Design and Chemical Synthesis of Sequence Specific DNA Cleaving Metalloproteins

Thesis by

David Phillip Mack

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1991

(Submitted August 14, 1990)

To Michelle

and

My Parents

Acknowledgements

I would like to thank my advisor Peter Dervan for his enthusiasm, understanding, and friendship during my stay at Caltech. I appreciated the freedom to explore my own ideas and the guidance to do thorough science.

I would also like to thank the members of the Dervan group, past and present, for providing a friendly and stimulating environment in which to do research. I would especially like to thank Brent Iverson, Jim Sluka, Kevin Leubke, and Warren Wade for their helpful scientific contributions and personal friendship.

Thanks also go to Stephen Kent, Mel Simon, and their research groups at Caltech without whom this work would not be possible.

Special thanks also go to Jackie Barton for her scientific help and insight as well as her personal friendship.

Finally, thanks to my parents for their encouragement and support, and to Michelle for her love, understanding, and companionship without which this work would have been impossible.

Abstracts

Part I: Affinity Cleaving Studies of the DNA Binding Domain of Hin Recombinase.

Previous studies have shown that attachment of ethylenediaminetetraacetic acid (EDTA) to the NH_2 terminus of the DNA binding domain of Hin recombinase creates an affinity cleaving molecule. The DNA cleavage pattern produced by this molecule has allowed a model for the DNA binding domain to be put forward which includes the NH_2 terminus binding in the minor groove and a helix-turn-helix structure binding in the major groove. In this work the affinity cleaving methodology has been extended by placing the EDTA on a lysine side chain. Differential protection of the α– and ε- amino groups of lysine has allowed EDTA to be attached at the COOH terminus and at an internal position of the protein. Using these techniques EDTA has been attached at the COOH and NH_2 terminus of the recognition helix of the DNA binding domain of Hin recombinase. Affinity cleaving studies with these molecules allows the orientation of the recognition helix on the DNA to be determined, as well as refine our model of the protein-DNA complex. The techniques developed are general and offer a powerful tool for examining the nature of protein-DNA complexes in the absence of crystallographic or NMR data.

Part II: Design of a DNA Cleaving Protein Consisting of Only Natural Amino Acids.

Attachment of the metal binding tripeptide, Gly-Gly-His, to the amino terminus of the DNA binding domain of Hin recombinase (residues 139-190) creates a new 55 residue protein containing only naturally occurring α-amino acids, GGH(Hin 139-190), with two structural domains each with distinct functions, sequence-specific recognition and cleavage of double helical DNA. This protein has been shown by footprinting to be competent to bind at four Hin sites, each 13 base pairs in length. In the presence of Cu(II), hydrogen peroxide and sodium ascorbate strong cleavage of the DNA occurs at one of the four sites by oxidative degradation of the deoxyribose backbone. Further, in the presence of Ni(II) and an oxygen atom source (eg. monoperoxyphthalic acid), the sequence specificity and efficiency of the DNA cleavage are remarkably altered. The nickel-mediated cleavage occurs at all four binding sites, is more rapid and efficient, and requires only one equivalent of an oxygen atom donor. At the hixL site, cleavage occurs predominantly at a *single* deoxyribose position on *one strand* of each binding site. Studies show that modification of the ligand system, Gly-Gly-His, can have tremendous effects on the intensity and position of DNA cleavage, and further, that for DNA cleavage mediated by the Ni(II)-protein a free amino terminus is required. Mechanistic studies indicate that the DNA cleavage by Ni(II)•GGH(Hin 139-190) likely results from abstraction of the C-4' hydrogen atom from the deoxyribose backbone by a high valent nickel-oxo species to produce a base labile modification. The tripeptide Gly-Gly-His thus is a metal specific structural domain with the function of substrate directed oxidation.

Table of Contents

List of Figures and Tables

Part I

Part II

Appendix A

Appendix B

Part I:

Affinity Cleaving Studies of the DNA Binding Domain of Hin Recombinase

Introduction

DNA recognition is an essential biological process responsible for many cellular functions. Many different types of molecules interact with DNA ranging from small organic antibiotics to large proteins. The role of protein-DNA interactions is critical in many DNA manipulations including packaging in nucleosomes, genomic rearrangements, transcription, and replication. The sequence specificity of DNA binding proteins can furthermore range from none to more than 12 specific base pairs. Of particular interest from either a control or design standpoint, is how proteins recognize DNA sequence specifically. The studies described here are generally directed toward an understanding of protein-DNA interactions. Questions which are specifically addressed include the ability to determine the location of a given amino acid or secondary structural element of a protein on the DNA using solution techniques and the ability to design a sequence specific DNA cleaving protein consisting wholly of naturally occurring amino acids.

The most common conformation of DNA is that of B-form DNA, with its well known double helical nature and approximately 10 base pairs per turn of helix. The general features of duplex DNA allow one to predict the types of interactions which may play a part in its recognition.[1] Nonspecific interactions could include ionic interactions between positively charged amino acid side chains and the DNA phosphodiester backbone as well as hydrogen bonds between the peptide backbone or side chain residues and the phosphodiester backbone. Specific interactions could include interactions between the

peptide backbone or amino acid side chains with the edges of the DNA base pairs,[2-4] hydrophobic contacts between amino acid side chains and sugars or bases, and stacking interactions of aromatic side chains with the bases. Any one or combination of these interactions could be important for the overall affinity of a protein for DNA or for the sequence specificity of the protein.

The most direct way to ascertain how proteins bind to DNA is by crystallographic techniques. The structures of a number of non-sequence specific DNA binding proteins have been determined by x-ray crystallography. These include the single stranded DNA binding protein product of Gene 5 from bacteriophage fd,[5,6] the eukaryote nucleosome particle cocrystal with DNA,[7-9] the procaryote histone-like protein HU from *Bacillus Stearothermophilus*,[10] DNase I,[11] and the Klenow fragment of DNA polymerase I from *E. Coli*.[12,13] Although some of these structures have been solved to high resolution no generalities have been discerned for the ability to bind to DNA.

Helix-Turn-Helix Binding Motif. In the early 1980's a number of x-ray crystal structures of sequence specific DNA binding proteins were solved. The first three structures solved were the DNA regulatory proteins λ Cro[14-16] and λ repressor (λ cI)[17,18] from λ phage, and the catabolite gene activator protein (CAP) (Figure 1).[19-21] Each of the three structures was observed to have a number of common features. All were dimeric and C_2 symmetric structures with each monomer containing an α-helix-turn-α-helix which was superimposable on the other two structures. The second helix of the helix-turn-helix structure was on the solvent exposed face of the protein and the distance between the two symmetry related helices in the dimers was approx. 34 Å. Furthermore the exposed helices in λ Cro and λ cI are tilted approx. 30° from the long dimension of the dimer. The exposed helices in the CAP structure, however, are tilted in the opposite direction.

The separation and tilt of the exposed helices closely resemble the separation and slope of two adjacent major grooves in B-form DNA. This similarity, along with data obtained from chemical protection data, led to a binding model wherein the exposed helices

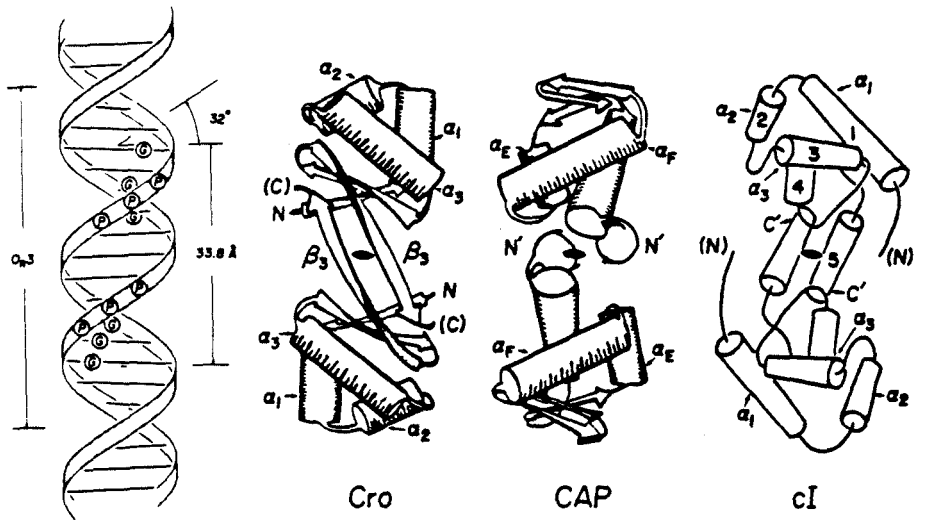


Figure 1 Schematic drawings of B-form DNA and the structures of λ Cro, CAP and λ repressor (cI). The helix turn helix domains are α_2-α_3 for λ Cro and λ repressor, and α_E-α_F for CAP.[22] The view of the proteins is of the face which interacts with the DNA. In the stylized DNA drawing, phosphates which when ethylated hinder Cro binding are marked, as well as G's which are protected from dimethyl sulfate by Cro (major groove).

of the dimers bind in adjacent major grooves of B-form DNA.[22-24] The side chains of the helix would make sequence specific contacts with the base pairs on the floor of the major groove imparting sequence specificity to the protein. The first helix of the helix-turn-helix would then lie across the major groove making nonspecific contacts with the phosphodiester backbone of the DNA.

The conservation of the helix-turn-helix structure between these three proteins led to searches of other protein sequences for conserved amino acid residues which make up this motif. It was found that homologous sequences were very common in DNA regulatory proteins but also were found in other proteins.[24] Figure 2 shows an alignment of the putative helix-turn-helix sequences of a number of sequence specific DNA binding proteins. A number of these examples have recently been shown to contain the helix-turn-helix motif and will be discussed later.

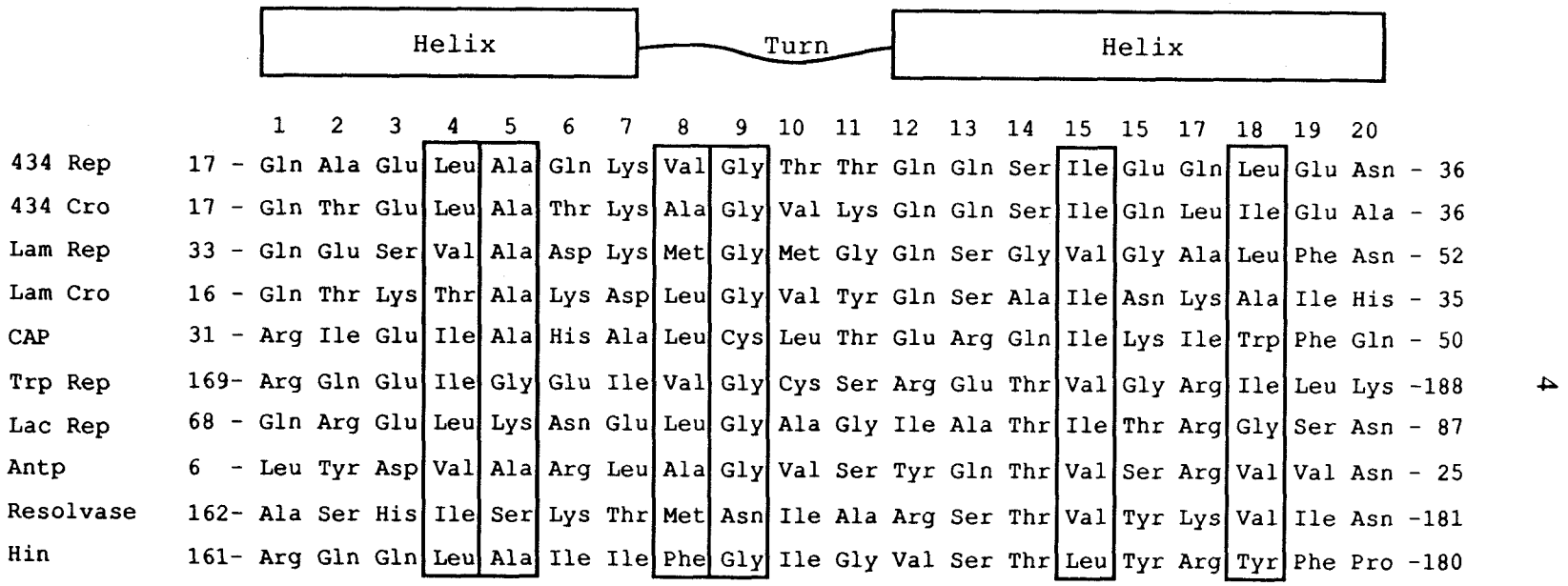


Figure 2: Helix-turn-helix alignment of some sequence specific DNA binding proteins. The sequences are shown N to C. The numbers at the beginning and end of each line are the residue numbers in the full protein sequence. Characteristic residues of the helix-turn-helix motif are boxed. For references, see sequence compilations in reference 24.

EcoRI-DNA Cocrystal. The helix-turn-helix structure, however, is not the only DNA binding motif to be proposed in recent years. The only other type of protein-DNA structure to be solved by x-ray crystallographic techniques is the DNA cocrystal of the EcoRI endonuclease apoprotein (the protein is missing a magnesium ion necessary for DNA strand scission).[25,26] The EcoRI interactions with the DNA are very different from those proposed for the helix-turn-helix motif. The two most striking features of the complex are the severe distortions of the DNA backbone and the tight packing of a large amount of protein into a small region of the DNA (Figure 3). Sequence specific DNA recognition is thought to be mediated by the amino terminal ends of four α helices which are wedged into the major groove. This four-helix parallel bundle may constitute a DNA binding motif but more structures will be necessary for such a classification.

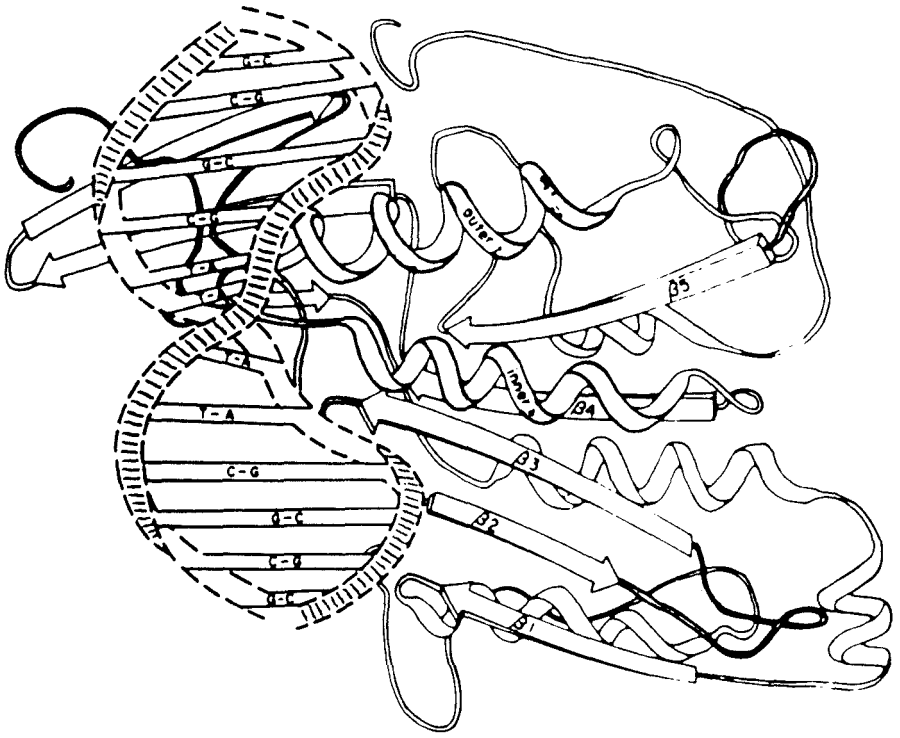


Figure 3 Schematic drawing of one subunit from the dimeric EcoRI-DNA complex.[25] The other subunit in the complex would be rotated 180 degrees about a horizontal line through the center of the first subunit in the plane of the page. The DNA oligomer sequence is 5'-TCTCGAATTCGCG-3'.

