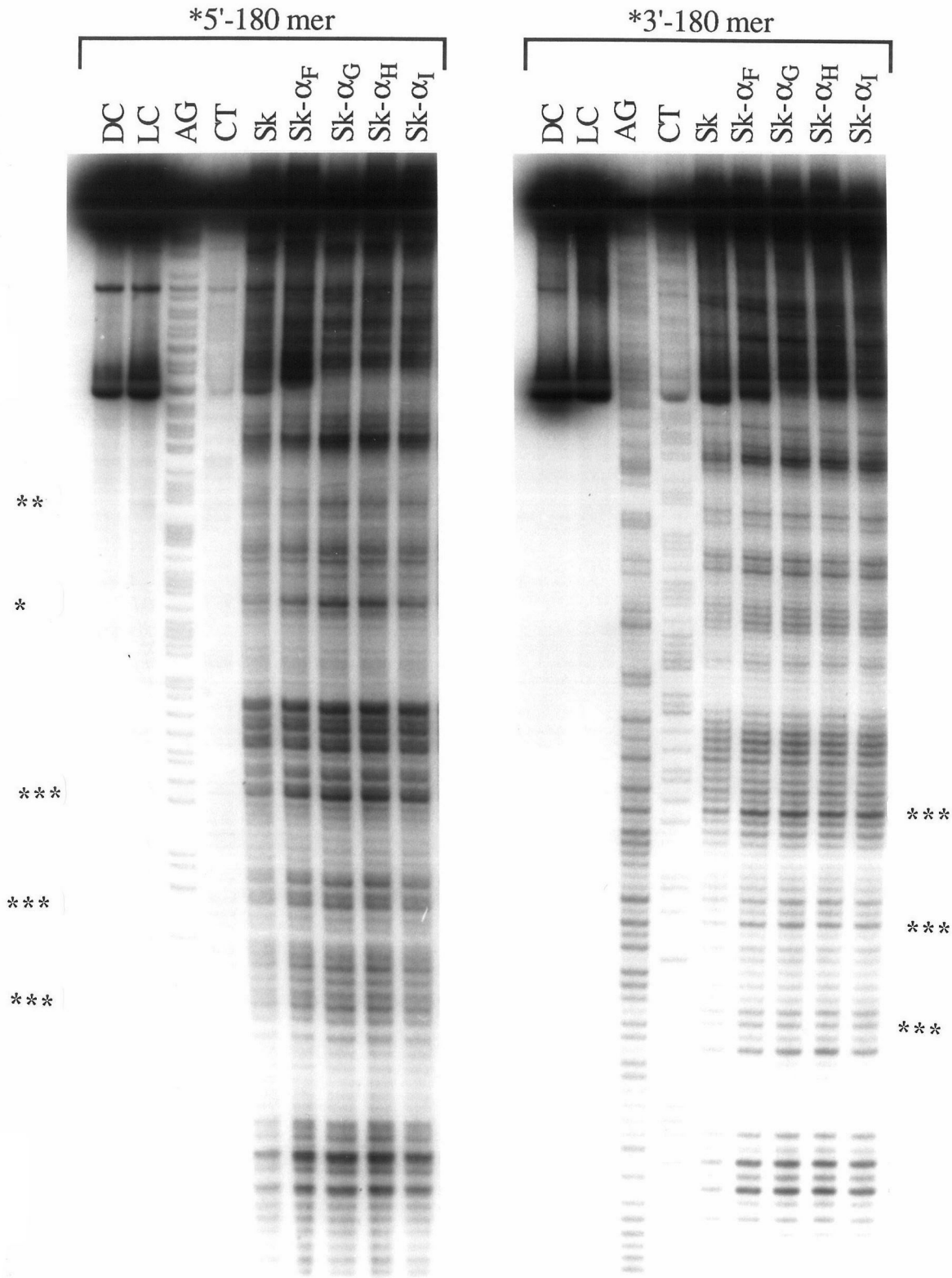
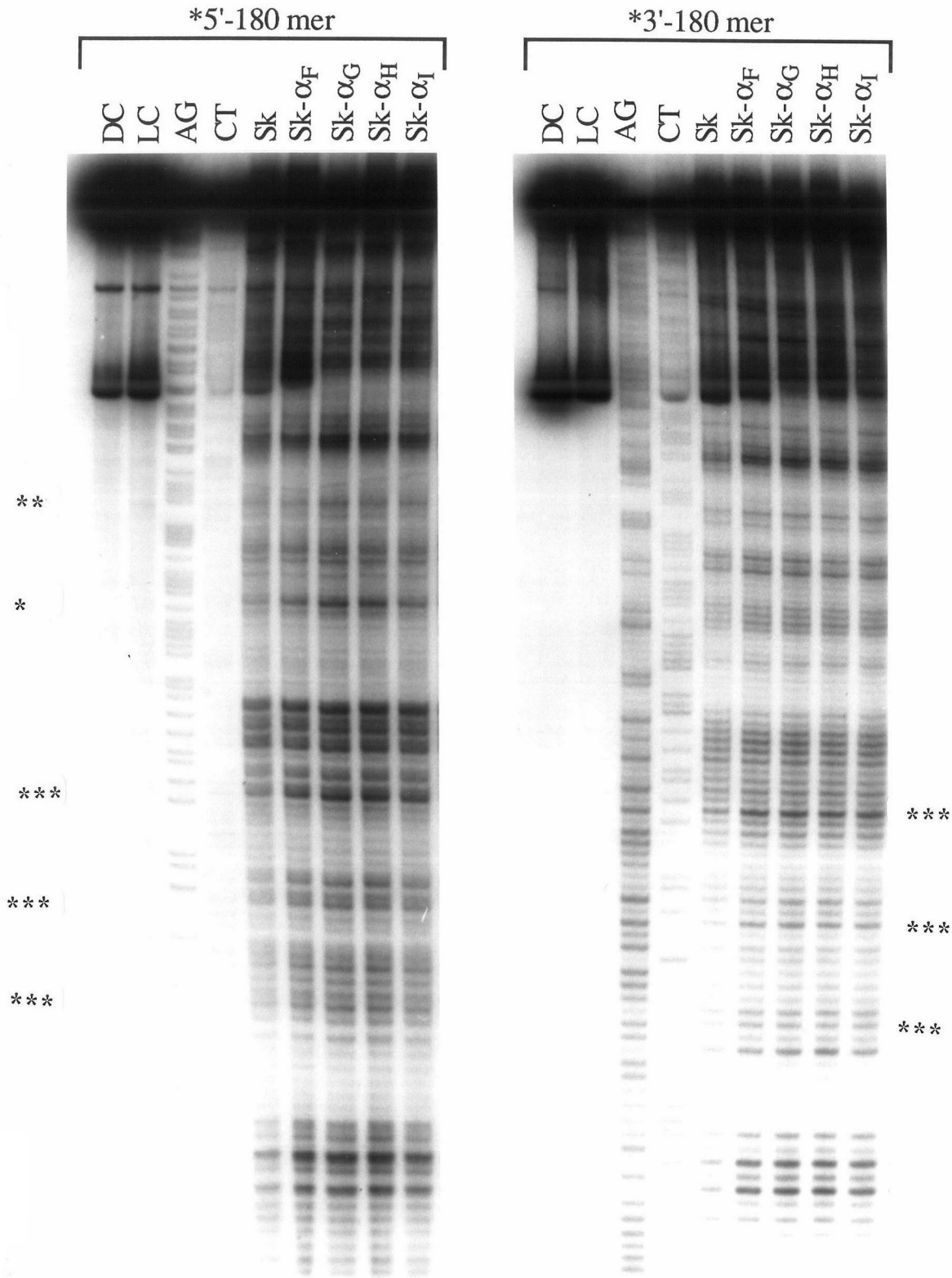

494
550
5'-ATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG-3'
3'-TAGGCGAGTGTTAAGGTGTGTTGTATGCTCGGCCTTCGTATTTCACATTTCGGACCC-5'
551
595
5'-GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC-3'
3'-CACGGATTACTCACTCGATTGAGTGTAATTAACGCAACGCGAGTG-5'

Figure 4.21. Autoradiogram of an 8% denaturing polyacrylamide gel showing photocleavage of the 5'-(*EcoR I**) and 3'-(*EcoR I**) end labeled 180 bp restriction fragment by Sk, Sk-α_F, Sk-α_G, Sk-α_H, and Sk-α_I. Reaction samples contained 1 μM rhodium complex and 50 μM CT-DNA in 50 mM sodium cacodylate, pH 7.0. The samples were incubated at rt for 5 min and irradiated at rt for 5 min (313 nm). DC: untreated fragment; LC: fragment irradiated in the absence of the rhodium complex; AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions. The 5'-NACAA/TTGTN-3' sites (***), the 5'-ATAAAG-3' site (*) and the 5'-CTGGGG-3' site (**) are indicated.