Zinc-Binding Fingers. Another DNA binding structure which is clearly different from the helix-turn-helix motif is that of the zinc-fingers. These domains are structurally organized around zinc ions coordinated to invariant cysteine and/or histidine residues.[27,28] This motif appears to be much more prominent than the EcoRI four-helix bundle and two major classes have been identified. The first class is the transcription factor IIIA-type proteins. These contain between 1 and 37 repeats which match an approx. 27 amino acid sequence which contains two cysteines and two histidines at invariant positions.[27] The second class is the steroid receptor superfamily. This class contains a single domain with two fingers each containing four cysteine residues.[29,30] The best structurally characterized class is that of the transcription factor IIIA-type proteins. NMR studies of single zinc fingers[31,32] along with predictions from amino acid sequences[33-35] have indicated a common structural motif which consists of an α helix and two strands of antiparallel β sheet (Figure 4A). A model for DNA binding suggests the α helix lies in the major groove making specific contacts in analogy to the helix-turn-helix proteins has been (Figure 4B). A more detailed structural analysis of both classes of zinc fingers will be necessary in order to determine the exact nature of the DNA interactions.

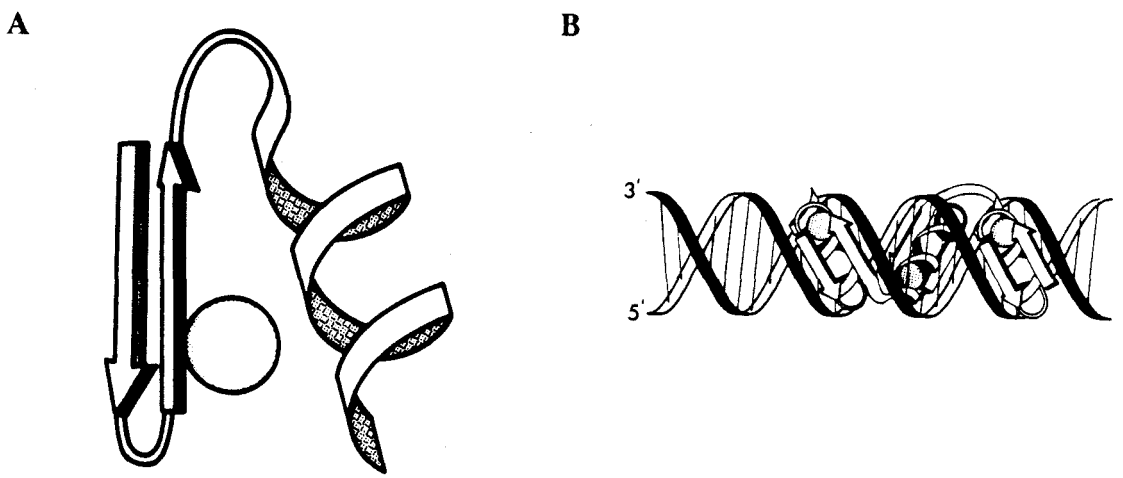


Figure 4 A) Schematic drawing of the zinc finger structure with two strands of antiparallel β-sheet, and an α-helix. B) Model for the interaction of a protein consisting of tandemly repeated zinc finger domains and DNA based upon the structure of the individual domains.

Leucine Zippers. The leucine zipper proteins constitute another structural motif for DNA binding which may be more general. This class of proteins includes the GCN4 transcriptional activator, the jun and fos oncoproteins, and the C/EBP enhancer binding protein.[36] This structural motif consists of a dimerization domain made up of a pair of amphipathic α helices which contain leucines spaced every seven residues and a basic region responsible for DNA binding. Although this structural motif has not been characterized by NMR or crystallographic techniques, solution studies employing footprinting[37] and affinity cleaving[38] techniques are consistent with a "scissor-grip" model as shown in Figure 5.

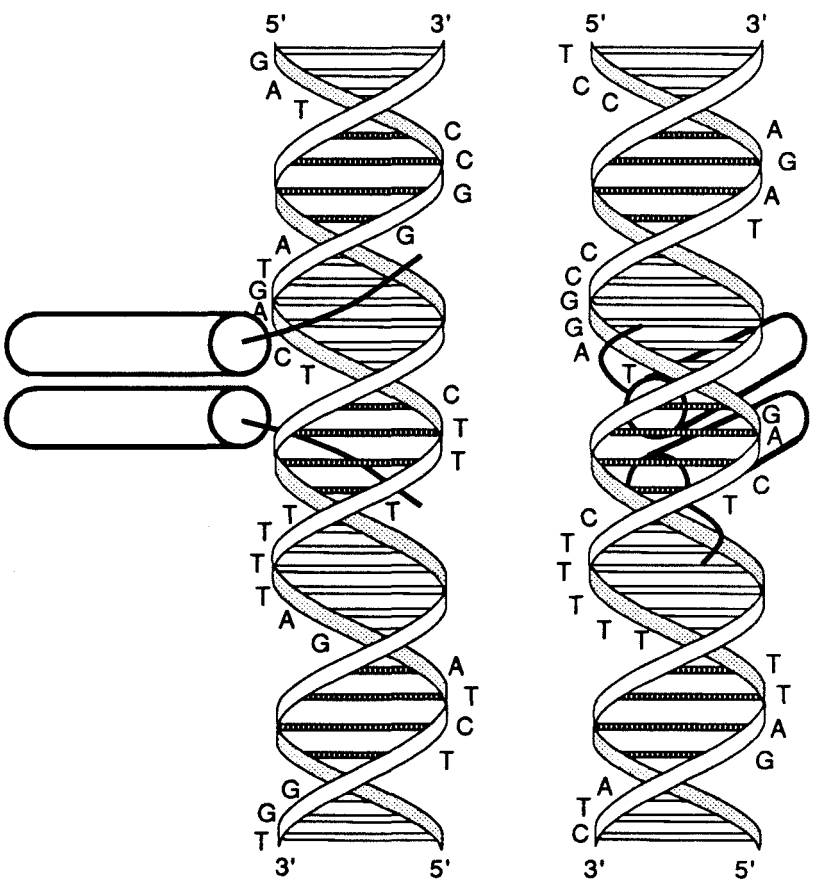


Figure 5 Side and front views of a Y-shaped model for the dimer of the DNA binding domain of GCN4. The leucine zipper dimerization domains are represented by a pair of cylinders pointing into the major groove of DNA.

Helix-Turn-Helix-DNA Cocrystals. Recently the DNA cocrystals of a number of helix-turn-helix proteins have been solved, supporting the proposed binding model.[39-41] These structures taken together indicate the complexity in structure and differences in sequence specific recognition that can occur even within a well-defined recognition motif.[42] The high resolution DNA cocrystal of λ repressor represents a straightforward means of recognition (Figure 6A).[39] The DNA duplex is not distorted or bent very much away from the idealized B-form. A network of main chain and side chain hydrogen bonds to the phosphate backbone position the helix-turn-helix unit on the DNA. The protein then makes specific hydrogen bonds with the edges of the base pairs in the major groove which are responsible for sequence specific recognition. Although this seems like a simple model for DNA binding the authors point out that there does not appear to be any strict rule telling how an amino acid will be used or how a particular base will be recognized.

The picture of how a helix-turn-helix protein recognizes its DNA sequence becomes even more complicated upon examination of the DNA cocrystal of the 434 repressor (Figure 6B).[40] In this structure there are again specific hydrogen bonds from amino acid side chains to the edges of the DNA base pairs in the major groove which help define the sequence specificity of the protein. However, in this structure an extended set of contacts between the protein and the phosphate backbone constrain a bent and twisted DNA structure. Thus the protein appears to be binding with an induced fit to the DNA, that is not only recognizing the DNA by direct hydrogen bonding to the base pairs, but also recognizing a particular DNA sequence by its ability to adopt the correct conformation for binding.

The third and probably most complicated structure to interpret is that of the DNA cocrystal of the trp repressor (Figure 6C).[41] Again, hydrogen bonds from the protein to the backbone anchor the protein to the DNA. However, in this structure there are very few specific hydrogen binds to the base pairs which specify a particular DNA sequence. For

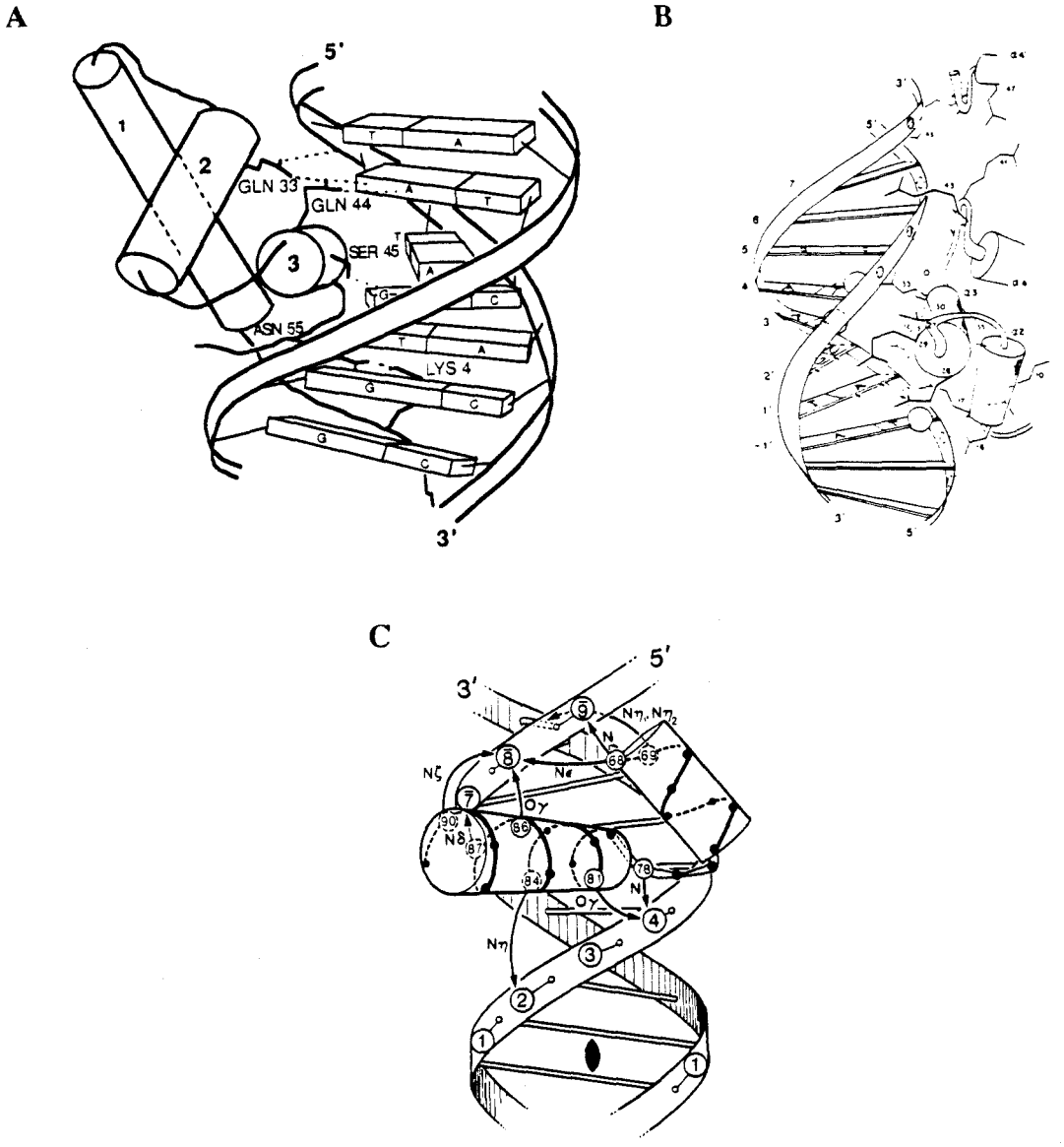


Figure 6 Stylized drawings of the helix-turn-helix regions of A) λ repressor,[39] B) 434 repressor,[40] and C) Trp repressor[41] on the DNA as obtained from the protein-DNA cocrystal structures. α helices are represented as cylinders.

this structure, two additional sources of specificity have been suggested. First, there are a number of solvent-mediated hydrogen bonds from amino acid side chains to base pairs through tightly bound water molecules, and second, the DNA again appears to be deformed to fit the surface of the repressor, indicating an induced fit based on DNA sequence. Further complicating this example is the fact that the recognition helix does not lie flat in the

major groove as in the two previous examples, but instead has its amino terminal end pointed in toward the groove.

These high resolution views reveal the complexity of protein-DNA interfaces. The α helices, linked by turns of varying length, contact the sugar-phosphate backbone as well as the base pairs in the major groove. The specificity depends on a set of correlated interactions and changing any one may affect others. The combination of direct protein-DNA contacts mediated by multiple hydrogen bonds and the sequence-dependent conformational effects in DNA limits our ability to make detailed structural predictions, even if a new protein can be assigned to a structural class such as the helix-turn-helix. Thus in the absence of high resolution crystallographic and nuclear magnetic resonance data, solution methods are needed to determine the topology of protein-DNA complexes and correlate sequence similarities with known structural classes.

Footprinting and Affinity Cleaving. The most common solution technique used to examine the nature of protein-DNA contacts is that of chemical or enzymatic footprinting.[43-45] These techniques allow the binding sites of a given molecule on DNA to be determined by the protection from cleavage imparted to the DNA by the bound molecule. Depending on the DNA cleavage reagent used, the binding site can be determined to nucleotide resolution. Footprinting data, however, does not give detailed information on where a particular amino acid or secondary structural element of a protein chain lies with respect to the DNA. Thus it is not possible to derive any information about how the protein binds to the DNA; it is only possible to determine where the protein lies along the DNA. In order to answer some of these more detailed questions the complimentary technique of affinity cleaving has been developed in our group.

Attachment of EDTA-Fe to a DNA binding moiety creates a DNA cleaving molecule that functions at physiologically relevant pH, temperature, and salt conditions.[46,47] The cleavage reaction can be initiated by the addition of a reducing agent such as dithiothreitol or sodium ascorbate. If the DNA-binding molecule is sequence specific, the EDTA-Fe

cleaves at highly localized sites on DNA restriction fragments and plasmids.[44,48-52] Because the cleaving moiety is not sequence specific, the cleavage specificity is determined only by the binding specificity of the molecule being investigated. EDTA•Fe-equipped DNA binding molecules cleave DNA by oxidation of the deoxyribose backbone via a diffusible oxidant, presumably hydroxyl radical.[44-47,49,53] Cleavage of both DNA strands is observed and typically extends over 4-6 base pairs.[44,48,49] If the DNA restriction fragment has been labeled with ^{32}P at a single end, the cleavage products can be analyzed by electrophoresis on high resolution denaturing polyacrylamide gels and visualized by autoradiography. Comparison of the individual cleavage lanes to sequence marker lanes then allows the assignment of the DNA cleavage events to nucleotide resolution (Figure 7). From this analysis binding site preference and size, as well as orientational preference information, can be obtained.

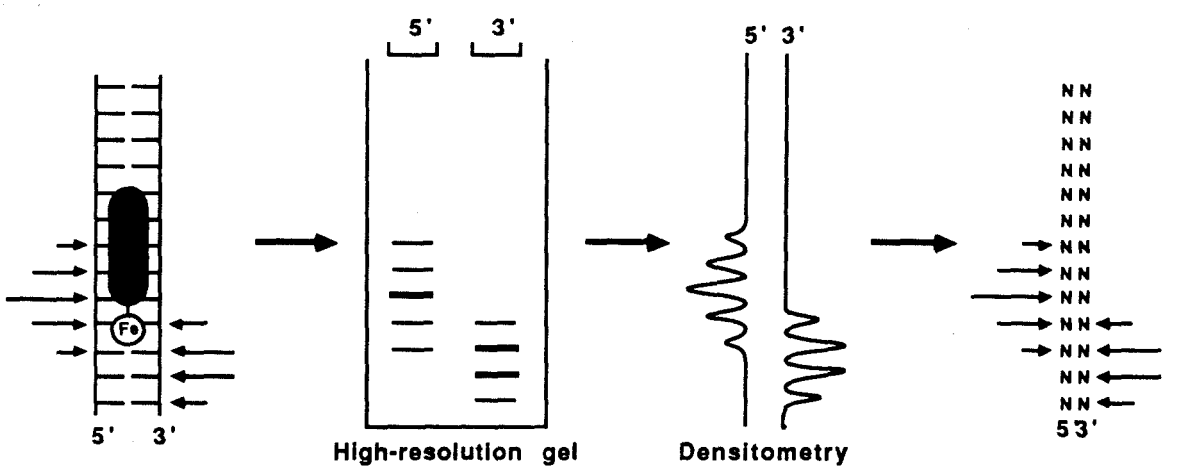


Figure 7 Outline of the high resolution assay used in affinity cleaving.

Due to the right-handed nature of double-helical DNA, the groove in which the EDTA•Fe is located can also be identified by complimentary strand analysis of the cleavage pattern. An asymmetric cleavage pattern with maximal cleavage loci shifted to the 3' side on opposite strands corresponds to an EDTA•Fe location in or above the minor groove

(Figure 8). When the EDTA is located in the major groove, the maximal cleavage loci are 5'-shifted; in addition, there appears cleavage of lower efficiency on the distal strands of adjacent minor grooves. This results in a pair of 3'-shifted asymmetric cleavage loci of unequal intensity on opposite strands (Figure 8). These patterns can be explained if the diffusible radical generated from the localized EDTA•Fe reacts with the major and minor grooves of DNA with unequal rates and preferentially (though not exclusively) in the minor groove. Several lines of evidence support this view. Tullius[53] has shown that a sinusoidal cleavage pattern is obtained when a DNA is bound to a precipitate of calcium phosphate and allowed to react with EDTA•Fe, demonstrating that the two grooves have different reactivity toward EDTA•Fe. From work in our own group, we find that EDTA•Fe attached to minor-groove-binding molecules afford cleavage patterns with a single 3'-shifted asymmetric cleavage locus per EDTA•Fe position.[44,49] However, for the case where EDTA•Fe is bound in the major groove of DNA by oligonucleotide triple helix formation, cleavage occurs along both strands of the adjacent minor grooves to afford a pair of cleavage loci of unequal intensity.[52,54] Cleavage is more intense along the strands proximal to the EDTA•Fe than along the distal strands of the adjacent minor grooves (Figure 8).

Affinity cleaving has been used to study the sequence-specific recognition of double helical DNA by naturally occurring DNA-binding antibiotics,[48,49] designed peptide analogues for the minor groove,[55-57] oligonucleotide-directed cleavage of single-stranded DNA,[50] and oligonucleotide-directed recognition of the major groove by triple helix formation.[52,54,58,59] From these studies, which involve less complicated DNA-binding motifs, numerous cleavage patterns caused by a diffusible oxidant in either the minor or major groove of duplex DNA have been analyzed. This data base allows interpretation of the affinity cleaving results from conformationally more complex proteins.

Incorporation of EDTA•Fe at discrete amino acid residues within a protein should allow the position and groove location of the modified residues to be mapped to nucleotide resolution. The secondary and tertiary structures of DNA-binding proteins can be analyzed

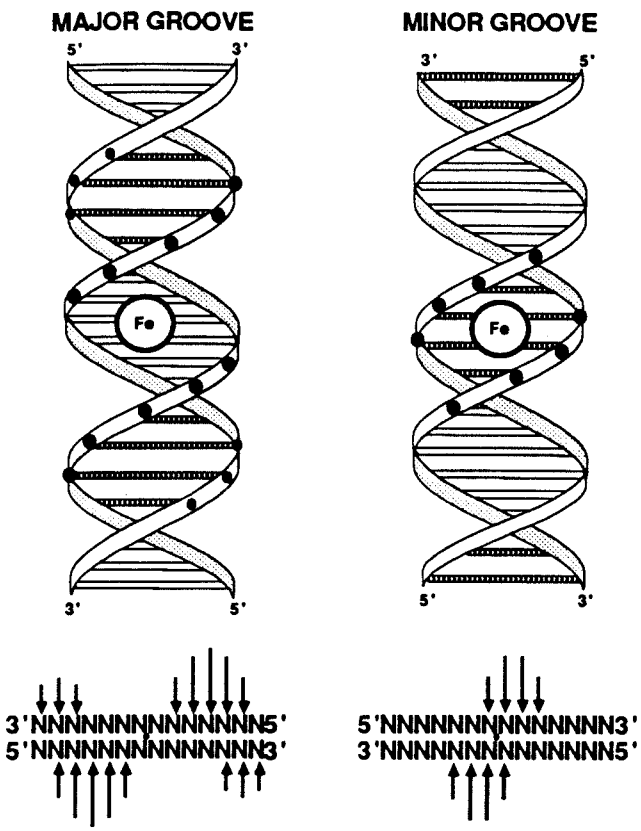


Figure 8 Cleavage patterns produced by a diffusible oxidant generated by Fe•EDTA located in the major and minor grooves of right-handed DNA. Filled circles represent points of cleavage along the phosphodiester deoxyribose backbone. Sizes of the circles represent extent of cleavage.

by affinity cleaving by using two approaches. First, the amino acid location of the EDTA on a polypeptide chain of constant length may be varied in order to reveal key topological features of the protein-DNA complex (e.g., location of the NH_2 versus COOH terminus). Alternatively, the length of the polypeptide may be incrementally changed while the EDTA is kept at the same teminus to examine the influence of substructures on the binding affinity of the protein, the base and groove location of the modified position, and the conformational flexibility of a specific peptide segment.

Hin Recombinase. Affinity cleaving was initially used to study the DNA binding domain of Hin recombinase. Hin recombinase is a 190 amino acid enzyme that inverts a

segment of DNA which controls the expression of the flagellin genes of *Salmonella typhimurium*.[60] Recombination occurs between two crossover sites, designated hixL and hixR, on supercoiled DNA.[61] Each hix site is 26 base pairs long and has nearly 2-fold symmetry. Hin binds to a hix site as a dimer and protects bases -13 through +13, inclusive, from chemical cleavage reagents (consensus half site: 5'-TTNTCNNAAACCA-3').[61,62] A synthetic 52 amino acid protein identical with the COOH terminus of Hin(139-190) has been shown by DNAse I and dimethyl sulfate protection experiments to contain the sequence-specific DNA-binding activity of Hin.[63] This protein fragment, Hin(139-190) protects the same DNA sequence as does the complete Hin protein except for the central three bases from -2 to +1.

The 52-residue DNA-binding region of Hin contains extensive sequence similarity to the helix-turn-helix DNA-binding proteins and is thought to recognize DNA through this motif in the major groove.[24] However, dimethyl sulfate footprinting experiments have shown that Hin recombinase and Hin(139-190) protect not only the expected major groove contacts but also residues -6 to -4 , and 4 to 6 of hixL from methylation. These results have been taken as evidence that protein contacts also occur in the minor groove.[62,63] In an attempt to study the DNA binding domain of Hin recombinase by affinity cleaving, methods have been developed in our group to attach a protected EDTA derivative to the amino terminus of a protected, resin-bound protein synthesized by solid-phase techniques.

EDTA Derivatives for Solid-Phase Peptide Synthesis. Protecting groups for the EDTA carboxylic acids were chosen based on standard side chain protection methods used for glutamic and aspartic acids in Merrifield solid phase protein synthesis employing N-t-butoxycarbonyl (Boc) amino acid derivatives.[64-67] The first derivative tribenzyl-EDTA-GABA (BEG),[68] has three of the four carboxyl groups of EDTA protected as benzyl esters. The fourth carboxyl is coupled via an amide bond to a γ-aminobutanoic acid (GABA) linker designed to minimize disruptions of the protein structure by the attached EDTA•Fe chelates. BEG can be coupled to the amino terminus of a protected resin-bound

peptide as its hydroxybenzotriazole (HOBt) ester in good yield (>98%) to produce upon cleavage with HF the EDTA-GABA-Hin(139-190) species (Figure 9). The second EDTA derivative is the tricyclohexyl ester of EDTA (TCE)[68] which has been developed as an alternative to BEG. TCE offers more flexibility than BEG in that a variety of amino acid linkers (or no linker at all) may be used. TCE like BEG couples in good yield to the amino terminus of a resin-bound peptide. Deprotection with HF again affords the EDTA-protein (Figure 10). Though BEG and TCE are formally amino acids, they are suitable only for capping an amino terminus or a side chain amino group since they do not have a primary or secondary amino group for further chain extension.

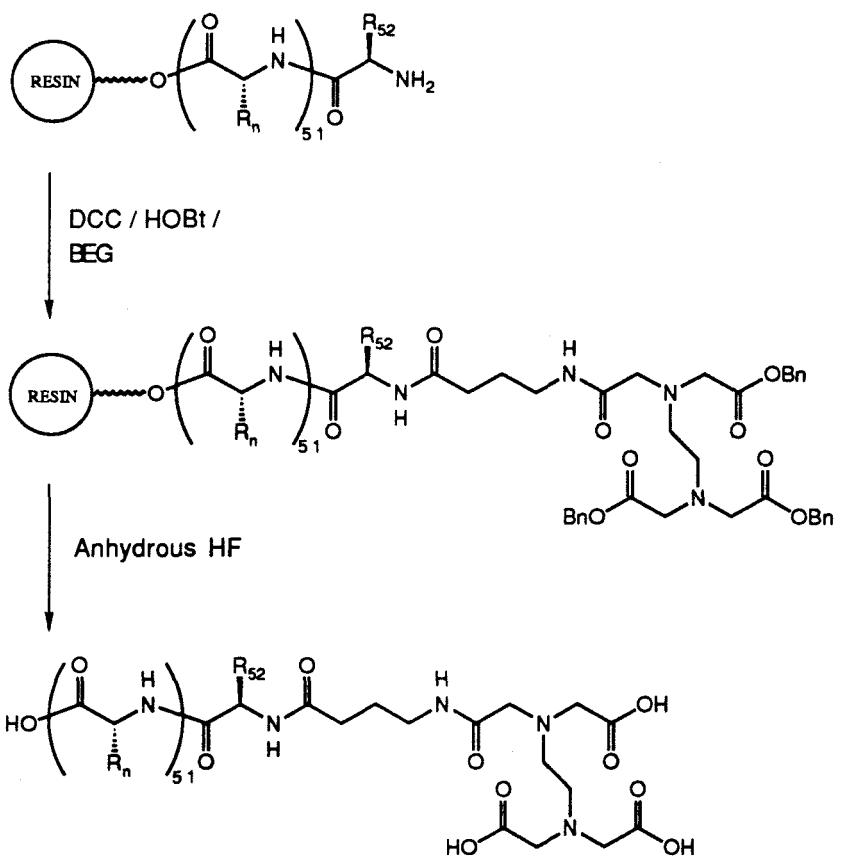


Figure 9 Scheme for attaching BEG to the amino terminus of a protein using solid phase methods.

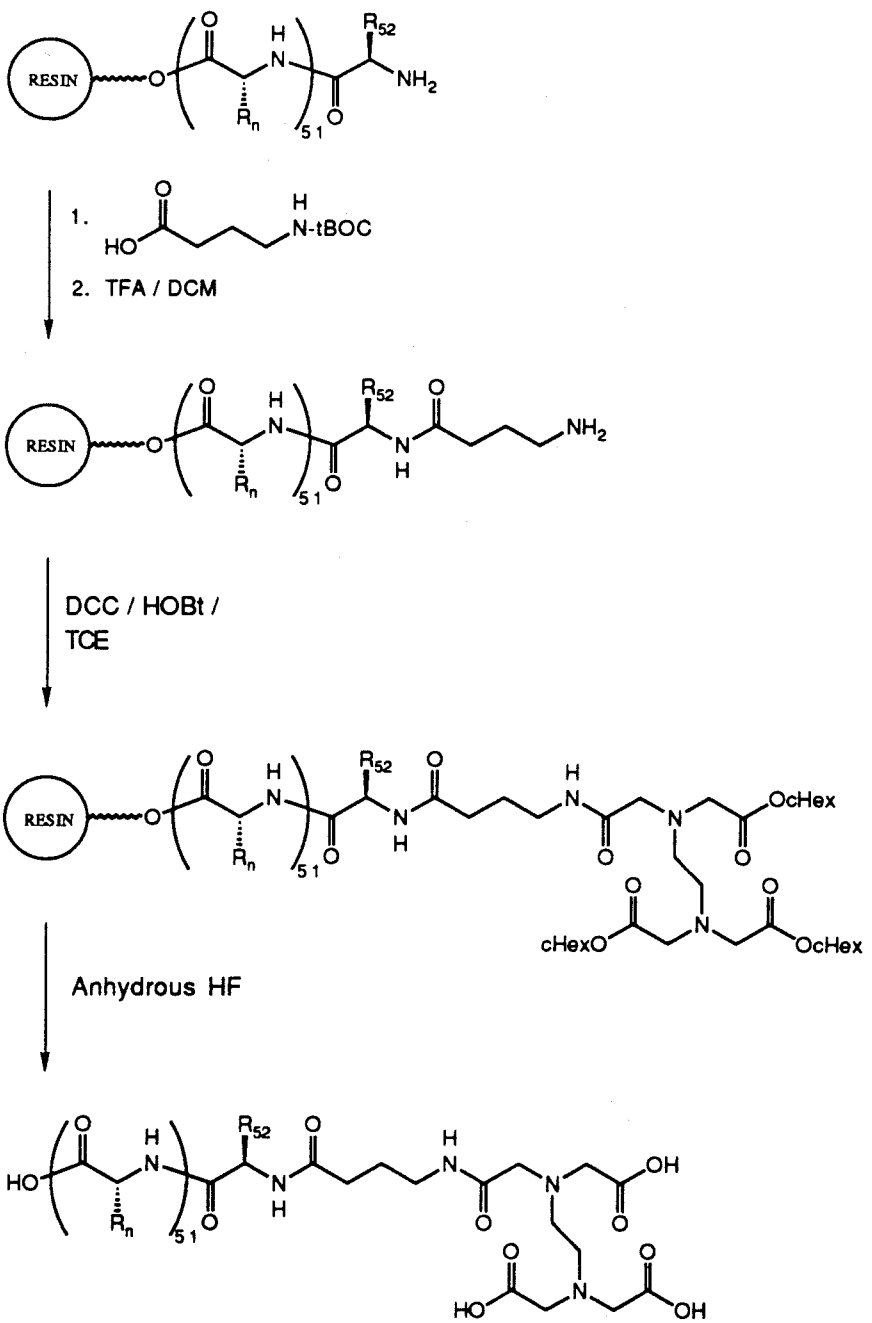


Figure 10 Scheme for the attachment of TCE to the amino terminus of a protein using solid-phase methods. In this example GABA is used as a linker between the amino terminus and the EDTA.

Affinity Cleaving Studies with EDTA-Hin(139-190). BEG was attached to the amino terminus of the DNA binding domain of Hin recombinase by slolid-phase techniques to

afford upon deprotection the EDTA containing protein.[51,69] Affinity cleaving studies with Fe•EDTA-Hin(139-190), show cleavage of the DNA near the symmetry axis of the Hin recombination sites.[51,69] The cleavage pattern is shifted to the 3' side, suggesting that the NH_2 terminus of Hin(139-190) is bound near or in the minor groove of DNA. As the bulk of the DNA-binding domain is likely located in the adjacent major groove as a helix-turn-helix motif, residues close to the NH_2 terminus must extend across the DNA phosphodiester backbone to the minor groove. This is consistent with results from DMS protection studies which indicate minor groove protection by Hin. The four residue segment at the amino terminus of the binding domain, Gly^{139}-Arg^{140}-Pro^{141}-Arg^{142}, contains two basic side chains and could be involved in such interactions. In further affinity cleaving studies directed toward investigating the importance of these positively charged residues for DNA binding, amino terminal EDTA proteins which have been truncated by 1, 2, or 3 amino acids at the amino terminus were studied.[69] Footprinting and affinity cleaving reveal that deletion of Gly^{139} results in a protein with affinity and specificity similar to that of Hin(139-190) but that the deletion of Gly^{139}-Arg^{140} affords a protein with altered affinities and sequence specificities for the five binding sites studied. It appears that Arg^{140} in the DNA binding domain of Hin is important for the recognition of the 5'-AAA-3' sequence in the minor groove of DNA.