*5'-180 mer
*3'-180 mer
DC
LC
AG
CT
Sk
Sk-αF
Sk-αG
Sk-αH
Sk-αI
**
*
***
***
***
***
***
***

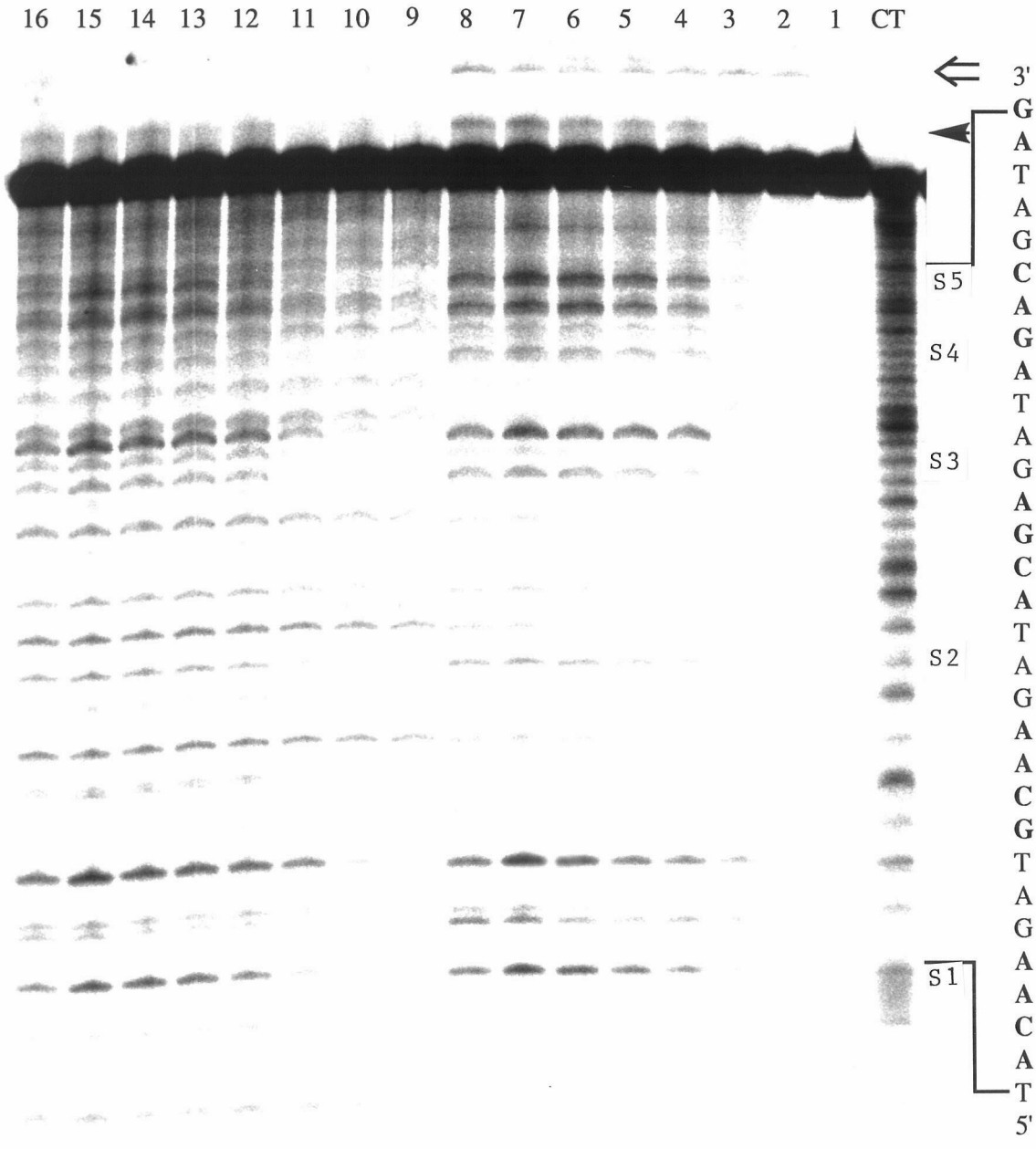

*5'-180 mer
*3'-180 mer
DC
LC
AG
CT
Sk
Sk-αF
Sk-αG
Sk-αH
Sk-αI
**
*
***
***
***
***
***
***

fragment; but the same is true for Sk too. We synthesized a 51 bp DNA oligonucleotide duplex containing six 4 bp variants of the 5'-ACAA-3' sequence in order to maximize the DNA sequence discrimination by the four metal-peptide complexes. The sites were separated by 5'-GAT-3' sequences. The six sites with a flanking 5'-T are: 5'-TACAA-3' (S1); 5'-TGCAA-3' (S2); 5'-TACGA-3' (S3); 5'-TAGAC-3' (S4); 5'-TAGAA-3' (S5); and 5'-TAGGA-3' (S6). Initial computer modeling and subsequent photocleavage data from experiments with restriction fragments predicts that the metal-peptide complexes intercalate at the 5'-TN-3' base step and cleave at the 5'-**TN/NA**-3' bases.

Photocleavage of the 5'-end labeled 51T oligonucleotide by the family is compared in Figure 4.22. The metal-peptide complexes cleave at the 5'-T at all sites, but the cleavage intensity varies. 5'-**TACAA**-3', 5'-**TACGA**-3', and 5'-**TAGAA**-3' sites are cleaved more strongly than 5'-**TGCAA**-3', 5'-**TAGGA**-3' and 5'-**TAGAC**-3' sites. While this selectivity is true for all the metal-peptide complexes, the absolute intensity of cleavage is dependent on the peptide sequence. Sk-α_G and Sk-α_H show the strongest cleavage. Sk does not cleave the sites appreciably. Comparison of the photocleavage by Sk-α_F at 4° C and 24° C (Figure 4.22 lanes 4 and 8 respectively) indicates that somewhat greater cleavage is observed at the higher temperature. The difference in temperature can affect the k_{on} for the metal-peptide complexes binding to DNA and increase cleavage in that manner.

Also shown in Figure 4.22 is the effect of treating the reaction samples with piperidine. Two sets of retarded bands are observed in the untreated lanes that migrate slower than the uncut fragment. The first set (open arrow) is from light induced damage to DNA and is seen in all lanes that have been irradiated (lanes 2-8); irrespective of the presence of rhodium complex in the reaction sample. This photoproduct is cleaved upon piperidine treatment (lanes 9-16) and migrates faster in the gel along with the other DNA products. The second DNA photoproduct (discussed in Section 4.3.6) is only observed in

Figure 4.22. Autoradiogram of a 20% denaturing polyacrylamide gel showing photocleavage of the 5'-51T* end labeled 51T/51B oligonucleotide duplex by Sk, Sk-α_F, Sk-α_G, Sk-α_H, and Sk-α_I. Reaction samples contained 0.2 μM rhodium complex and 0.2 μM oligonucleotide duplex in 25 mM sodium cacodylate, pH 7.0. The samples were incubated at 4° C (on ice) for 10 min and irradiated at 4° C (on ice) for 10 min (313 nm). CT: Maxam-Gilbert C+T sequencing reaction; lane 1: untreated oligonucleotide duplex; lane 2: oligonucleotide duplex irradiated in the absence of the rhodium complex; lanes 3-7: oligonucleotide duplex irradiated in the presence of Sk, Sk-α_F, Sk-α_G, Sk-α_H, and Sk-α_I respectively; lane 8: oligonucleotide duplex incubated and irradiated in the presence of Sk-α_F at 24° C. Lanes 9-16 are samples from lanes 1-8 respectively following piperidine treatment. The six operator site variants are labeled S1: 5'-TACAA-3'; S2: 5'-TGCAA-3'; S3: 5'-TACGA-3'; S4: 5'-TAGAC-3'; S5: 5'-TAGAA-3'; S6: 5'-TAGGA-3' and marked at the position of the 5'-T cleavage site. Note the two retarded bands (arrows) migrating slower than the uncut fragment. The band corresponding to the "open arrow" is piperidine labile while that corresponding to the solid arrow is not.