A binding model for the DNA-recognition domain of Hin might therefore involve at least two sets of interactions: (i) a helix-turn-helix structure binding in the major groove and (ii) a region that connects the helix-turn-helix domain with the other 140 residues of Hin by following the adjacent minor groove sequence 5'-AAA-3' toward the center of the hix site. A model that incorporates these features was put forward[51,69] and is shown in Figure 11. The three α helices proposed in this model are based on the α_1-α_2-α_3 helices of λ $Cro^{14\text{-}16}$ and 434 repressor.[40] The model also shows the NH_2 terminus of the 52mer extending along the minor groove toward the center of the dimer binding site. As

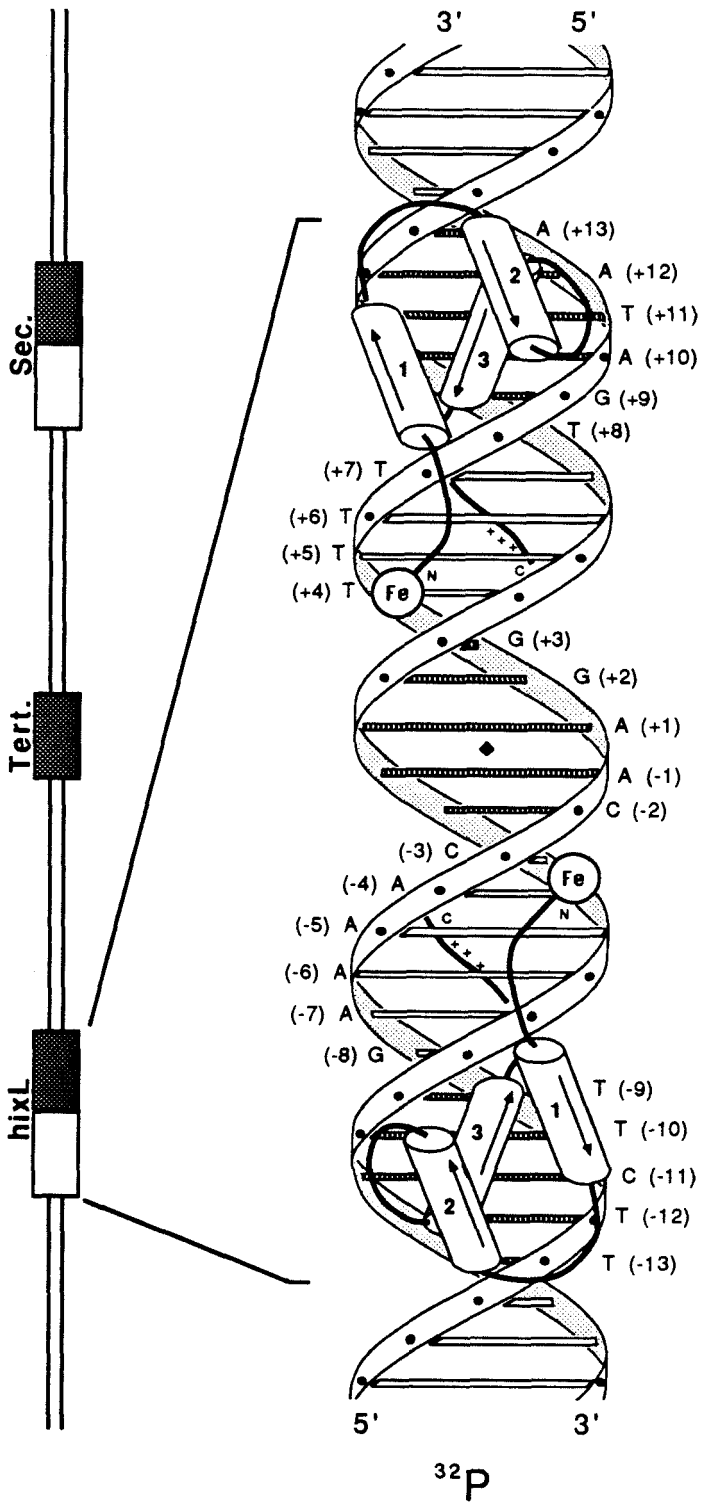


Figure 11 Model for Hin(139-190) binding to hixL.[69] The sequence of hixL is shown along the 5'-strand. The pseudo-C_2 axis is indicated by a solid diamond. Putative α helices are shown as cylinders with an arrow pointing from NH_2 to COOH terminus.

demonstrated by footprinting studies, the positions contacted by Hin(139-190) in the major grooves are separated by two turns of the DNA helix.[51]

By use of the affinity cleaving technique information has been gained as to the location of the NH_2 terminus of the DNA binding domain of Hin recombinase on the DNA binding site as well as importance of the amino terminal sequence of the protein for DNA binding. Results with amino terminal EDTA proteins indicate that the NH_2 terminus is located in or near the minor groove toward the symmetry axis of the binding site and that Arg^{140} is important for the recognition of the 5'-AAA-3' sequence in the minor groove. Detailed structural results such as these are normally only available through crystallographic or NMR studies. The use of affinity cleaving proteins thus appears to be an alternative technique for the characterization of protein-DNA complexes in the absence of other structural data.

Further Characterization of Protein-DNA Complexes by Affinity Cleaving

Previous results in our group led to binding models for Hin(139-190) which include a helix-turn-helix structure in the major groove with residues at the NH_2 terminus extending across the DNA phosphodiester backbone to the adjacent minor groove.[51,69] These results were obtained by attaching EDTA to the amino terminus of the DNA binding domain of Hin recombinase. In order to further study DNA binding proteins it will be necessary to position the EDTA at a different location along the polypeptide chain. By moving the EDTA to a different amino acid position it should be possible to determine the where that specific residue lies with respect to the DNA, and thus obtain a more detailed picture of the protein-DNA complex. Techniques for the attachment of EDTA to a lysine side chain which take advantage of the nature of BEG and TCE as capping reagents were developed. These techniques have been applied to the DNA binding domain of Hin

recombinase to determine the location of the COOH terminus of the recognition helix and to verify its orientation along the DNA.

Orientation of the Putative Recognition Helix of Hin Recombinase. λ Cro, λ repressor, and 434 repressor bind to pseudo-C_2-symmetric DNA sequences as dimers.[24] High-resolution X-ray crystal structures of λ repressor and 434 repressor (1-69) with their operators reveal that the COOH termini of the recognition helices are oriented toward the symmetry axis of the dimer-binding site.[39,40] However, NMR data with the *lac* repressor headpiece bound to operator DNA indicate an *opposite orientation* with the COOH terminus of the recognition helix facing away from the symmetry axis of the binding site.[70] Two models for the DNA binding domain of Hin were put forward that differ in the orientation of the recognition helix.[51] One is based on the orientation found in 434 repressor (1-69) and λ repressor (Figure 12A), and the other is based on lac repressor headpiece (1-56) (Figure 12B). In order to distinguish between these two models for Hin(139-190), we have carried out affinity cleaving experiments designed to determine the orientation of the putative recognition helix in the major groove of the Hin-binding site (Figure 12).

To determine the orientation of the putative recognition helix of the Hin DNA-binding domain in the major grove of the DNA binding site, residues 139-190 were synthesized with EDTA near the COOH terminus. However, the cleavage pattern observed on the hixL binding site by using this synthetic protein was significantly more diffuse than the typical 4-6-base-pair cleavage pattern consistent with a highly localized EDTA•Fe.[71] The lack of a localized cleavage pattern suggests that the COOH terminus is conformationally flexible or extends some distance from the grooves of the hixL site. This lack of a discrete cleavage pattern does not allow an orientational assignment of the COOH terminus of Hin(139-190).[71] One solution to this problem is to truncate the COOH terminus near the end of the postulated binding helix (Pro^{181}), which presumably will leave the COOH terminus in a relatively fixed position near the DNA. Before attaching EDTA to the COOH terminus of the truncated Hin DNA-binding domain, it was necessary to determine whether the shorter

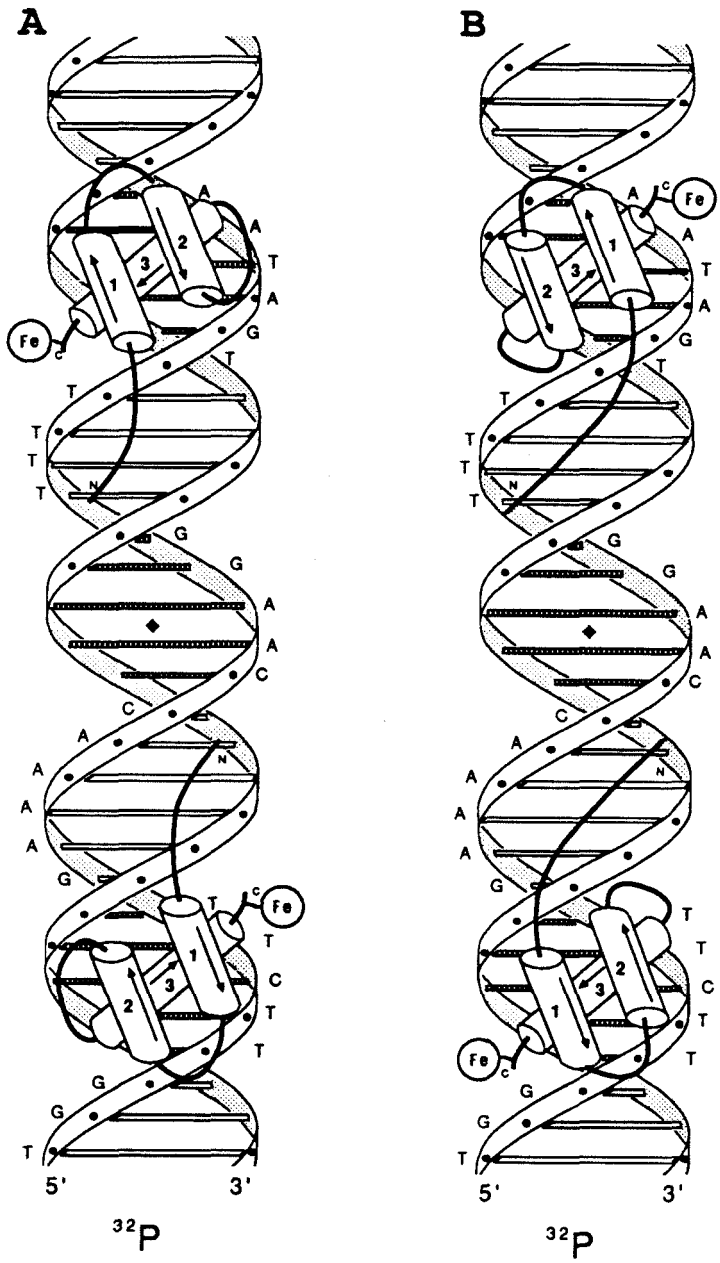


Figure 12 Two models for the DNA recognition domain of Hin(139-190) binding to hixL. The sequence of hixL is shown along the 5' strand. The pseudo-C2 axis is indicated by a solid diamond. Putative α helices are shown as cylinders with an arrow pointing from the NH_2 to the COOH terminus. Model A is based on the 434 repressor-DNA cocrystal and model B is based on the NMR studies of the lac repressor headpiece. EDTA•Fe is shown attached to the side chain of a lysine at position 183 which replaced a Ser residue.

protein would fold properly and retain the same affinity and sequence specificity for the hixL site as Hin(139-190).

Hin is a member of a family of recombinases which includes Gin from phage Mu,[72] Cin from phage P1,[73] and Pin from the e14 element of *E. coli* .[74] Examination of these protein sequences indicates that, after the position corresponding to Ser^{184} of Hin, there is variability present in the length and content of the remaining residues at the COOH termini of the proteins (Figure 13). We have synthesized a protein which lacks six residues at the COOH terminus of Hin(139-190) to afford a 46-residue protein, Hin(139-184). EDTA has been attached to the NH_2 terminus of this protein to yield EDTA-Hin(139-184) for comparison with EDTA-Hin(139-190). We find that Fe•EDTA-Hin(139-184) cleaves DNA in a sequence-specific fashion similar to that of Fe•EDTA-Hin(139-190). It is therefore possible to attach EDTA near the COOH terminus of the putative recognition helix of Hin(139-184) to determine its orientation.

Synthesis: Three 46-residue proteins were synthesized with EDTA at the NH_2 terminus, the COOH terminus, and both termini (Figure 14). Hin(139-184) was synthesized by automated solid-phase techniques on benzhydrylamine (BHA) resin using tBoc protected amino acids. BHA resin was chosen such that upon deprotection with HF a carboxamide would be formed at the COOH terminus instead of a free carboxylate, thus neutralizing a negative charge which would not exist in the native protein.[75] The average yield per cycle, as determined by quantitative ninhydrin analysis of the resin sample,[66] over 45 couplings was 99.4%. BEG was attached to the NH_2 terminus of the protected peptide-resin as previously described.[68] Attachment of EDTA near the COOH terminus was accomplished by a combination of tBoc and Fmoc protection schemes (Figure 15). N^{ε}-Fmoc-N^{α}-tBoc-lysine was substituted for the Ser^{183} of the protein. The Fmoc protecting

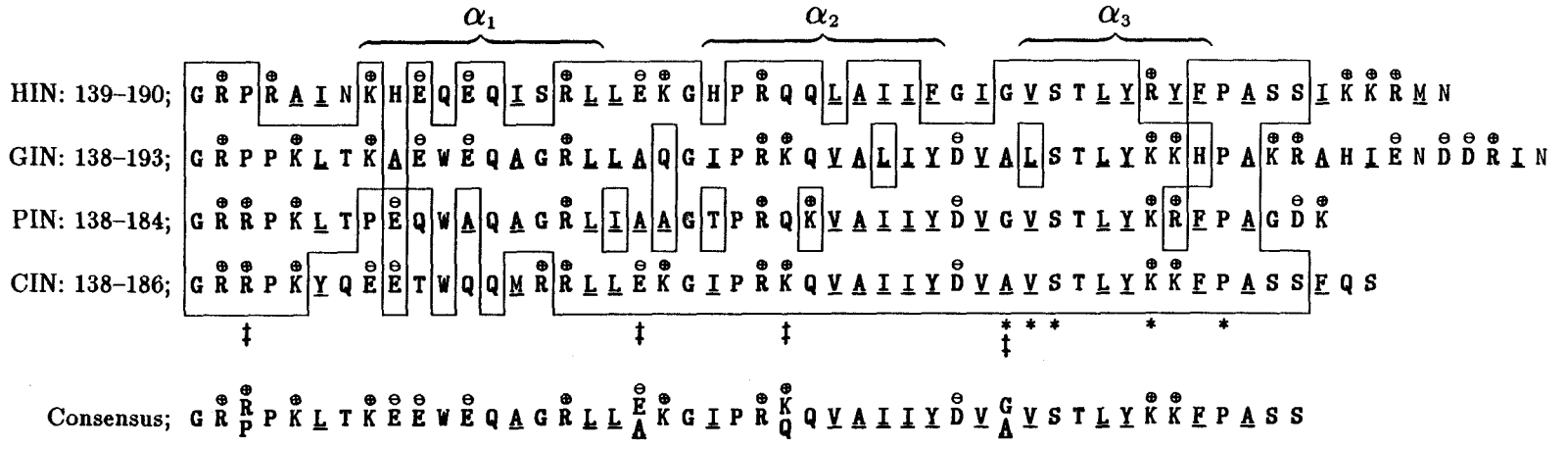


Figure 13 Protein sequences for the COOH-terminal domains of the Hin family of recombinases. Hydrophobic residues are underlined, and charged residues are marked with the appropriate sign. Net charges, excluding amino- and carboxy-terminal charges, for these domains are Hin +7, Gin +5, Pin +6, and Cin +6.

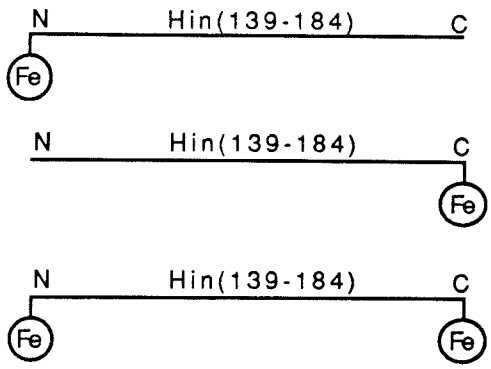


Figure 14 Synthetic proteins from the DNA-binding domain of Hin recombinase with EDTA•Fe positioned at the NH_2 terminus, the COOH terminus and at both termini of residues 139-184.

group was then selectively removed by using piperidine in DMF.[75] Coupling of TCE to the side-chain amine was accomplished by using DCC and HOBt in DMF.[68] Ninhydrin analysis indicated a 99.9% yield in 1 hour. The tBoc group was then removed by using TFA in DCM. The synthesis was continued manually for two residues and completed by using automated techniques with an average yield of approximately 99.1% per cycle. BEG was attached to the NH_2 terminus of a small portion of this protected peptide resin to afford EDTA-Hin(139-184)-EDTA (Figure 14). All proteins were deprotected with anhydrous HF and purified by reverse-phase HPLC.

Affinity Cleaving: The proteins were allowed to react at micromolar concentrations (22 °C, pH 7.5, 20 mM NaCl) for 1 hour in the presence of DTT with a ^{32}P-end-labeled restriction fragment containing the Hin-binding sites hixL and the secondary Hin.[63]. The DNA cleavage products were separated by high-resolution polyacrylamide gel electrophoresis and visualized by autoradiography (Figure 16). By comparison of Hin(139-184) and Hin(139-190), both with EDTA at the NH_2 terminus, we find that Fe•EDTA-Hin(139-184) cleaves DNA with the same sequence specificity as Fe•EDTA-Hin(139-190) but with reduced efficiency (Figure 16, compare lanes 7 and 8 with lanes 5 and 6). When EDTA•Fe is moved from the NH_2 to the COOH terminus of Hin(139-184), we find that the DNA

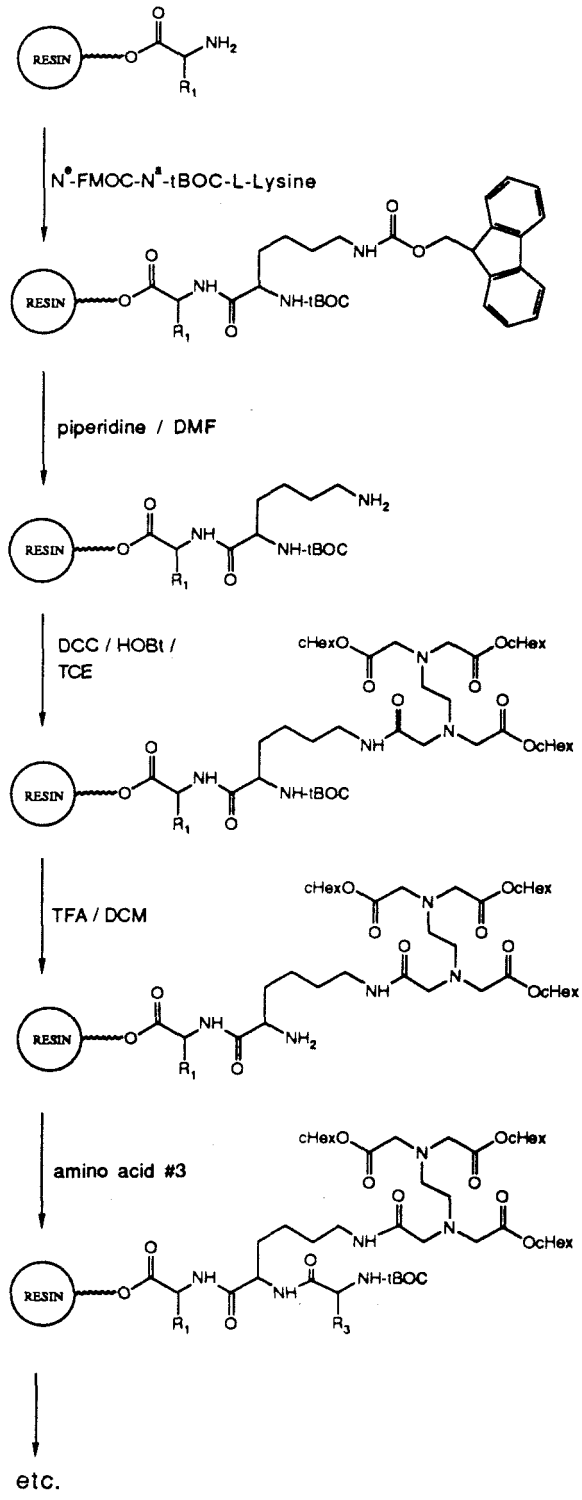


Figure 15 Synthetic scheme for the attachment of TCE to the ε-amino group of Lys^{183} in the second position on the resin.

cleavage locus for Hin(139-184)-EDTA•Fe is shifted further from the symmetry axis at the binding sites compared to Fe•EDTA-Hin(139-184) (Figure 16, lanes 9 and 10). Remarkably, the maximal cleavage patterns observed on opposite strands of the DNA for Hin(139-184) with EDTA•Fe at the COOH terminus are shifted to the 3' direction. A protein equipped the EDTA at the NH_2 and COOH termini, Fe•EDTA-Hin(139-184)-EDTA•Fe, affords a pair of cleavage patterns consistent with the combination of patterns from EDTA-Hin(139-190) and Hin(139-184)-EDTA (Figure 16, lanes 11 and 12 and Figure 17). The fact that the cleavage pattern for EDTA•Fe at Gly^{139} (NH_2 terminus) appears unchanged when EDTA•Fe is present or absent at the COOH terminus suggests that Hin(139-184) maintains the same structure independent of which terminus is modified with EDTA•Fe.

Histograms of cleavage at the hixL and secondary sites, which consist of imperfectly conserved 12-bp inverted repeats, on a DNA restriction fragment are shown in Figure 17. Although the binding-site specificity for Hin(139-184) and Hin(139-190) are identical, the binding affinities are different, suggesting that nonspecific affinity has been lost in the shorter protein. Three of the six residues removed contain positive charges, and it is possible that these residues, Lys^{186}-Lys^{187}-Arg^{188}, make important electrostatic contributions (at 20 mM NaCl). Thus, the removal of six residues (185-190) from the COOH terminus of the protein apparently leaves the tertiary structure and the sequence-specific binding elements of protein intact.

The specific cleavage pattern observed for Hin(139-184)-EDTA•Fe is shifted to the 3' side and indicates that the Fe•EDTA attached near the COOH terminus of the recognition helix is positioned within the Hin-binding site in or above the minor groove or the sequence 5'-GGTT-3'. The 3' shift of the cleavage patterns is unexpected for an EDTA positioned in the major groove, as would be expected for the helix-turn-helix structure. There are two possible explanations which alone or together could account for the observed pattern. First, the COOH terminus of the recognition helix does not have to be centered in the major

Figure 16 Autoradiogram of a high-resolution denaturing polyacrylamide gel of affinity cleaving reactions on a ^{32}P-end-labeled fragment (XbaI/EcoRI) from pMFB36.[63] Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in bp) calf thymus DNA, 5 mM DTT, and ≈15,000 cpm of end-labeled DNA in a total volume of 20 μL. Reactions were run for 60 min. at 25 °C and terminated by ethanol precipitation. Cleavage products were analyzed on an 8%, 1:20 cross-linked, 50% urea polyacrylamide gel. Odd and even numbered lanes contain 5'- and 3'-end-labeled DNA, respectively. Lanes 1 and 2 are DNA controls, and lanes 3 and 4 are A-specific sequencing lanes. Lanes 5 and 6 contain 5 μM Fe•EDTA-Hin(139-190), lanes 7 and 8 contain 10 μM Fe•EDTA-Hin(139-184), lanes 9 and 10 contain 5 μM Hin(139-184)-EDTA•Fe, and lanes 11 and 12 contain 10 μM Fe•EDTA-Hin(139-184)-EDTA•Fe.

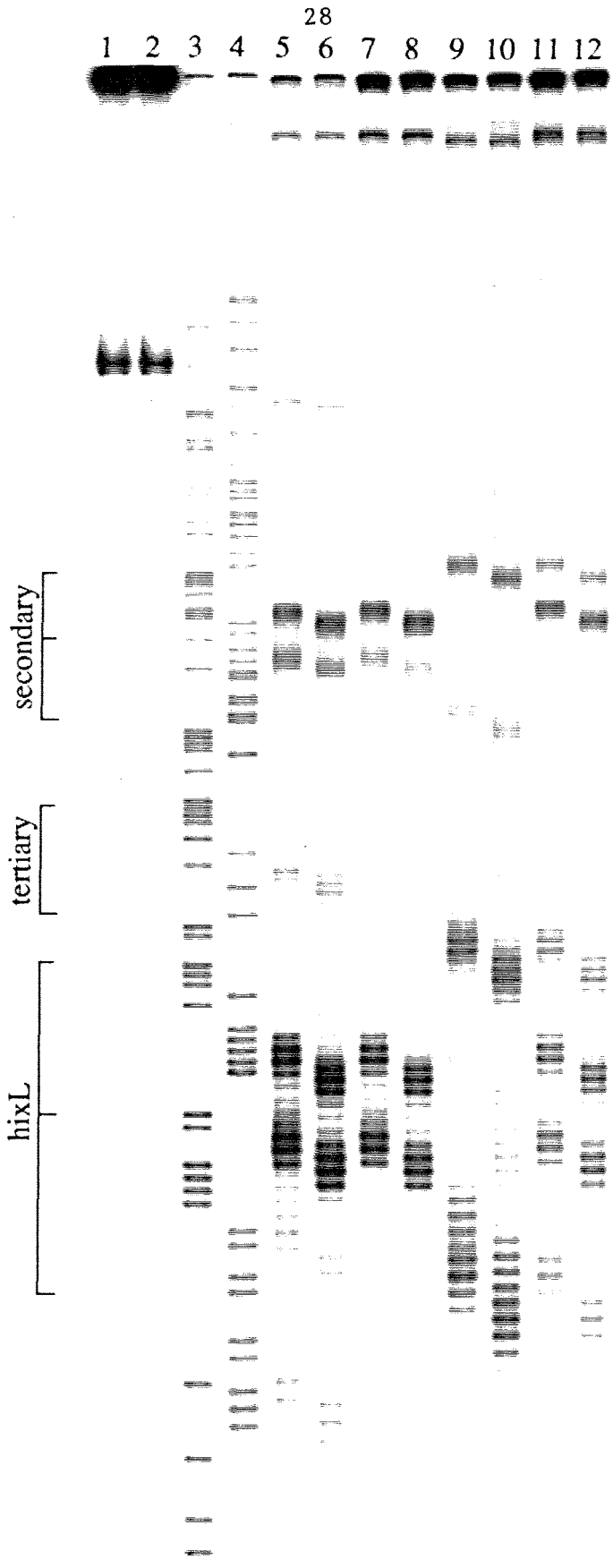
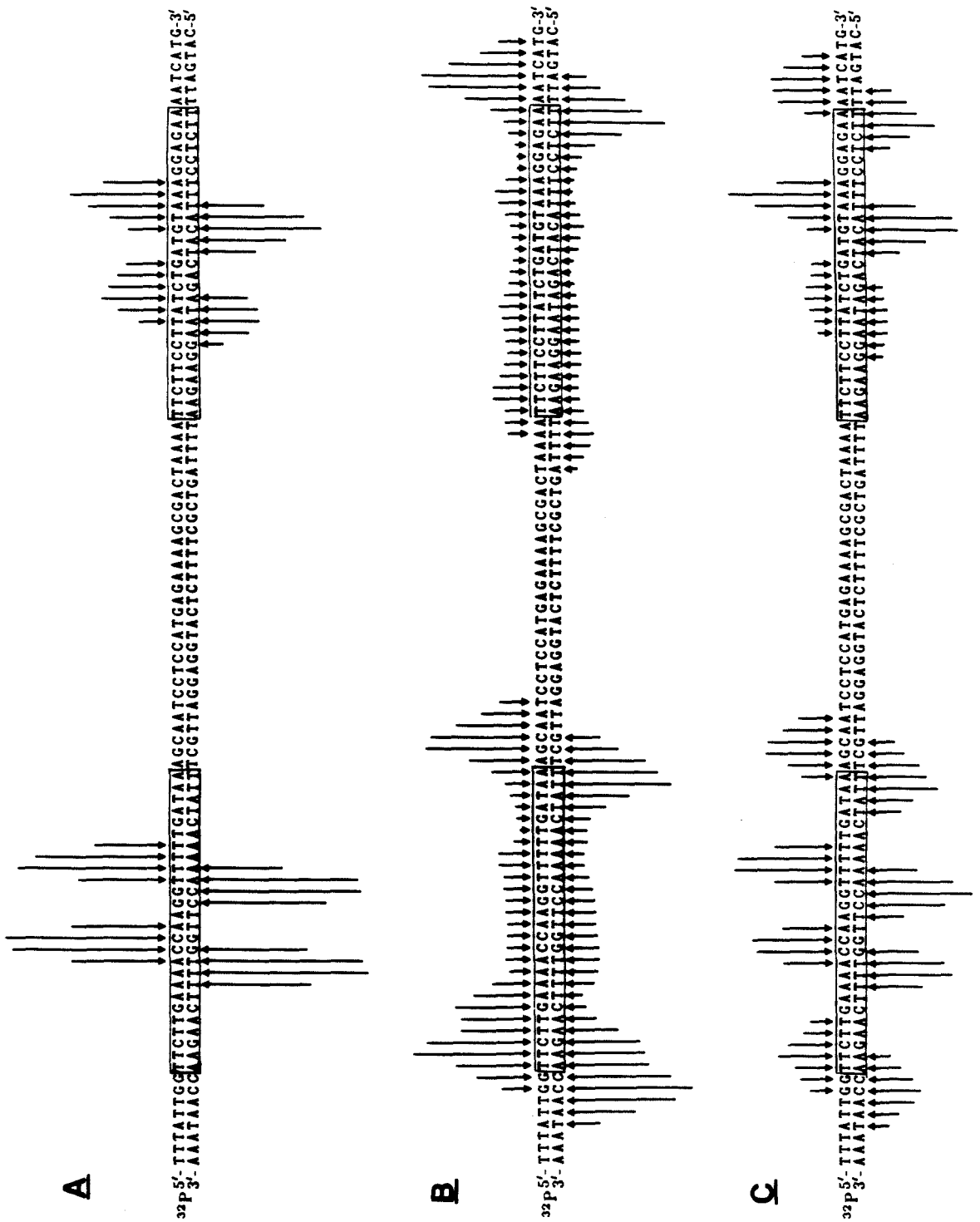

1 2 3 4 5 6 7 8 9 10 11 12
secondary
tertiary
hixL

Figure 17 Histograms of the cleavage data from Figure 16. The sequence left to right is from the bottom to the middle of the gel. The left and right boxed sequences represent hixL and the secondary site, respectively. Arrow heights indicate the extent of cleavage at the indicated base. Panels A-C are cleavage patterns for Fe•EDTA-Hin(139-190), Hin(139-184)-EDTA•Fe, and Fe•EDTA-Hin(139-184)-EDTA•Fe.


A
B
C
32P 5′- TTTATTGGTTCTTGAAAACCAAGGTTTTTGATATAAGCAATCCTCCATGAGAAAAGCGACTAAAATTCTTCCTTATCTGATGTAAAGGAGAGAAATCATG-3′
3′- AAATAACCAAGAACTTTTGGTTCCAAAAACTATTTCGTTAGGAGGTACTCTTTTCGCTGATTTTAAGAAGGAATAGACTACATTTCCTCTCTTTAGTAC-5′

groove. If the EDTA•Fe were located toward one side of the major groove, one of the pairs of 3' shifted patterns expected for an EDTA located in the major groove would be expected to be increased due to proximity of the EDTA•Fe to that minor groove, while the other 3' shifted pattern should be decreased. This difference could be further enhanced if the relatively long lysine side chain which contains the EDTA extends toward the closer minor groove. Second, if one of the minor grooves is in some way blocked, the cleavage pattern would be expected to decrease. This explanation also seems likely in the case of the DNA binding domain of Hin since it is known that the amino terminal end of the sequence contacts the adjacent minor groove at the sequence 5'-AAA-3'.[51,69] This possibility is further born out by the experimental data (Figure 17B) where, in fact, a region of DNA protection can be seen at the 5'-AAA-3' sequence where the amino terminus is proposed to bind. It seems likely that both of these explanations in part account for the observed cleavage pattern.

Regardless of the explanation for the cleavage pattern, its strong intense nature allows us to examine the two models for the orientation of the recognition helix in relation to the data. Due to the opposite orientation of the recognition helix in each model, the EDTA•Fe should be located on opposite faces of the DNA (Figure 12). Therefore, the location of maximal cleavage expected for the two orientations should be different; the λ and 434 repressor orientation is closer to the symmetry axis of the inverted repeats. Two models for the DNA-binding domain of Hin, one based on λ and 434 repressor-DNA cocrystals[39,40] and the other on *lac* repressor headpiece-DNA NMR studies,[70] were superimposed on the experimentally observed cleavage data (Figure 18). Comparison of the two models and the cleavage data obtained with an EDTA•Fe attached near the COOH terminus of Hin(139-184) supports the model where the recognition helix of Hin is oriented *toward the center of the hixL binding site* as in λ and 434 repressors (Figure 18A) and not as assigned for the *lac* repressor headpiece (Figure 18B).