16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 CT
3'
G
A
T
A
G
C
A
G
A
T
A
G
A
G
C
A
T
A
G
A
A
C
G
T
A
G
A
A
C
A
T
5'
S5
S4
S3
S2
S1

the reaction lanes containing metal-peptide complexes. The product, absent in lanes containing Sk or no rhodium complex (lanes 1-3), is not affected by piperidine treatment.

Photocleavage of the 3'- end labeled 51T and 51B oligonucleotide duplexes is shown in Figure 4.23. The hierarchy of cleavage at the six sites is as follows:
(3'*-51T): 5'-TAGAA-3' >> 5'-TACGA-3' > 5'-TAGAC-3' > 5'-TAGGA-3' > 5'-TGCAA-3'; the 5'-TACAA-3' is not discernible as it migrates close to the uncut fragment and is difficult to resolve on the gel.
(3'*-51B): 3'-ATGCT-5' (5'-TACGA-3') > 3'-ATGTT-5' (5'-TACAA-3') > 3'-ACGTT-5' (5'-TGCAA-3') ~ 3'-ATCTT-5' (5'-TAGAA-3') > 3'-ATCTG-5' (5'-TAGAC-3'); the 3'-ATCCT-5' (5'-TAGGA-3') is not well resolved.

The overall DNA site selectivity appears to follow the trend: 5'-ACGA-3' > 5'-ACAA-3' > 5'-AGAA-3' ~ 5'-AGAC-3' > 5'-AGGA-3' > 5'-GCAA-3'. The strongest sites are the same for all four metal-peptide complexes but the intensity and rank of the intermediate/weak sites may vary with the peptide sequence. Quantitative measurements were made of the cleavage intensities at the six intercalation sites. A 4 to 6-fold enhancement in cleavage intensity is observed between the strongest and weakest of the six sites. This difference translates into a ~ 0.6-0.8 kcal/mol difference in the binding energy for the DNA sites. Sk, in contrast, shows overall reduced binding and only a 2 to 3-fold difference in cleavage amongst the six sites.

4.3.10. Gel Electrophoretic Mobility Retardation Assay for Studying Operator Binding. In parallel to the photocleavage experiments on oligonucleotide duplexes, we also investigated the affinity of the metal-peptide complexes for their DNA sites by gel retardation. The 51 bp oligonucleotide duplex was used in the experiment. As shown in Figure 4.24, the mobility of the oligonucleotide duplex is incrementally retarded by titrating in Sk-α_I. This observation of a band shift is a first for small metal complexes bound to small oligonucleotide duplexes and represents a parallel to DNA binding by

Figure 4.23. Autoradiogram of a 20% denaturing polyacrylamide gel showing photocleavage of the 3'-51T* and 3'-51B* end labeled 51T/51B oligonucleotide duplexes by Sk, Sk-α_F, Sk-α_G, Sk-α_H, and Sk-α_I. Reaction samples contained 0.1 μM rhodium complex and 0.1 μM oligonucleotide duplex in 25 mM sodium cacodylate, pH 7.0. The samples were incubated at 4° C (on ice) for 10 min and irradiated at 4° C (on ice) for 10 min (313 nm). AG, CT: Maxam-Gilbert A+G and C+T sequencing reactions; DC: untreated oligonucleotide duplex; LC: oligonucleotide duplex irradiated in the absence of the rhodium complex; lanes containing oligonucleotide duplex irradiated in the presence of Sk, Sk-α_F, Sk-α_G, Sk-α_H, and Sk-α_I, respectively, are appropriately labeled. The 3'-ends of the duplexes are at the bottom and the 5'-ends at the top. The six operator site variants are labeled S1: 5'-TACAA-3'; S2: 5'-TGCAA-3'; S3: 5'-TACGA-3'; S4: 5'-TAGAC-3'; S5: 5'-TAGAA-3'; S6: 5'-TAGGA-3' for 3'-51T* and S1'-S6', respectively, for the complementary 3'-51B* strand.

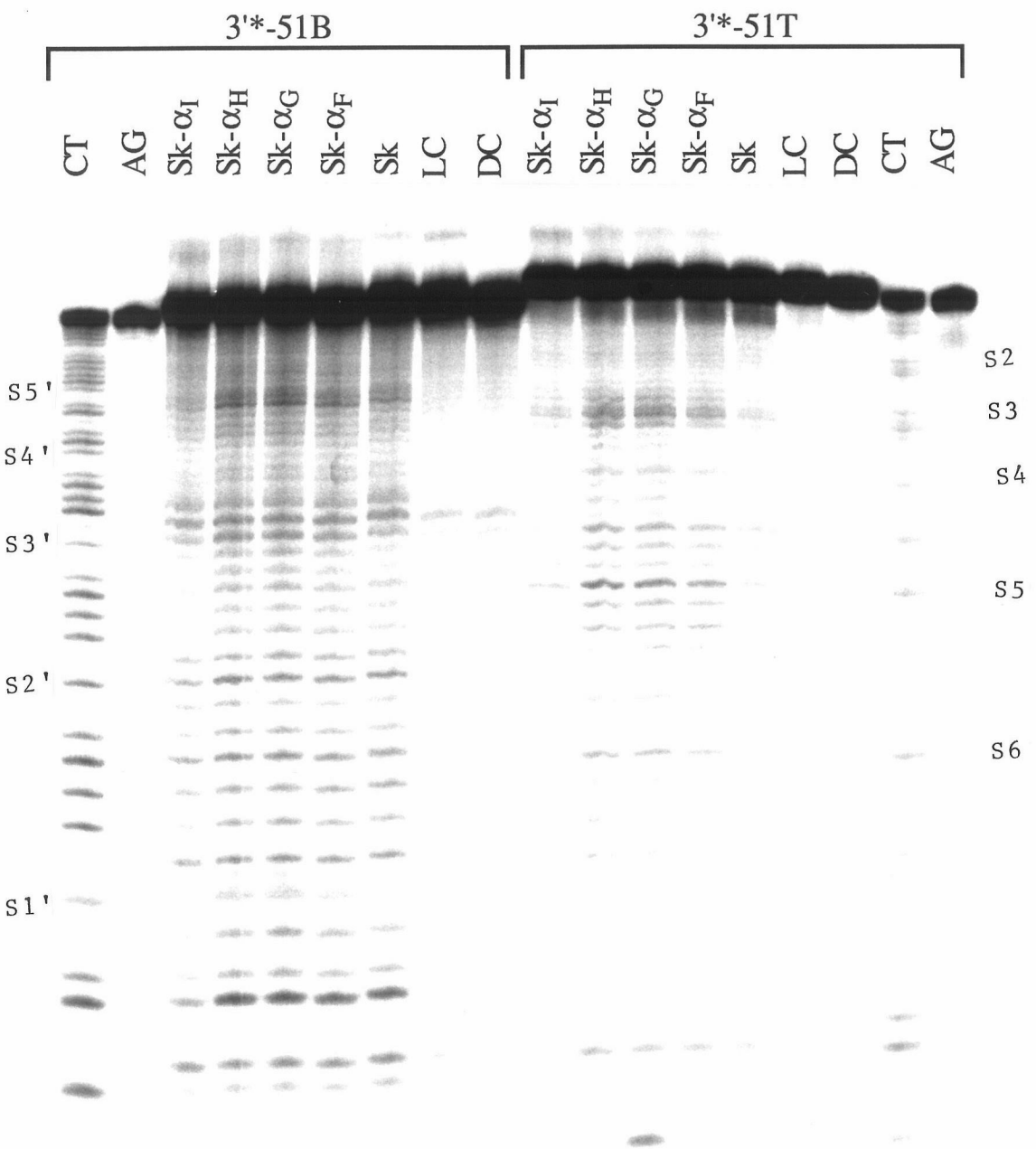
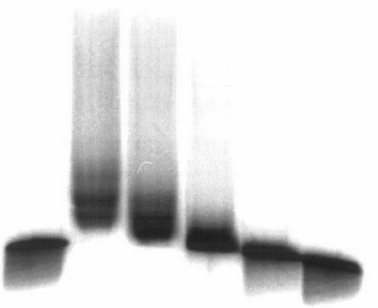
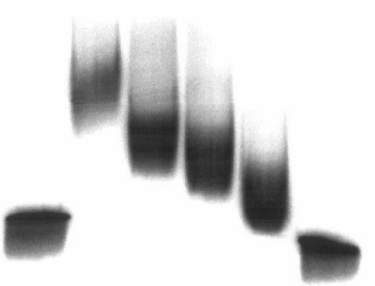