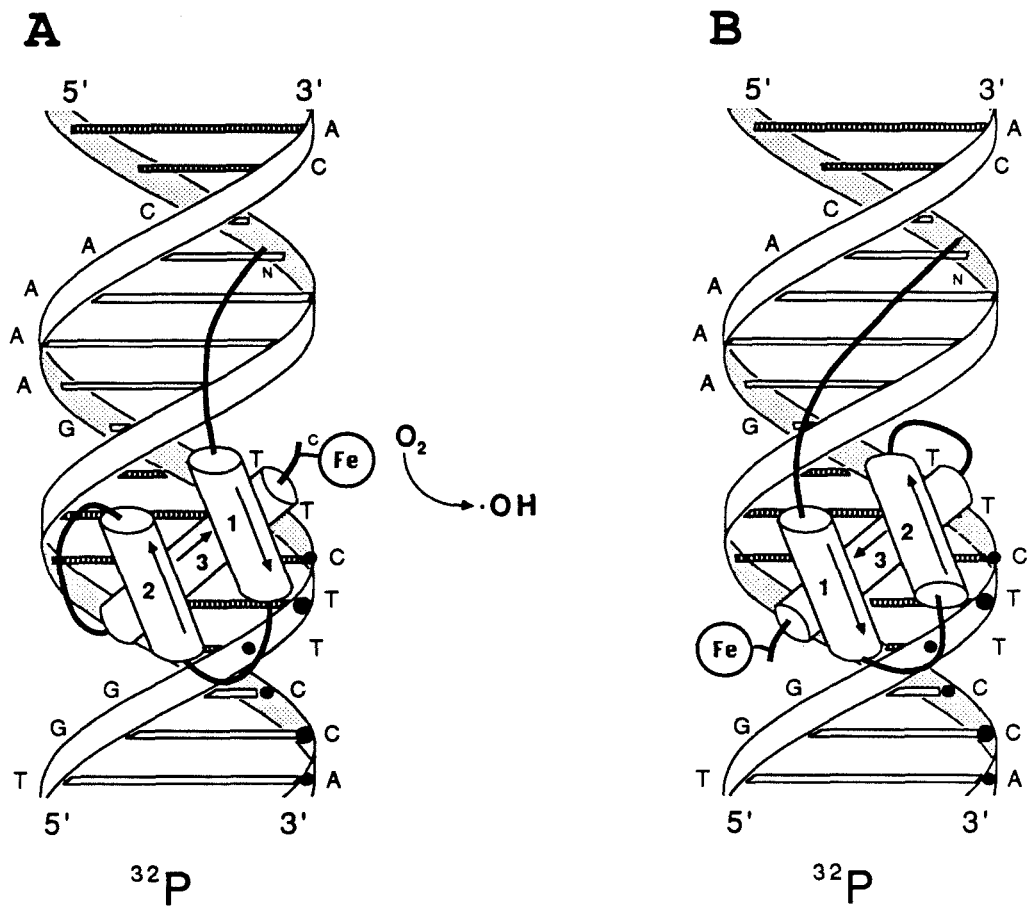


Figure 18 Schematic representation of two models of Hin(139-184)EDTA•Fe bound to the bottom half of hixL, with filled circles indicating the sites of maximal cleavage produced by Hin(139-184)-EDTA•Fe. The model on the left (A) seems to fit the experimental cleavage pattern the best.

Placement of EDTA at an Internal Position of the DNA Binding Domain of Hin. By placing EDTA at the amino or carboxy terminus of the DNA binding domain of Hin recombinase, we have taken the first steps toward mapping the tertiary structure of a protein on the DNA. The affinity cleaving technique provides the location of the Fe•EDTA complex to within one base pair on the DNA, and the exact nature of the cleavage pattern gives us information about the groove location, proximity to the DNA, and conformational rigidity of the amino acid to which it is attached. In order to gather further information about the structure of the DNA binding domain of Hin recombinase, it will be necessary to place the EDTA at an internal position along the amino acid chain instead of at one of the ends.

Attachment of EDTA to an internal position along the protein chain should be able to be accomplished by using the methodology developed to attach EDTA to the COOH terminus. Substitution of N^{ε}-Fmoc-N^{α}-tBoc-lysine for any residue along the chain should allow EDTA to be incorporated at that position by attachment to the lysine side chain. The position chosen for substitution was between helix 2 and helix 3 at Ile^{171}. This position was selected for two reasons: 1) since the location of the COOH terminus of the putative recognition helix had been determined, it was anticipated that the new data would complement that previously obtained by providing the location of the other end of the helix, and 2) because the lysine side chain with an EDTA attached will constitute a rather large group, it would seem logical to choose a position which already contains a large side chain in the original sequence. Although the protected lysine could be incorporated into the chain at this position, the Fmoc group could be removed, and EDTA could be attached in very good yield, it was not possible to successfully extend the chain after the tBoc group was removed (data not shown). An alternative strategy to attach the EDTA would be to leave the Fmoc protecting group in position throughout the remainder of the synthesis, deprotect it selectively due to the absence of any other base labile protecting groups, and attach the EDTA. Although this methodology has been shown to be successful in other cases,[76,77]

under the conditions of DIEA in DMF employed for reneutralization of the resin for second couplings in our synthesis, the Fmoc group was not stable (data not shown).

To attach EDTA to the side chain amino group of a lysine residue at an internal position along the chain it appeared necessary to do the coupling after the entire chain is assembled and still fully protected. This technique then requires that a protecting group for the lysine side chain be chosen that will be stable to the synthesis conditions and can be removed in a manner which will not affect the peptide-resin bond or the side chain protecting groups. Although a number of orthogonal schemes for main-chain and side-chain protection are established,[76-78] we chose a group which was only minimally different from the one normally used for lysine in our synthesis scheme but offered a new deprotection pathway. In particular, we chose the photolabile 2-nitrobenzyl carbamate,[79-82] which differs from the normal lysine protecting group 2-chlorobenzyl carbamate by a simple substitution of a nitro for a chloro group. This change should not affect the protecting groups performance during peptide synthesis; in fact, it should be a better protecting group against acid due to the more electron-withdrawing nature of the nitro group than the chloro group. Furthermore, the use of a photolabile group in solid phase synthesis has been documented with the use of a 2-nitrobenzyl ester as the peptide resin bond, which could be cleaved with light while leaving all of the side chain protecting groups in place.[83,84]

Synthesis: N^{ε}-2-nitro-CBZ-N^{α}-tBoc-lysine was prepared using standard Schotten-Baumann conditions with 2-nitro-benzyloxycarbonyl chloride[81] and N^{α}-tBoc-lysine in 55% isolated yield. This amino acid was then used to incorporate EDTA to an internal position in the 52 residue DNA binding domain of Hin recombinase as shown in Figure 19. The peptide chain was extended by automated solid-phase synthesis to the desired position. The resin was removed from the machine and N^{ε}-2-nitro-CBZ-N^{α}-tBoc-lysine was coupled as its HOBt ester. The following two amino acids were coupled manually and the resin returned to the machine for the remainder of the synthesis. N^{α}-CBZ protected

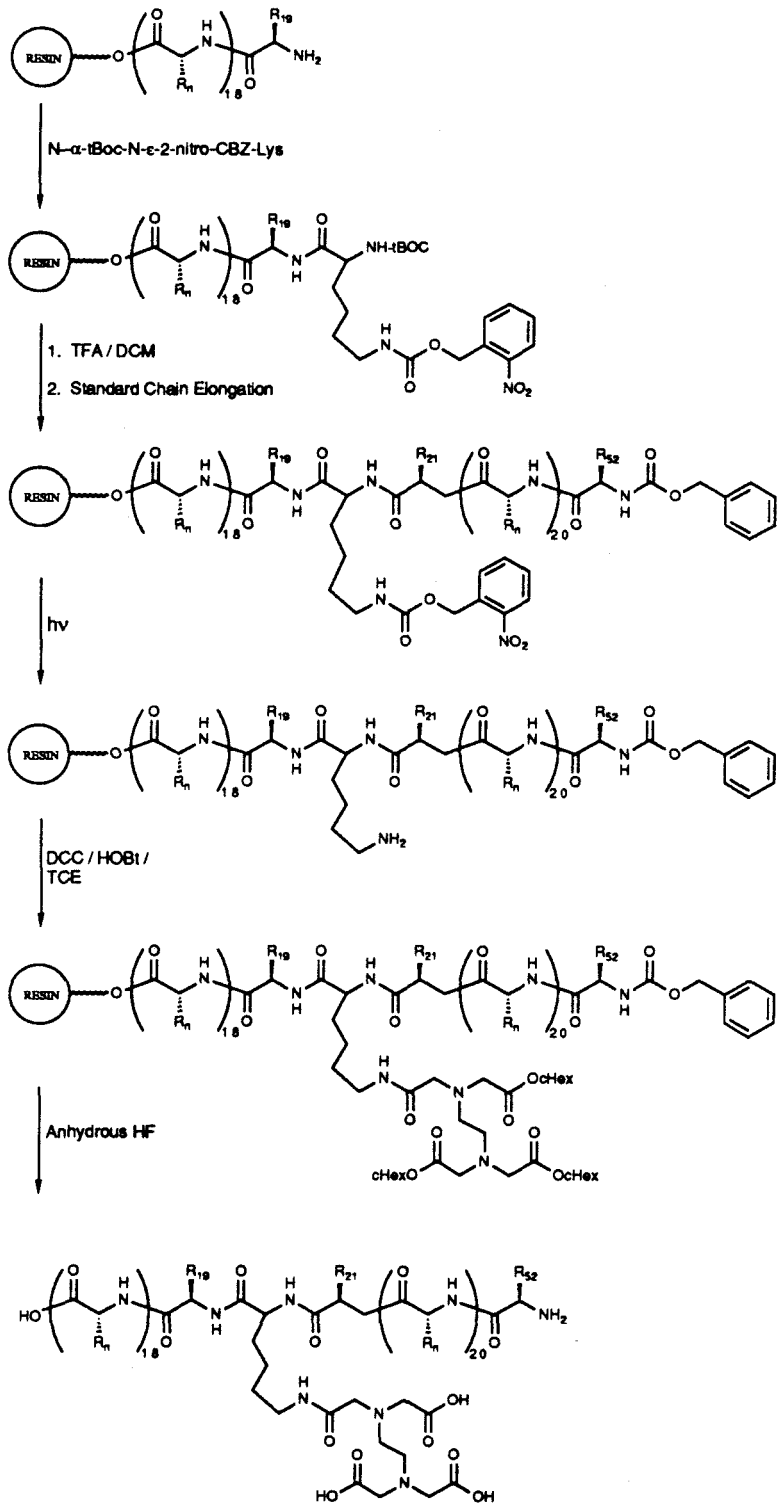


Figure 19 Synthetic scheme for the attachment of TCE to the ε-amino group of Lys^{171} in the middle of the protein sequence of Hin(139-190).

glycine was used for the last amino acid due to the instability of the standard tBoc protecting group to the photolysis conditions (data not shown). The 2-nitro-CBZ protecting group was removed by photolysis with a 1000W Hg/Xe lamp at wavelengths > 340nm in 95% ethanol over 45 minutes with 10 equivalents of semicarbazide•HCl as scavenger.[80] Quantitative ninhydrin analysis indicated that ≈55% of the protecting group had been removed under these conditions. TCE was coupled to the free amino group of the lysine side chain under standard conditions[68] in 98.5% yield after 4 hours. The protein was deprotected by standard HF procedures and purified by reverse phase HPLC.

Although the photo-protecting group was only removed from the resin-bound peptide in ≈55% yield, this reaction has not been optimized. A number of examples in the literature [80,81] indicate that the 2-nitro-CBZ group can be removed from amines in yields approaching 100%. Only a single attempt at photodeprotection of the resin-bound peptide was made and conditions could probably be optimized. This poor deprotection yield did not, however, interfere with the purification of the protein. The strong electron withdrawing nature of the nitro group has apparently made the 2-nitro-CBZ group stable to HF cleavage conditions, and thus the retention time of any protected material has significantly increased (> 3 minutes) by the presence of the aromatic protecting group and this material was easily separated from the desired product (data not shown).

Affinity Cleaving: The DNA cleavage reactions of Hin(139-190, Lys^{171}-EDTA) as well as Hin(139-190, Gly-GABA-EDTA) and Hin (139-184, Lys^{183}-EDTA) for comparison were studied. The proteins were allowed to react at micromolar concentrations (22 °C, pH 7.5, 20 mM NaCl) for 1 hour in the presence of DTT with a ^{32}P-end-labeled restriction fragment containing the Hin-binding sites hixL and secondary Hin.[63] The DNA cleavage products were separated by high resolution polyacrylamide gel electrophoresis and visualized by autoradiogrphy (Figure 20). When the EDTA•Fe is moved from the NH_2 to the COOH terminus of Hin(139-184) it was found that the DNA cleavage locus for Hin(139-184, Lys^{183}-EDTA) is shifted further from the symmetry axis at the binding sites

Figure 20 Autoradiogram of a high-resolution denaturing polyacrylamide gel of affinity cleaving reactions on a ^{32}P-end-labeled fragment (XbaI/EcoRI) from pMFB36.[63] Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in bp) calf thymus DNA, 5 mM DTT, and ≈15,000 cpm of end-labeled DNA in a total volume of 20 μL. Reactions were run for 60 min. at 25 °C and terminated by ethanol precipitation. Cleavage products were analyzed on an 8%, 1:20 cross-linked, 50% urea polyacrylamide gel. Odd and even numbered lanes contain 5'- and 3'-end-labeled DNA, respectively. Lanes 1 and 2 are DNA controls, and lanes 3 and 4 are A-specific sequencing lanes. Lanes 5 and 6 contain 1 μM Hin(139-190, Gly^{139}-GABA-EDTA•Fe), lanes 7 and 8 contain 1 μM Hin(139-184, Lys^{183}-EDTA•Fe), and lanes 9 and 10 contain 5 μM Hin(139-190, Lys^{171}-EDTA•Fe).