3'*-51B
3'*-51T
CT
AG
Sk-αI
Sk-αH
Sk-αG
Sk-αF
Sk
LC
DC
Sk-αI
Sk-αH
Sk-αG
Sk-αF
Sk
LC
DC
CT
AG
S5'
S4'
S3'
S2'
S1'
S2
S3
S4
S5
S6

Figure 4.24. Autoradiogram of a 5% non-denaturing polyacrylamide gel showing the band shift induced in a labeled 51 bp oligonucleotide duplex upon the addition of increasing amounts of Sk-α_1. The numbers indicate concentration of rhodium complex (in μM) added to 0.2 μM oligonucleotide duplex.

0 1.2 1.0 0.6 0.2 0 (μM Sk-α_I)



0 3.0 2.0 1.6 1.2 0 (μM Sk-α_I)



proteins. The fact that we see a band shift also attests to the relatively low k_{off} for these metal-peptide complexes. A high k_{off} might lead to the rapid dissociation of the rhodium-DNA complex upon application of an electric field and migration of the two species towards opposite poles. However, unlike with DNA binding proteins, we do not see a "free DNA" band distinct from the "bound DNA." Also, we do not see saturation binding even at a 15 rhodium:1 duplex (51 bp) ratio. Having several binding sites (at least six) makes the data difficult to fit to derive a binding constant.

4.4. Discussion

4.4.1. Peptide Conformation in the Metal-Peptide Complexes. The ability of the attached peptides to adopt a helical conformation is crucial for mimicking DNA recognition by the parent protein, 434R, and in the rational design of molecules with defined specificity for DNA. The predicted helicity of the wild-type peptide, α_3, is low based on the P_α and σ values for the individual residues comprising the peptide.[13] Not surprisingly, in the absence of extensive stabilizing interactions with the rest of the protein core, the circular dichroism of α_3 indicates that it is not significantly helical in solution. By introducing Leu and Ala residues at specific locations in the sequence, we designed peptides with increased helicity. Furthermore, the metal-peptide complex, Sk-α_A, has a greater helical propensity compared to α_A; possibly because of N-capping the peptide.[24]

The second family of Sk-α_n (n=F, G, H or I) metal-peptide complexes designed to carry an increased positive charge and contain specific mutations for altering DNA recognition also shows an increased helical propensity (ca. 40-50%) by CD measurements. The metal-peptide complexes show no ellipticity above 240 nm indicative of a diastereomeric mixture at the rhodium center. The wild type analogue of this family (Sk-α_F) was studied by CD and NMR spectroscopy and a TFE induced change in conformation is observed by both the methods. The NMR spectra of Sk-α_F

(and the absence of an induced CD) suggest that the skeletal complex and the peptide are independent in solution. The peptide does not appear to wrap around the skeletal complex in these metal-peptide complexes. This form of aggregation may be a possibility with the negatively charged peptides.

4.4.2. Photochemical Properties of the Metal-Peptide Complexes. The photochemical behaviour of the metal-peptide complexes is dependent on the net charge on the peptide. Upon photoirradiation, the metal-peptide complexes undergo differential amounts of degradation as judged by analytical HPLC experiments. In particular, the comparison between Sk-α_A and Sk-α_A-OMe is striking. The two complexes have the same sequence with a difference of two charge units on the peptide which makes Sk-α_A-OMe distinctly more stable to photoirradiation.

The loss in HPLC signal is likely due to photoanation of the phi ligand leading to the degradation of the metal complex. The remarkable differences amongst these metal-peptide complexes and Sk illustrates the peptide-sequence dependent perturbation of pK_as of the phi-imine protons. Sk has a pK_a of ~6.4. The closely related complex, $[Rh(phi)_2(bpy)]^{3+}$, which lacks the free carboxylate has a slightly higher pK_a of ~6.8-6.9. Having a net charge of -2 on the peptide possibly lowers the pK_a of Sk further so that photoanation in Sk-α_A is rapid even at pH 6.0. Neutralizing the negative charges through esterification or substitution of Glu^6 and amidation of the C-terminus possibly raises the pK_a relative to Sk-α_A and stabilizes the complex. The Sk-P_n complexes (discussed in Chapter 3) carry a neutral charge on the peptide moieties and consequently show less degradation compared to Sk-α_A. The amount of metal-peptide complexes degraded is reduced in the presence of DNA.**

** For comparison, ca. 70% Sk-PA is degraded upon irradiation for 5 min at 55° C and a 100% loss in signal is observed after 15 min of irradiation at 55° C. In contrast, ca. 56% degradation is observed in the presence of DNA.[21]

The faster rate of photoanation manifests itself in the observed DNA cleavage too in the form of an unusual photoproduct seen only in reactions containing the metal-peptide complexes. The DNA photoadduct may be formed, possibly through base coordination, by the reaction of metal-peptide complexes that have lost a phi ligand in solution. Alternatively, the adduct may be a second reaction product following the usual phi chemistry leading to strand scission seen with these and other phi complexes of rhodium.[25] In either case, this alternate mode of interacting with DNA reduces the amount of DNA cleavage products seen in gels by retarding the cleaved DNA or by lowering the effective concentration of metal-peptide complex for DNA recognition and concomitant cleavage chemistry.

4.4.3. DNA Binding by the Metal-Peptide Complexes. The Sk-α_n family of metal-peptide complexes exhibits major groove chemistry by comparison to the cleavage observed with Sk at DNA sites that are cleaved by both complexes. Denaturing polyacrylamide gels (20%) of photocleavage of 5'-end labeled oligonucleotide duplexes by the Sk-α_n complexes show formation of a 3'-phosphoglycaldehyde product, in addition to the 3'-phosphate, similar to the products characterized for cleavage by $[Rh(phi)_2(bpy)]^{3+}$ and $[Rh(phen)_2(phi)]^{3+}$.[25] The metal-peptide complexes show greater overall cleavage compared to Sk at concentrations below 0.2 μM. A portion of this greater photocleavage may be from the higher rates of photoanation, particularly true for peptides with a negative charge. However, a significant portion comes from an increased specific binding affinity. We can observe specific photocleavage at 50 or 10 nM rhodium concentration with the metal-peptide complexes, concentrations at which no significant cleavage is seen with Sk.

4.4.4. Site Specificity of the Metal-Peptide Complexes. Tethering a peptide on Sk affects the site specificity of the parent complex. Significantly, 5'-NACAA-3' sites show

the greatest discrimination for cleavage by Sk-α_n *versus* Sk. These sequences are also the 434R operator half sites. Even the wild-type peptide Sk-α_3, with a low helical propensity, shows a preference for the 434R operator sites. That this selectivity is due to the particular peptide sequence is demonstrated by the comparison with Sk-PA and Sk'-PA (Figure 4.8). These two complexes have a different peptide sequence and do not show the preference for 5'-NACAA-3' sites over Sk. Changes in the peptide sequence of Sk-α_n influence the cleavage intensity at the operator sites. Thus, for example, Sk-α_A having a greater helical propensity cleaves at the 5'-NACAA-3' sites more strongly than Sk-α_3. Similarly, Sk-α_A-OMe cleaves more strongly at the operator sites compared to Sk-α_A possibly as a result of the relatively greater charge on the peptide (Figure 4.9).

DNA recognition by the metal-peptide complexes has two components viz., recognition by the peptide and contributions from Sk. Sk has a high binding affinity for DNA relative to the free peptide and consequently makes the major contribution to binding. The background Sk-like selectivity seen in all the gels is not surprising. 5'-ACA-3' sites are preferred Sk sites making the distinction between DNA recognition by Sk or by the peptide more difficult. Furthermore, for the metal-peptide complex to bind to a particular site, the skeletal complex must intercalate at that site. Sk alone may do so too. Therefore, the absolute cleavage pattern for Sk-α_n and Sk will be identical but the relative intensities of bands will change depending on the preferences of the peptide. The fact that we do not see identical cleavage with Sk and Sk-α_n demonstrates the contributions of the peptide in recognition. Conversely, the metal-peptide complexes may bind to strong Sk sites (5'-ACA/TGT-3' or 5'-TAATGAG / ACTCATTA-3'). The peptides in the Sk-α_n family are not highly helical. The conformational flexibility may allow the peptides to remain away from the DNA in solution or interact non-specifically with the phosphate backbone. The major contribution to DNA recognition in this mode will be from the intercalated rhodium complex.

What is the source of the 5'-ACAA-3' recognition by the Sk-α_n family of metal-peptide complexes? The 5'-ACAA-3' selection cannot be from shape selection by a random coil peptide attached to a rhodium complex because sequence changes should alter the conformation and change the recognition as in Sk-PA or Sk'-PA. Moreover, the enhancement in cleavage cannot be due to an overall increase in affinity because only the 5'-ACAA-3' sites are enhanced. *We believe that the predicted operator site recognition is due to the peptide binding in the correct orientation as an α-helix to the DNA site.* The other sites, which are Sk-like, are for reasons above accessible to the metal-peptide complex in a different conformation. The effect of flexibility is exemplified in the comparison between Sk-PA and Sk'-PA: two metal-peptide complexes that share the same peptide sequence but have different carrier molecules and tethers. The complex with the greater conformational flexibility in the linker, Sk'-PA, is totally similar to the parent complex in DNA recognition. Therefore, exclusive operator recognition may be achieved by delivering the functional groups as a rigid α-helix.

4.4.5. DNA Recognition by Sk-α_F. Sk-α_F has been studied in greatest detail amongst the Sk-α_n family of metal-peptide complexes. Sk-α_F is distinct from Sk in several important ways. The complex has at least a 10-fold higher binding affinity than the parent skeletal complex, binds selectively to 5'-ACAA-3' sequences over Sk, and exhibits a 2-fold increase in diastereoselectivity for binding of the delta-diastereomer to these sites over that observed for Sk. Furthermore, Δ-Sk-α_F protects 7-10 bp of DNA from MPE-Fe cleavage around the operator sites and the region of protection correlates well with the sites of photocleavage (Figure 4.20). The footprints extend in the 3'-direction from the photocleavage site suggesting that the N-->C orientation of the appended peptide is parallel to the 5'-ACAA-3' sequence in a manner analogous to 434R. In contrast, the lambda-diastereomer of Sk-α_F and the enantiomers of Sk only show hypersensitivity to MPE-Fe cleavage under similar conditions (Figure 4.17).

Preliminary kinetic evidence (Figure 4.14) suggests that the k_{off} is only on the order of 0.2 min^{-1} or less ($t_{1/2} \sim 3.5$ min) for Sk-α_F. This slow off rate is in contrast to that observed for other phi complexes of rhodium which are in the 10-100 s^{-1} regime.[26] Perhaps, because the binding interactions are spread over a greater surface area, the dissociation rates for the metal-peptide complex from DNA are slower compared to that for Sk even though the difference in binding affinity may only be 10-fold. The high binding affinity and slower dissociation rate for Sk-α_F are strong parallels to DNA binding proteins. The slower dissociation rates may also account for the observed electrophoretic mobility shifts with Sk-α_I on oligonucleotide duplexes (Figure 4.24).

4.4.6. Operator Site Discrimination by the Sk-α_n Family of Metal-Peptide Complexes and Comparison to 434R. 434R binds to six naturally occurring dimeric operator sites and together with 434Cro controls the lysis-lysogeny decision in the life cycle of bacteriophage 434. The two proteins bind with different affinity to two sites O_R1 and O_R3. Of the 12 half sites, 11 (including both half sites of O_R1) have the 5'-ACAA-3' conserved sequence at the four outer base pairs contacted by the proteins and the remaining half site (one of the two in O_R3) has 5'-ACAG-3'. The sequence of the recognition helices of 434R and 434Cro and the proposed protein-DNA contacts with the 434R operator are shown in Figure 4.25.[27] Multiple protein-DNA contacts are used by the proteins to contact the operator. In an attempt to understand the source of the discrimination in operator affinity, Koudelka and Lam[27] compared binding of mutant 434R and 434Cro to position 3 and 4 mutants of the DNA operator. 434R is highly specific for 5'-ACAA-3' with the difference in relative affinity for binding to other 5'-ACAN-3' or 5'-ACNA-3' sequences being >150. This preference for 5'-ACAA-3' holds true with mutant repressors also: e.g., Q33L, Q33A, and multiple mutations at positions 27, 30, 32, or 33 all show a 20 to greater than 150 fold preference for 5'-ACAA-3' over other 5'-ACAN-3' or 5'-ACNA-3' sites. These mutations lower the

absolute affinity of 434R binding but retain the specificity.*** In fact three mutations, T27K/E32Q/Q33A, are required for 434R to lose its preference for 5'-ACAA-3' and bind to 5'-ACAG-3', 5'-ACAC-3' or 5'-ACAT-3' sites with equal affinity.

We observe a similar behaviour in the Sk-α_n (n=F, G, H, I) series of metal-peptide complexes. Single site mutants at Q^7 or Q^3 do not affect the relative specificity of the metal-peptide complexes but change the affinity.**** We observe a 4 to 6-fold difference in photocleavage intensity at the six operator site variants and a 2-fold difference between the metal-peptide complexes at 0.1-0.2 μM rhodium concentration. The 5'-ACAA-3' sites are cleaved strongly by the metal-peptide complexes. However, in contrast to 434R, Sk-α_n (n=F, G, H, I) show a slight 5'-ACGA-3' > 5'-ACAA-3' preference. This preference may be dependent on the sequence or may arise from an altered orientation of DNA binding. The modified 434R sequences appended to Sk-α_n (Figure 4.25 b) contain I5K, E6K and G11K mutations. K^5 and K^6 may play a role in the altered 5'-ACGA-3' recognition. Koudelka and Lam speculate that E^{32} helps in the 5'-ACAA-3' recognition in 434R by affecting the DNA configuration or the structure of the recognition helix in the protein-DNA complex. Thus E^{32} contributes indirectly to the differences in recognition between 434R and 434Cro. If so, in our E6K mutants these specific orienting interactions are absent and a different set of peptide-DNA interactions may follow.