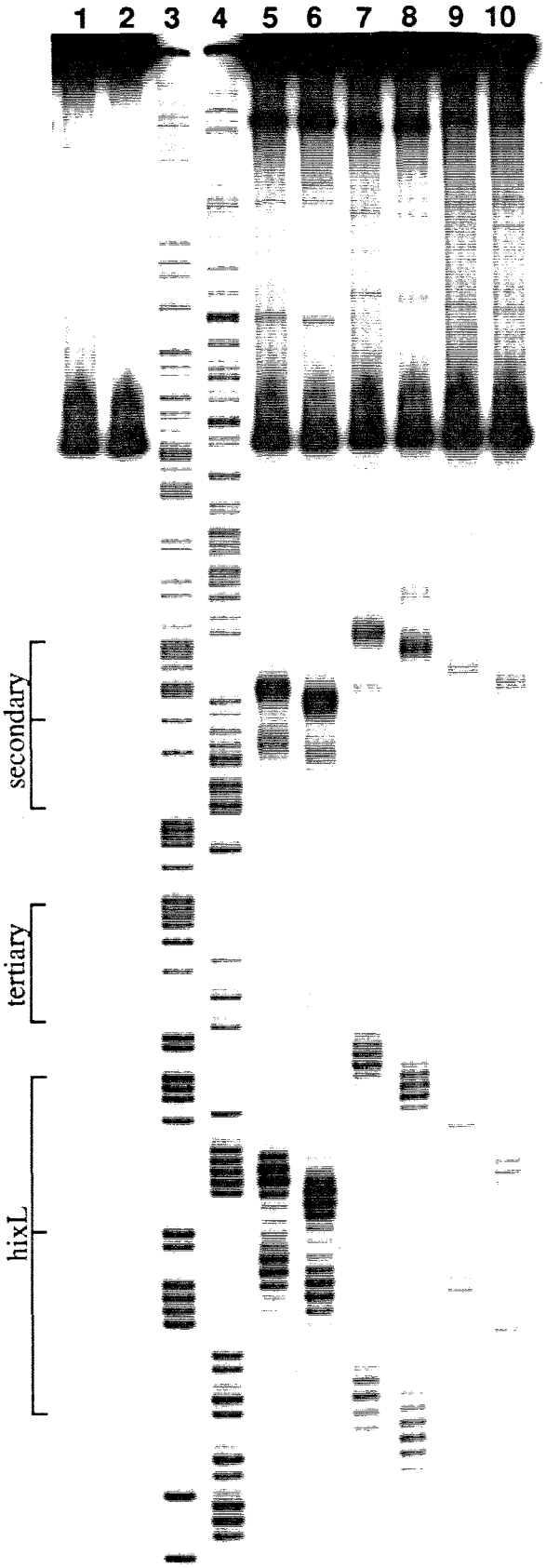
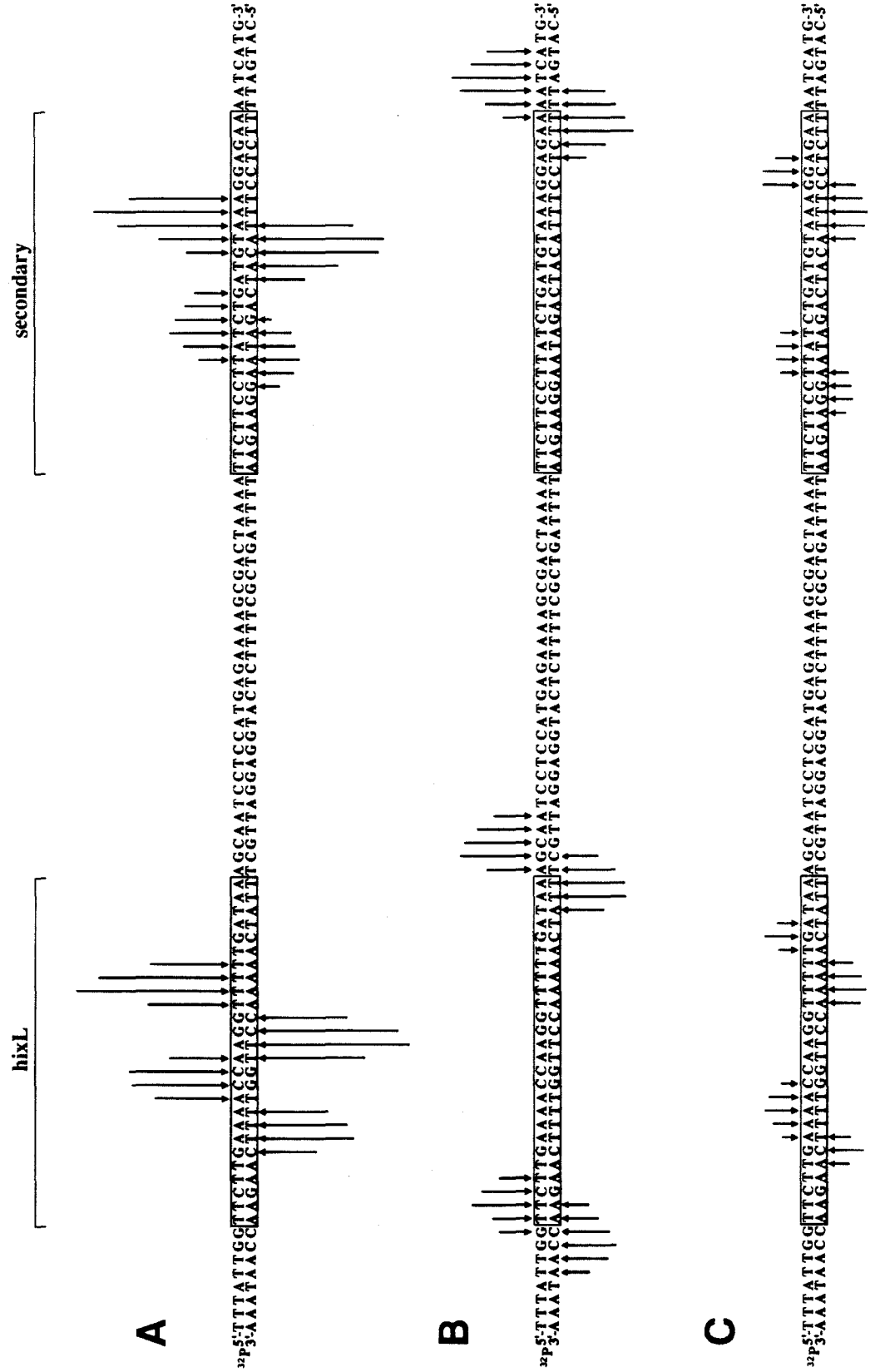

1 2 3 4 5 6 7 8 9 10
secondary
tertiary
hixL

Figure 21 Histograms of the cleavage data from Figure 20. The sequence left to right is from the bottom to the middle of the gel. The left and right boxed sequences represent hixL and the secondary site, respectively. Arrow heights indicate the extent of cleavage at the indicated base. Panels A-C are cleavage patterns for Hin(139-190, Gly^{139}-GABA-EDTA•Fe), Hin(139-184, Lys^{183}-EDTA•Fe), and Hin(139-190, Lys^{171}-EDTA•Fe).


hixL
secondary
A
B
C
32P 5'-TTTATTGGTTCTTGAAAACCAAGGTTTTTGATAAAGCAATCCTCCATGAGAAAAGCGACTAAAATTCTTTCCTTATCTGATGTAAAGGAGAAAATCATG-3'
3'-AAATAACCAAGAACTTTTGGTTCCAAAAACTATTTCGTTAGGAGGTACTCTTTTCGCTGATTTTTAAGAAAGGAATAGACTACATTTCCTCTTTTAGTAC-5'

compared to Hin(139-190, Gly^{139}-GABA-EDTA) (Figure 20, compare lanes 5 and 6 with lanes 7 and 8). When the EDTA is moved to a position between helix 2 and 3, in Hin(139-190, Lys^{171}-EDTA), the DNA cleavage locus moves back in toward the symmetry axis of the site compared to Hin(139-184, Lys^{183}-EDTA) (Figure 20, compare lanes 9 and 10 with lanes 7 and 8). Furthermore, the maximal cleavage patterns observed on opposite strands of the DNA for all three EDTA positions are shifted to the 3' direction.

Histograms of cleavage at the hixL and secondary sites, which consist of imperfectly conserved 12-bp inverted repeats, on a DNA restriction fragment are shown in Figure 21. The specific cleavage pattern observed for Hin(139-190, Lys^{171}-EDTA) is weak compared to the that observed for both Hin(139-190, Gly^{139}-GABA-EDTA) and Hin (139-184, Lys^{183}-EDTA) which are at 5 times higher concentrations. This weak cleavage may indicate that the EDTA•Fe is not in close proximity to the DNA or that the EDTA•Fe is in close proximity to the body of the protein where the hydroxyl radical could possibly also react with the protein backbone.[85-87] The cleavage pattern itself however, is quite localized and specific with a 3' asymmetric shift, indicating that the EDTA•Fe is rigidly held in space in proximity to the minor groove of the DNA.

Models of the DNA binding domain of Hin recombinase which are based on the the helix-turn-helix binding motif and the affinity cleaving data for Hin(139-190, Gly^{139}-GABA-EDTA) and Hin(139-184, Lys^{183}-EDTA) are shown in Figures 22 A and B respectively.[51,58] Figure 22C shows a revised model based on the cleavage data for Hin(139-190, Lys^{171}-EDTA). The data indicate that the loop between helix 2 and 3 must be located farther up the helix toward the symmetry axis of the binding site. This positioning then requires that the recognition helix be oriented in a fashion more perpendicular the the DNA helix axis, in a position which is even more consistent with the cleavage data obtained from Hin(139-184, Lys^{183}-EDTA) (Figure 22B). Although our models are shown in two dimensions, it is possible that the NH_2 terminus of the recognition helix is pointed in toward the groove while the COOH terminus is pointed out

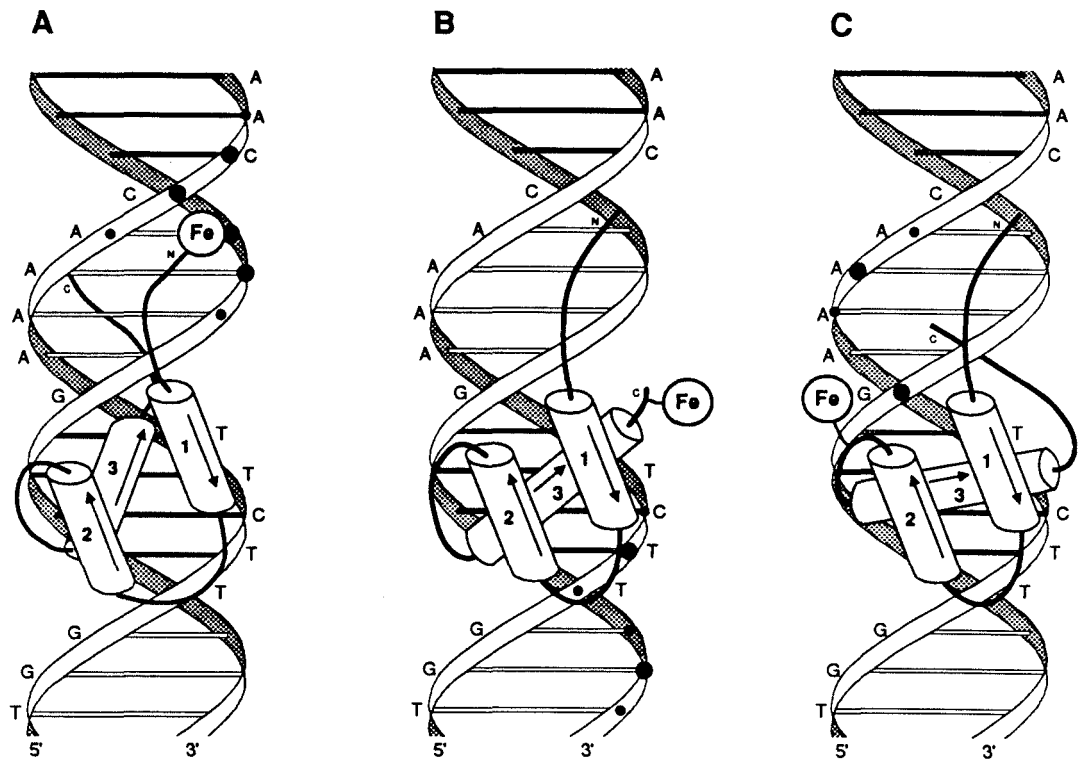


Figure 22 Schematic representations of the DNA binding domain of Hin recombinase bound to the bottom half of the hixL site. The filled circles represent the sites of maximal cleavage produced by A) Hin(139-190, Gly^{139}-GABA-EDTA•Fe), B) Hin(139-184, Lys^{183}-EDTA•Fe), and C) Hin(139-190, Lys^{171}-EDTA•Fe). The Figure shows how the model of the binding domain was revised with the addition of affinity cleaving data from EDTA's at different positions A) NH_2 terminus (1987),[51] B) COOH terminus (1990), [88] and C) between helix 2 and 3 (1991).

away from the DNA in a fashion similar the that observed in the Trp-repressor DNA cocrystal structure.[41]

Conclusions

Previous studies have shown that EDTA can be attached to the amino terminus of a protein creating an affinity cleaving molecule. The work described here extends this methodology to be able to place an EDTA on a lysine side chain at the COOH terminus or at any position along the protein. The techniques developed for EDTA attachment are general and can be applied to any protein which can be synthesized by solid-phase methodologies.

In particular, the DNA binding domain of Hin recombinase has been synthesized with EDTA attached to the NH_2 terminus.[51,69] Based on the affinity cleaving data obtained from the NH_2 terminal EDTA a model was put forward in which the NH_2 terminus binds in the minor groove and a helix-turn-helix structure binds in the major groove. In this work the model has been refined further. By attaching EDTA on a lysine side chain at the COOH terminus of the recognition helix it was possible to determine the orientation of the helix on the DNA. By placing EDTA at the NH_2 terminus of the recognition helix (an internal position of the protein) the orientation has been confirmed and the structure of the complex refined further (Figure 22). The complicated nature of the cleavage patterns (Figure 17), due to the diffusible nature of the oxidizing agent and self footprinting by the protein, will likely not allow the 3-dimensional structure of a protein-DNA complex to be determined, de novo, from affinity cleaving experiments. This technique does, however, offer a straight forward method by which any protein-DNA model may be tested, as has been shown for Hin recombinase[51,69,88] and GCN4.[38] Given the large number of proteins and binding motifs possible, and the difficulties associated with obtaining structural information from crystallography or NMR, affinity cleaving offers a powerful, yet relatively simple, option for examining the structures of protein-DNA complexes.

Experimental

General. Nuclear magnetic resonance spectra were recorded on a JEOL JNM-GX 400 and are reported in parts per million downfield from tetramethylsilane. UV-Vis spectra were recorded on a Beckman model 25 spectrophotometer, a Cary model 219 spectrophotometer, or a Perkin-Elmer lambda 4C spectrophotometer. Laser densitometry of gel autoradiograms was performed on an LKB Ultrascan XL densitometer.

Manual peptide synthesis was carried out in 20-mL vessels fitted with coarse glass frits as described by Kent.[65] Automated syntheses were performed on an ABI 430A synthesizer,[89,90] modified by the removal of the in-line filters to the top and bottom of the

reaction vessel, with a 20 mL Teflon/KelF reaction vessel. The synthetic protocals used were developed at the California Institute of Technology.[65,91-93] Protected amino acid derivatives were purchased from Pininsula Laboratories. tBoc-L-His (DNP protected) was purchased from Fluka, N^{α}-tBoc-Lys and tBoc-L-His (Bom protected) from Bachem Bioscience, and N^{α}-tBoc-N^{ε}-Fmoc-L-Lys, from chemical Dynamics Corp. Benzhydrylamine resin (BHA) was purchased from U.S. Biochemical Corp. Asn-(phenylacetamido)methyl (PAM) resin, dimethylformamide (DMF), diisopropylethylamine (DIEA), dicyclohexylcarbodiimide (DCC) in dichloromethane, N-hydroxybenzotriazole (HOBt) in DMF, and trifluoroacetic acid (TFA) were obtained from Applied Biosystems. Dichloromethane (DCM) and methanol (HPLC grade) were purchased from Mallinkrodt, p-cresol and p-thiocresol were from aldrich, and diethyl ether was from Baker.

Doubly distilled water was used for all aqueous reactions and dilutions. Plasmid pMFB36 DNA[63] was isolated from *E. coli* and purified by CsCl centrifugation. Calf thymus (CT) DNA was purchased from Sigma and was sonicated, deproteinized, and dialyzed. Enzymes were purchased from Boehringer Mannheim or New England Biolabs.

Peptide Synthesis. N^{α}-tBoc-L-amino acids were used with the following side chain protecting groups: Arg(Tos), Asp(OBzl), Glu(OBzl), His(DNP), Lys(Cl-Z), Ser(Bzl), Thr(Bzl), and Tyr(Br-Z). Manual assembly of the protected peptide on the solid support was carried out at 25 °C via a three-step reaction cycle. First the tBoc protecting group was removed from the α-amino group of the resin-bound amino acid by using TFA (65% TFA in DCM for 1 and 15 min). The deprotected peptide resin was then neutralized with 10% DIEA in DCM (2 X 1 min). Amino acids (except Asn, Gln, and Arg) were coupled to the free α-amino group as the symmetric anhydrides. For the first coupling, the symmetric anhydride was formed in situ in DCM. DMF was added after a sufficient activation time (10 min), and the reaction was allowed to proceed for a total of 30 min. Coupling yields were determined by quantitative ninhydrin monitoring[66] with acceptable values being ≥99.7% in the beginning of the synthesis and falling off gradually to 99% near the end. If

a second coupling was necessary, the resin was neutralized with 10% DIEA in DMF. The symmetric anhydride was formed outside the reaction vessel in an absolute minimum of DCM. The solution was filtered into the vessel to remove the dicyclohexylurea, topped off with DMF, and allowed to react for 1-2 hours. If the yield was not acceptable at this point, a third coupling was performed by the same protocol as the second coupling. Asn, Gln, and Arg were coupled as the HOBt esters in DMF. The HOBt ester was formed in situ by combining the amino acid, DCC, and HOBt in DMF in the reaction vessel. Reaction times were longer due to the slower nature of the coupling reaction, typically 1-2 hours. When second couplings were required for Asn and Gln, the same HOBt procedure was followed after neutralization with 10% DIEA in DMF. Second couplings for Arg were performed by the preformed symmetric anhydride procedure above.