The studies with α_3-434R described in this chapter demonstrate that we can reproduce protein-DNA recognition in a small system. These metallointercalator mimics of 434R recognize the 5'-ACAA-3' sequence and discriminate between related operator

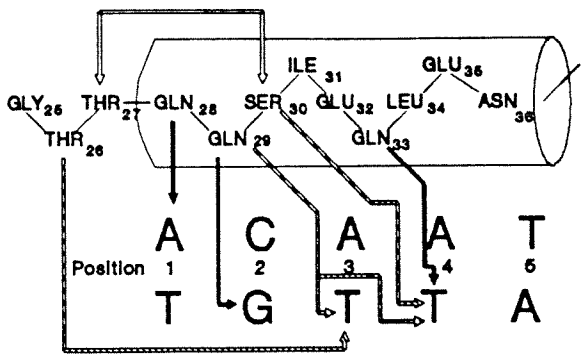
---

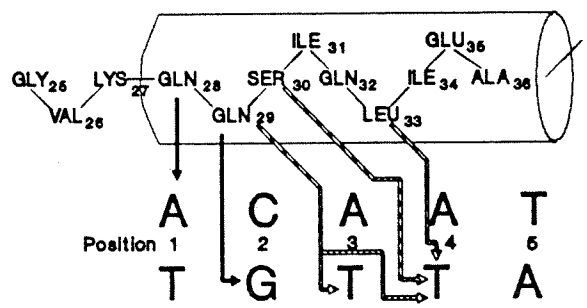
*** The binding affinity as defined by the concentration of protein monomers required to half-maximally occupy a reference operator site (5'-ACAATAT-3') for 434R is $2x10^{-8}$ M. The binding affinity of mutant 434R proteins for the reference site ranges from $6x10^{-8}$ M (Q33A) to $4x10^{-7}$ M (T27K/E32Q/Q33A triple mutant).[27]

**** The 434R numbering scheme has been followed to describe mutations in the protein while the residues in the metal-peptide complexes have been numbered from 1-14 and referred as such (Figure 4.25 b), e.g., T^{27} in 434R is T^1 in Sk-α_F.

Figure 4.25. **(a)** Proposed protein-DNA contacts between the recognition helices of 434R (top) and 434Cro (bottom) and the bases in the conserved regions of the 434R operator. Solid lines: hydrogen bonds; hatched lines: hydrophobic or van der Waals interactions; hollow line: proposed intramolecular hydrogen bond. This figure is adapted from reference 27 and is derived from crystal structure information presented in references 10 and 28. Note: The multiple protein-DNA contacts and that different residues may make different contacts to recognize the same base (Q^{33} and L^{33}). **(b)** Numbering scheme for α_3 in 434R and in the Sk-α_n family of metal-peptide complexes with their modified peptide sequences.

a.





b.

```
                27                                  37
434R-α3:....T Q Q S I E Q L E N G.....
Sk-αF:    Sk-T Q Q S K K Q L Q N K A A A-CONH2
Sk-αG:    Sk-T Q Q S K K R L Q N K A A A-CONH2
Sk-αH:    Sk-T Q R S K K Q L Q N K A A A-CONH2
Sk-αI:    Sk-T Q A S K K Q L Q N K A A A-CONH2
                1                               14
```

sites. The greatest difference between 434R and these mimics is the rigidity of the recognition helix and unique orientation with which it is delivered to the DNA. Both these factors are difficult to control in the metal-peptide complexes and perhaps therein lies the explanation for the modest operator selectivity seen.

The metal-peptide complexes bind as monomers and cleave operator sequences at 10 nM concentration. Like DNA binding proteins, the metal-peptide complexes protect DNA binding sites from MPE-Fe cleavage and retard the gel electrophoretic mobility of oligonucleotide duplexes containing binding sites. Given these observations, we believe that tethering DNA recognition peptides onto metallointercalator anchors is a viable strategy for constructing artificial repressors.

4.5. References

1. (a) Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1992**, 61, 1053. (b) Steitz, T. A. *Q. Rev. Biophys.* **1990**, 23, 205. (c) Harrison, S. C.; Aggarwal, A. K. *Annu. Rev. Biochem.* **1990**, 59, 933. (d) Johnson, P. F.; McKnight, S. L. *Annu. Rev. Biochem.* **1989**, 58, 799.

2. Helix-turn-helix motif. (a) Brennan, R. G.; Matthews, B. W. *J. Biol. Chem.* **1989**, *264*, 1903. (b) Wolberger, C.; Dong, Y.; Ptashne, M.; Harrison, S. C. *Nature* **1988**, *335*, 789. (c) Jordan, S. R,; Pabo, C. O. *Science* **1988**, *242*, 893. (d) Luisi, B. F.; Sigler, P. B. *Biochim. Biophys. Acta* **1990**, *1048*, 113. (e) Schultz, S. C.; Shields, G. C.; Steitz, T. A. *Science* **1991**, *253*, 1001. (f) Hughes, K. T.; Gaines, P. C. W.; Karlinsey, J. E.; Vinayak, R.; Simon, M. I. *EMBO J.* **1992**, *11*, 2695.

3. Zinc finger motif. (a) Berg, J. M. *Ann. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 405. (b) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, 252, 809. (c) Pavletich, N. P.; Pabo, C. O. *Science* **1993**, *261*, 1701. (d) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635. (e) Fairall, L.;

Harrison, S. D.; Travers, A. A.; Rhodes, D. *J. Mol. Biol.* **1992**, *226*, 349. (f) Vallee, B. L.; Coleman, J. E.; Auld, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, 88, 999.

4. Leucine Zipper motif. (a) Landschultz, W. H.; Johnson, P. F.; McKnight, S. L. *Science* **1988**, 240, 1759. (b) Vinson, C. R.; Sigler, P. B.; McKnight, S. L. *Science* **1989**, 246, 911. (c) O'Neil, K. T.; Hoess, R. H.; DeGrado, W. F. *Science* **1990**, 249, 774. (d) O'shea, E. K.; Rutkowski, R.; Kim, P. S. *Science* **1989**, 243, 538. (e) Oakly, M.; Dervan, P. B. *Science* **1990**, 248, 847. (f) Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223.

5. Homeodomain. (a) Gehring, W.J.; Muller, M.; Affolter, M.; Percival-Smith, A.; Billeter, M.; Qian, Y. Q.; Otting, G.; Wuthrich, K. *Trends Genet.* **1990**, 6, 323. (b) Kissinger, C. R.; Liu, B.; Martin-Blanco, E.; Kornberg, T. B.; Pabo, C. O. *Cell* **1990**, 63, 579.

6. Blackwell, T. K.; Bowerman, B.; Priess, J. R.; Weintraub, H. *Science* **1994**, *266*, 621.

7. (a) Bruist, M. F.; Horvath, S. J.; Hood, L. E.; Steitz, T. A.; Simon, M. I. *Science* **1987**, *235*, 777. (b) Sluka, J. P.; Horvath, S. J.; Bruist, M. F.; Simon, M. I.; Dervan, P. B. *Science* **1987**, *238*, 1129.

8. (a) Talanian, R. V.; McKnight, C. J.; Kim, P. S. *Science* **1990**, *249*, 769. (b) Talanian, R. V.; McKnight, C. J.; Rutkowski, R.; Kim, P. S. *Biochem.* **1992**, *31*, 6871 and ibid. **1993**, *32*, 1688. (c) Cuenoud, B.; Schepartz, A. *Science*, **1993**, *259*, 510. (d) Morii, T.; Simomura, M.; Morimoto, S.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 1150. (e) Ueno, M.; Sawada, M.; Makino, K.; Morii, T. *J. Am. Chem. Soc.* **1994**, *116*, 11137.

9. (a) Bushman, R.; Anderson, J. E.; Harrison, S. C.; Ptashne, M. *Nature* **1985**, *316*, 651. (b) Koudelka, G. B.; Harrison, S. C.; Ptashne, M. *Nature* **1987**, *326*, 886. (c) Koudelka, G. B.; Harbury, P.; Harrison, S. C.; Ptashne, M. *Proc. Natl. Acad.*

Sci. U. S. A. **1988**, *85*, 4633. (d) Wharton, R. P.; Ptashne, M. *Nature* **1987**, *326*, 888. (e) Koudelka, G. B.; Lam, C.-Y. *J. Biol. Chem.* **1993**, *268*, 23812.

10. (a) Anderson, J. E.; Ptashne, M.; Harrison, S. C. *Nature* **1985**, *316*, 596. (b) Anderson, J. E.; Ptashne, M.; Harrison, S. C. *Nature* **1987**, *326*, 846. (c) Aggarwal, A. K.; Rodgers, D. W.; Drottar, M.; Ptashne, M.; Harrison, S. C. *Science* **1988**, *242*, 899. (d) Harrison, S. C.; Anderson, J. E.; Koudelka, G. B.; Mondragón, A.; Subbiah, S.; Wharton, R. P.; Wolberger, C.; Ptashne, M. *Biophys. Chem.* **1988**, *29*, 31.
11. (a) Wharton, R. P.; Brown, E. L.; Ptashne, M. *Cell* **1984**, *38*, 361. (b) Wharton, R. P.; Ptashne, M. *Nature* **1985**, *316*, 601.
12. Caruthers, M. H.; Barone, A. D.; Beaucage, S. L.; Dodds, D. R.; Fisher, E. F.; McBride, L. J.; Matteucci, M.; Stabinsky, Z.; Tang, J-Y. *Methods Enz.* **1987**, *154*, 287.
13. (a) Fasman, G. D. ed. *Prediction of Protein Structure and the Principles of Protein Conformation*, Plenum Press, New York, **1989**. (b) Creighton, T. E. *Proteins: Structures and Molecular Principles*, W. H. Freeman, New York, **1983**.
14. Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning*; Cold Spring Harbor Laboratory, **1982**.
15. Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci.* USA **1977**, *74*, 560.
16. Hertzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23*, 3934.
17. (a) Carey, J. *Methods Enz.* **1991**, *208*, 103. (b) Brown, B. M.; Sauer, R. T. *Biochemistry* **1993**, *32*, 1354.
18. (a) Lehrman, S. R.; Tuls, J. L.; Lund, M. *Biochemistry* **1990**, *29*, 5590. (b) Hermans, J. *J. Am. Chem. Soc.* **1966**, *88*, 2418. (c) Brooks III, C. L.; Nilsson, L. *J. Am. Chem. Soc.* **1993**, *115*, 11034.
19. (a) Weiss, M. A.; Ellenberger, T.; Wobbe, C. R.; Lee, J. P.; Harrison, S. C.; Struhl, K. *Nature* **1990**, *347*, 575. (b) Weiss, M. A. *Biochemistry* **1990**, *29*, 8020.

(c) Patel, L.; Abate, C.; Curran, T. *Nature* **1990**, *347*, 572. (d) O' Neil, K. T.; Shuman, J. D.; Ampe, C.; DeGrado, W. F. *Biochemistry* **1990**, *30*, 9030. (e) Johnson, N. P.; Lindstrom, J.; Baase, W. A.; von Hippel, P. H. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 4840.

20. Wuthrich, K. *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York, **1986**.

21. Zimmerman, K. *Laboratory Notebook II*, Caltech, **1992-1993**.

22. (a) Chapter 2, this thesis. (b) Sardesai, N. Y.; Lin, S. C.; Zimmerman, K.; Barton, J. K. *Bioconjugate Chem.* **1995**, *submitted for publication.*

23. (a) Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432. (b) Campisi, D.; Morii, T.; Barton, J. K. *Biochemistry* **1994**, *33*, 4130.

24. (a) Presta, L. G.; Rose, G. D. *Science* **1988**, *240*, 1632. (b) Fairman, R.; Shoemaker, K. R.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Proteins: Structure, Function and Genetics* **1989**, *5*, 1. (c) Lyu, P. C.; Zhou, H. X.; Jelveh, N.; Wemmer, D. E.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1992**, *114*, 6560. (d) Forood, B.; Reddy, H. K.; Nambiar, K. P. *J. Am. Chem. Soc.* **1994**, *116*, 6935.

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

26. (a) Collins, J. G.; Shields, T. P.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 9840. (b) Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 10286.

27. Koudelka, G. B.; Lam, C.-Y. *J. Biol. Chem.* **1993**, *268*, 23812.

28. Mondragón, A.; Harrison, S. C. *J. Mol. Biol.* **1991**, *219*, 321.

Chapter 5:
Conclusions and Perspectives

5.1. Conclusions

The design of artificial repressors and mimics for DNA binding proteins is a challenge before biochemists today. The design involves a critical understanding of the principles of protein-DNA recognition which are becoming clearer as more and more information is available with the solution of each new co-crystal structure of a protein-DNA complex. Conversely, studying DNA recognition in small systems allows us to understand some of the principles governing DNA-site specific interactions in proteins. This study describes our approach for constructing small molecular analogs of DNA binding proteins which may be developed into artificial repressors.

The two families of metal-peptide complexes studied in Chapters 3 and 4, Sk-P_n and Sk-α_n respectively, illustrate several general principles:

(a) *Inducible Peptide Dependent Specificity:* First and foremost, peptides may be appended onto relatively sequence neutral metallointercalators to change the DNA specificity of the parent metal complex with the specificity dependent on the attached peptide. This property is important from the standpoint of a general strategy to target varying DNA sequences. Thus rhodium complexes may be utilized as high affinity (K_a>10^6 M^{-1}) carriers for peptide side chain functionalities to the DNA major groove. Peptide attachment appears to increase the specific binding affinity of the metallointercalator at least 10-fold.

(b) *Conformational Rigidity Versus Flexibility in DNA Recognition:* Conformational flexibility may be desirable for screening a large number of DNA sequences and defining a hierarchy of binding sites. A flexible molecule may bind to several DNA sites by invoking a different set of interactions. However, the degree of site specificity is linked

to the conformational rigidity of the metal-peptide complex. The rigidity may be reflected in the form of α-helical content of the tethered peptide. For example, in the Sk-P_n series peptides with a higher helical content and containing the essential E^{10} show greater specificity for the 5'-CCA-3' sites. Similarly, one of the factors for the relatively modest 5'-ACAA-3' recognition observed for Sk-α_n is the low helical content of these peptides. Another factor is the rigidity of the linker connecting the peptide to the parent metal complex. Sk-PA {$[Rh(phi)_2(phen')]^{3+}$-AANVAISQWERAA-CONH2} and Sk'-PA {$[Rh(phi)_2(bpy')]^{3+}$-AANVAISQWERAA-CONH2} contain the same peptide sequence but different carrier molecules. The more conformationally restricted di-amide linker based on a phen ligand in Sk-PA helps define the 5'-CCA-3' selectivity completely absent in Sk'-PA which contains a carboxybutyl linker on a bpy ligand. The problem of helicity may be overcome by the *de novo* design of a predominantly helical peptide scaffold. Sk-PJ {$[Rh(phi)_2(phen')]^{3+}$-AANVAIAAWARAA-CONH2} is one such construct that may be explored further. Similarly, the linkers may be made more rigid by introducing olefinic, acetylenic or phenyl intervening groups.

(c) <u>*α-Helices in DNA Recognition*</u>*:* The α-helix is a frequently used structural element for making major groove base specific contacts and is found in most of the well-characterized families of DNA-binding proteins. Pabo and Sauer[1] conclude, "There is no evidence that an isolated helix from any of the known motifs can bind DNA in a site-specific fashion, and no regulatory system has been discovered that uses an isolated helical peptide (i.e., a single α-helix) as a site-specific DNA-binding protein. In every reported complex involving α-helical motifs, it appears that the overall binding specificity results from a set of interactions, with some contacts from a *recognition* helix and some from surrounding regions of the protein." While the above is true because isolated helices lack the requisite binding affinity to bind specifically, this study shows that a protein scaffolding need not be a requisite for site-specific binding *in vitro* and that

other non-protein carriers, such as $[Rh(phi)_2(phen')]^{3+}$, may be used for providing the overall binding affinity and specificity.

(d) *Helix Orientation in the Major Groove:* Helix orientation is critical in site-specific recognition of DNA. While the diameter of an α-helix is correct to allow it to fit snugly into the major groove of DNA, it may do so in several orientations. The orientations of the α-helices vary significantly in different DNA-binding families. Even within the same family, the orientations may differ as with the λ-repressor and the homeodomain or the Trp repressor. The helices may lie in the major groove and run along the groove or they may be inclined at various angles. Alternately, as in the zinc-finger domains, only the N-terminal portion of the helix may lie in the major groove. Helix orientation is certainly the most difficult aspect to design in small peptide systems. The metal-peptide complexes studied here are no different. In the absence of constraints imposed by the rest of the protein, the Sk-P_n family recognizes a different DNA site 5'-CCA-3' from the parent P22R and recognition of 5'-ACAA-3' sites by the Sk-α_n family of metal-peptide complexes is modest. Nonetheless, native repressor-like recognition is observed which may be improved by building in orientational constraints such as through a rigid linker.

5.2. Future Considerations for Design of Artificial Repressors

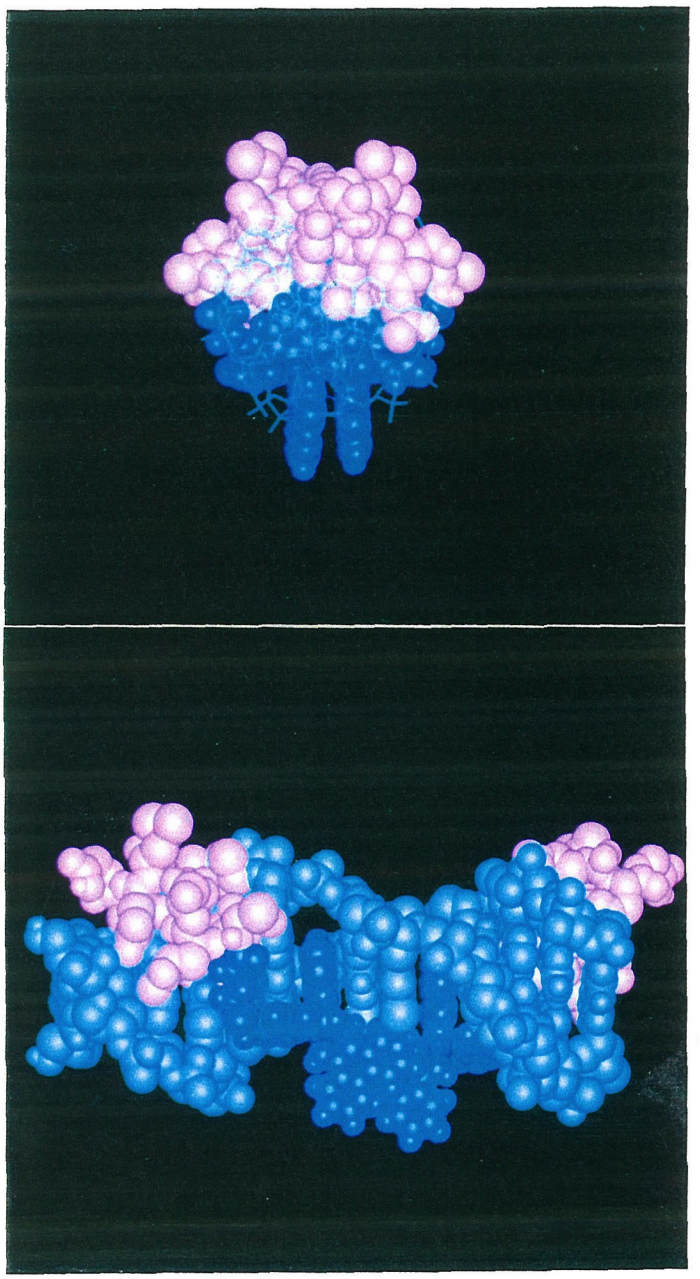
For a repressor to function effectively, it should possess high affinity and high specificity. The metal-peptide complexes bind with $> 10^7\ M^{-1}$ affinity which may be increased as the DNA recognition site is enlarged. Most DNA-binding proteins have 3-7 bp DNA recognition sequences per site (half site for dimers). Since they bind as dimers or multimers, the recognition sequence is larger. The metal-peptide complexes studied here show 3-4 bp recognition. This recognition site size may be increased by using a carrier that binds DNA with sequence specificity. Therefore, the recognition properties

of the carrier may be coupled to that of the peptide. The Barton laboratory has several phi complexes of rhodium with different 4 bp specificity that may serve as carriers upon suitably modifying one of the ancillary ligands, e.g., $[Rh(phen)_2(phi)]^{3+}$ for 5'-py-py-pu-3' sites such as 5'-CCAG-3';[2] or Δ–α-$[Rh[R,R)\text{-}Me_2trien]phi]^{3+}$ for 5'-TGCA-3' sites and other such amino acid derived aliphatic amine complexes.[3] These mono-phi complexes have the added advantage of C_2 symmetric binding about the Rh-phi axis. The carrier, $[Rh(phi)_2(phen)]^{3+}$, does not intercalate into DNA with C_2 symmetry and so the two enantiomers can each adopt two binding orientations. The alternate binding orientations available may affect the observed binding specificity in the metal-peptide complexes.

Use of longer peptides may in principle lead to recognition of larger DNA sequences. However, the length of a 13 residue α-helix is ~19.5 Å and spans the length of a major groove. Any longer helix will have to bend around the major groove to retain contact with the bases. Therefore, an α-helix can only contact a few adjacent base pairs unless a substantial kink is introduced in the peptide sequence. The glutamate at position 10 in Sk-P_n may serve as a second C-terminal contact to pin the metal-peptide complex in the major groove; the first being the intercalated metal complex at the N-terminus. This second contact may prove useful in extending the protein-DNA contact surface if a kink or a smooth bend were to be designed in the sequence. Alternately, the specificity of this "pinned" metal-peptide complex for DNA may be tuned by introducing residues such as Arg or His at specific positions from the glutamate for recognition of guanosine bases.

A third approach to increasing the DNA recognition sequence is by exploring non-covalent modes of dimerization on the DNA helix. The rhodium complex Δ-$[Rh(DPB)_2(phi)]^{3+}$ (DPB=4,4'-diphenyl-2,2'-bipyridyl) binds to an 8 bp palindromic site, 5'-CTCTAGAG-3', as a non-covalent dimer by aromatic-aromatic interactions between a phenyl group on one monomer and the bpy on the other.[4] This complex may be developed into a carrier for peptides by tethering the peptides from one of the phenyl

Figure 5.1. Computer model for "phi-mediated" dimerization of metal-peptide complexes on DNA. The pictures show the Δ-diastereomers of two metal-peptide complexes intercalated at adjacent sites separated by two DNA base pairs. The ancillary phi-ligands can stack by overlapping two of three ring systems on each other separated by 3.2 Å. The metal-peptide complexes have been canted towards opposite strands, consistent with photocleavage studies with Sk-P_n. (Left) front view; (right) Top view along the DNA principal axis. The DNA duplex is shown in aqua; rhodium complex, in blue; and peptide, in red. It should be pointed out that the picture would look similar if other modes of dimerization were to be used. Note: As non-covalent dimers, the metal-peptide complexes bear a strong resemblance to the b-zip proteins in their interaction with DNA.



substituents which can serve as a rigid linker. Interestingly, as shown in Figure 5.1, model building suggests that $[Rh(phi)_2]^{3+}$ complexes also have the potential to dimerize on DNA through stacking of the ancillary phi ligands. The model lacks experimental evidence as yet but nonetheless illustrates the potential of harnessing the combined recognition abilities of two metal-peptide complexes in a cooperative manner reminiscent of natural repressors.

Overall, this study illustrates that metallointercalators may be used as peptide delivery systems to DNA for site-specific recognition. The peptide helps lower the k_{off} of the complexes, a factor of importance in the design of therapeutics and competitors to DNA binding proteins. Finally, it should be emphasized that the metal-peptide complexes are very small compared to the DNA-binding domains of proteins (< 10-50 fold smaller) and yet bind to DNA as monomers with high affinity. These features make them attractive for developing into artificial repressors.

5.3. References

1. Pabo, C. O.; Sauer, R. T. *Ann. Rev. Biochem.* **1992**, *61*, 1053.

2. (a) Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520.
   (b) Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432.

3. Krotz, A.; Hudson, B. P.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, *115*, 12577.

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