Automated syntheses were carried out with modified cycles that are similar to the manual procedures. Double couplings were performed for every amino acid. tBoc protecting groups were removed from the α-amino group of the resin-bound amino acid by using 100% TFA. The deprotected peptide resin was neutralized with 10% DIEA in DMF. Amino acids (except Asn, Gln, and Arg) were coupled to the free α-amino group as the symmetric anhydrides. In the first coupling, the symmetric anhydride was formed in an activating vessel with DCC in DCM. The dicyclohexylurea was removed by filtering the solution into a concentrating vessel, where the DCM was removed and replaced with DMF. The solution was then transferred to the reaction vessel, where the resin had previously been deprotected and neutralized. After the first coupling, the resin was neutralized with DIEA in DMF for the second coupling. The symmetric anhydride was formed in the activation in DCM, filtered, and transferred to the reaction vessel. DMF was added at the midpoint of the reaction cycle. Yields for double couplings were determined by quantitative ninhydrin monitoring.[66]

BEG and TCE Coupling to the NH_2-terminus of Resin Bound Peptides. A typical procedure for the coupling of BEG[68] to the NH_2-terminus of resin bound peptide,

corresponding to the synthesis of EDTA-Hin(139-190)[51,68] was as follows. A ≈100mg sample of peptide-resin (*ca.* 100 μmol/g, total peptide *ca.* 10 μmol) was placed in a 12 X 80 mm reactor, and the resin was swollen in DCM for 15 minutes. The resin was washed with DCM (5X), the terminal Boc protecting group was removed, and the resin neutralized using standard procedures. BEG (125mg, 194 μmol) and HOBt (50mg, 330 μmol) were dissolved in DMF (2 mL) and DCC (42 mg, 203 μmol) was added. The solution was stirred 30 min at room temperature and added to the peptide-resin along with sufficient DMF to fill the reactor two-thirds full. Ninhydrin analysis indicated ≥98% reaction yield after 2-3 hours. The resin was washed with DMF (4X) and DCM (4X). TCE was coupled to the NH_2-terminus of resin-bound peptide using this same procedure.

TCE Coupling to the ε-NH_2 Side Chain of a Lysine Residue at the COOH Terminus of Resin Bound Peptides. N^{α}-tBoc-N^{ε}-Fmoc-L-lysine was coupled to the growing peptide as an HOBt ester using the standard procedure. The resin was washed with DCM (5X) followed by DMF (5X). The N^{ε}-Fmoc group was selectively removed using 20% piperidine in DMF (a 1 min reaction step followed by a 10 min step) and the resin was washed with DMF (5X). TCE[68] (1.08g, 2 mmol, 4 equiv. based on a 0.5 mmol synthesis) and HOBt (.46 g, 3.4 mmol, 6.8 equiv.) were dissolved in DMF (2 mL) and DCC (0.41 g, 2 mmol, 4 equiv.) was added. The solution was stirred 30 min at room temperature and added to the peptide resin along with sufficient DMF to fill the reactor two-thirds full. Ninhydrin analysis indicated 99.9% reaction after 1 hour. The peptide resin was washed with DMF (5X) and DCM (5X). The N^{α}-tBoc was removed as usual with TFA and the synthesis continued with the standard protocals.

TCE Coupling to the ε NH_2 Side Chain of a Lysine Residue at an Internal Position along the Peptide Chain. **N^{α}-tBoc-N^{ε}-2-nitorCBZ-L-lysine.** N^{α}-tBoc-L-lysine (5g, 20 mmol) and sodium bicarbonate (4.2 g, 50 mmol) were dissolved in water:dioxane (2:1, 100mL). A solution of 2-nitro-benzyloxycarbonyl chloride[81] in dioxane (20 mL) was added in 1 mL portions over 30 minutes to the stirrred solution. The reaction was stopped

after 1 hour and the dioxane removed by rotory evaporation. 100 mL ethyl acetate was added and the solution acidified to pH = 3 with KH_2SO_4. The ethyl acetate layer was separated and the aqueous layer washed with an additional 2 X 100 mL ethyl acetate. The organic layers were combined, dried with sodium sulfate, and the solvent removed by rotory evaporation to yield 5.5 g of crude product which by TLC contained some of the 2-nitro-benzylalcohol. 2.5 g of the material was purified by flash chomatography (1% CH_3OH in DCM followed by 10% CH_3OH in DCM) to afford 2.1 g (55% overall yield) of the desired product. ^1H NMR (DC_3OD) δ8.10 (m, 1H), 7.71 (m, 2H), 7.54 (m, 1H), 5.44 (s, 2H), 3.93 (m, 1H), 3.12 (m, 2H), 1.80 (m, 1H), 1.64 (m, 1H), 1.52 (m, 2H), 1.43 (broad s, 11H) ppm. HRMS (FAB) for $C_{19}H_{26}N_3O_8$, calc. 424.1720, found 424.1711.

The procedure for coupling N^{α}-tBoc-N^{ε}-2-nitro-CBZ-L-lysine at position 171 of the DNA binding domain of Hin recombinase to yield Hin(139-190, Lys^{171}-EDTA) was as follows. Residues 190-172 were coupled by standard automated techniques using tBoc-L-His(Bom) instead of tBoc-L-His(DNP). The resin was removed from the machine and N^{α}-tBoc-N^{ε}-2-nitro-CBZ-L-lysine coupled to the peptide as an HOBt ester using the standard procedure. The following two residues were coupled by manual procedures and the resin returned to the machine for the synthesis of residues 168-138. At this point the resin was again removed from the machine and the final residue, CBZ-Gly was coupled via the symmetric anhydride procedure. A ≈100 mg (ca.10 μmol) resin sample was photodeprotected using a 1000W Hg/Xe lamp with a $\lambda > 339.1$ nm cut-off filter in the presence of semicarbazide hydrochloride (13 mg, 100 μmol) as scabanger[80] in ≈3 mL 95% ethanol. After 45 minutes of photolysis ninhydrin analysis indicated ≈50% deprotection and the photolysis was stopped. The resin was placed in a 12 X 80 mm reactor and washed with DMF (5X). TCE (104 mg, 194 μmol) and HOBt (50 mg, 330 μmol) were dissolved in DMF (2 mL) and DCC (42 mg, 203 μmol) was added. The solution was stirred 30 min at room temperature and added to the peptide-resin along with sufficient

DMF to fill the reactor two-thirds full. Ninhydrin analysis indicated 98.5% reaction yield after 4 hours. The resin was washed with DMF (4X) and DCM (4X).

Deprotection and Purification. When the histidine protecting group, dinitrophenyl (DNP), was present it was removed by thiolysis at 25 °C using 20% 2-mercaptoethanol and 10% DIEA in DMF; this treatment was repeated twice for 30 min each time. After removal of the N^{α}-tBoc group with TFA and drying of the resin, all other side-chain protecting groups as well as the peptide-resin bond were cleaved by using anhydrous HF, in the presence of p-cresol and p-thiocresol as scavangers, for 60 min at 0 °C.[75] The HF was removed under vacuum. The crude protein was precipitated with diethyl ether, collected on a fritted funnel, dissolved in water, and washed through, leaving the resin on the frit. A small sample was then removed, filtered, and subjected to analytical HPLC (Vydac 25 cm X 4.6 mm C_4 column, 0-60% acetonitrile/0.1% TFA over 60 min). The remaining solution was frozen and lyophilized. Risidual DNP groups were removed from the crude protein by treatment in 4 M guanidine hydrochloride, 50 mM tris, pH 8.5, and 20% 2-mercaptoehtanol for 1 hour at 50 °C.[65] This solution was injected directly onto a semipreperative C_4 HPLC column (25 cm X 1 cm) and run in H_2O/0.1% TFA until the guanidine and 2-mercaptoethanol had eluted. A gradient of 0-60% acetonitrile/0.1% TFA was run over 240 min, and fractions were collected. Fractions were checked by analytical HPLC, and those containing the desired peak were pooled and lyophilized. The protein was packaged into 5 X 10^{-9} mol samples on the basis of OD_{275} (ε = 2800 for 2 Tyr).

Labeling of the XbaI/EcorRI Restriction Fragment from pMFB36. The plasmid pMFB36[63] was linearized by cleavage with restriction endonuclease XbaI. Labeling at the 3' end was accomplished with [α-^{32}P]dATP and the Klenow fragment of DNA polymerase I. The 5' end was labeled with ^{32}P by dephosphorylation with calf alkaline phosphatase (CAP) followed by treatment with [γ-^{32}P]ATP and T4 polynucleotide kinase. Cleavage with restriction endonuclease EcoRI yielded 3'- and 5'- end-labeled fragments 557 bp,

which were isolated by polyacrylamide gel electrophoresis. A specific sequencing reactions were then performed on the fragments.[94]

Affinity Cleaving Reactions. DNA cleaving reactions were done in a total volume of 20 μL. Final concentrations were 20 mM phosphate, pH 7.5, 20 mM NaCl, 100 μM (in base pairs) CT DNA, ≈15,000 cpm of ^{32}P-end-labeled restriction fragment, 5 mM dithiothreitol (DTT), and 5 μM Fe•EDTA-protein. The proteins were allowed to equilibrate with the DNA for 10 minutes at 25 °C, and cleavage was then initiated by the addition of DTT and allowed to proceed for 60 min at 25 °C. The reactions were terminated by ethanol precipitation, dried, and resuspended in 5 μL of 100 mM Tris-borate-EDTA and 80% formamide solution. The ^{32}P-labeled DNA products were analyzed by denaturing polyacrylamide gel electrophoresis followed by autoradiography. Densitometric analysis of the gel autoradiogram and comparison of individual lanes with sequence marker lanes allowed assignment of DNA cleavage to nucleotide resolution.

References

1. Ollis, D.L.; White, S.W., *Chem. Rev*, **87**, 981 (1987).
2. Helene, C.; Lancelot, G., *Prog. Biophy. Molec. Biol.*, **39**, 1 (1982).
3. Seeman, N.C.; Rosenberg, J.; Rich, A., *Proc. Nat. Acad. Sci., USA*, **73**, 804 (1976).
4. Von Hippel, P.H.; Berg, O.G., *Proc. Nat. Acad. Sci., USA*, **83**, 1608 (1986).
5. Grayer, G.D.; McPherson, A., *J. Mol. Biol.*, **169**, 565 (1983).
6. Brayer, G.D.; McPherson, A., *Biochemistry*, **23**, 340 (1984).
7. Finch, J.J.; Brown, R.S.; Rhodes, D.; Richmond, T.; Rushton, B; Lutter, L.C.; Klug, A., *J. Mol. Biol.*, **145**, 757 (1981).
8. Richmond, T.J.; Finch, J.T.; Rushton, B.; Rhodes, D.; Klug, A., *Nature*, **311**, 532 (1984).
9. Burlingame, R.; Love, W.; Wang, B.; Hanlin, R.; Xuong, N.; Moudranakis, E., *Science*, **228**, 546 (1985).
10. Tanaka, I.; Appelt, K.; Dijik, J.; White, S.W.; Wilson, K.S., *Nature*, **310**, 376 (1984).
11. Suck, D.; Oefiner, C.; Kabsch, W., *EMBO J.*, **3**, 2423 (1984).

12. Brick, P.; Ollis, D.L.; Steitz, T., *J. Mol Biol.*, **166**, 453 (1983).
13. Ollis, D.L.; Brick, P.; Hamlin, R.; Xuong, N.G.; Steitz, T., *Nature*, **313**, 762 (1985).
14. Ohlendor, D.H.; Anderson, W.F.; Takeda, Y.; Matthews, B.W., *Nature*, **290**, 754 (1981).
15. Matthews, B.W.; Ohlendorf, D.H.; Anderson, W.F.; Fisher, R.G.; Takeda, Y., *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 427 (1982).
16. Matthews, B.W.; Anderson, W.F.; Fisher, R.G.; Takeda, Y.; Ohlendorf, D.H., *Nature*, **298**, 718 (1982).
17. Pabo, C.O.; Lewis, M., *Nature*, **298**, 443 (1982).
18. Lewis, M.; Jeffrey, A.; Wang, J.; Ladner, R.; Ptashne, M.; Pabo, C.O., *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 435 (1982).
19. Steitz, T.A.; Weber, I.T.; Matthews, J.B., *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 419 (1982).
20. McKay, D.B.; Weber, I.T.; Steitz, T.A., *J. Biol. Chem.*, **257**, 9815 (1982).
21. McKay, D.B.; Steitz, T.A., *Nature* **290**, 744 (1981).
22. Ohlendorf, R.H.; Matthews, B.W., *Ann. Rev. Biophys. Bioeng.*, **12**, 259 (1983).
23. Sauer, T.T.; Yocum, R.R.; Doolittle, R.F.; Lewis, M.; Pabo, C.O., *Nature*, **298**, 447 (1982).
24. Pabo, C.O.; Sauer, R.T., *Ann. Rev. Biochem.*, **53**, 293 (1984).
25. McClarin, J.A.; Frederick, C.A.; Wang, B.C.; Greene, P.; Boyer, H.; Grable, J.; Rosenberg, J., *Science,* **234**, 1526 (1986).
26. Frederick, C.A., *et al.*, *Nature*, **309**, 327 (1984).
27. Evans, R.M.; Hollenberg, S.M., *Cell*, **52**, 1 (1988).
28. Klug, A.; Rhodes, D., *Trends Biochem. Sci.*, **12**, 464 (1987).
29. Evens, R.M., *Science*, **240**, 889 (1988).
30. Beato, M., *Cell*, **56**, 335 (1989).
31. Parraga, G.; Horvath, S.J.; Eisen, A.; Tayler, W.E.; Hood. L.E.; Young, E.T.; Klevit, R.E., *Science*, **241**, 1489 (1988).
32. Lee, M.S.; Gippert, G.P.; Soman, K.V.; Case, D.A.; Write, P.E., *Science*, **245**, 635 (1989).
33. Berg, J.M., *Proc. Nat. Acad, Sci., USA*, **85**, 99 (1988).
34. Brown, R.S.; Argos, P., *Nature*, **324**, 215 (1986).
35. Gibson, T.J.; Tastama, J.P.M.; Brown, R.S.; Argos, P., *Protein Eng.*, **2**, 209 (1988).
36. Landschulz, W.H.; Johnson, P.F.; Mcknight, S.L., *Science*, **240**, 1759 (1988).

37. Vinson, C.R.; Sigler, P.B.; Mcknight, S.L., *Science*, **246**, 911 (1989).
38. Oakley, M.G.; Dervan, P.B.; *Science*, **248**, 847 (1990).
39. Jordan, S.R.; Pabo, C.O., *Science*, **242**, 893 (1988).
40. Aggarwal, A.K.; Rodgers, D.W.; Drottar, M.; Ptashne, M.; Harrison, S.C., *Science*, **242**, 899 (1988).
41. Otwinowski, Z.; Schevitz, R.W.; Zhang, R.-G.; Lawson, C.L.; Joachimiak, A.; Marmorstein, R.Q.; Luisi, B.F.; Sigler, P.B., *Nature*, **335**, 321 (1988).
42. Harrison, S.C.; Aggarwal, A.K., *Ann. Rev. Biochem.*, **59**, 933 (1990).
43. Galas, D.J.; Schmitz, A., *Nucleic Acids Res.*, **5**, 3157 (1978).
44. Dervan, P.B., *Science*, **232**, 464 (1986).
45. Tullius, T.D.; Combroski, B.A.; Churchill, M.E.A.; Kam, L., *Methods in Enz.*, **155**, 537 (1987).
46. Hertzberg, R.P.; Dervan, P.B., *J. Am. Chem. Soc.*, **104**, 313 (1982).
47. Hertzberg, R.P.; Dervan, P.B., *Biochemistry*, **23**, 3934 (1984).
48. Schultz, P.G.; Taylor, J.S.; Dervan, P.B., *J. Am. Chem. Soc.*, **104**, 6861 (1982).
49. Taylor, J.S.; Schultz, P.G.; Dervan, P.B., *Tetrahedron*, **40**, 457 (1984).
50. Dryer, J.B.; Dervan, P.B., *Proc. Nat. Acad. Sci., USA*, **82**, 968 (1985).
51. Sluka, J.P.; Horvath, S.J.; Bruist, M.F.; Simon, M.I.; Dervan, P.B., *Science*, **238**, 1129 (1987).
52. Moser, H.; Dervan, P.B., *Science*, **238**, 645 (1987).
53. Tullius, T.D.; Dombrowski, B.A., *Science*, **230**, 679 (1985).
54. Griffin, L.C.; Dervan, P.B., *Science*, **245**, 976 (1989).
55. Youngquist, R.S.; Dervan, P.B., *Proc. Nat. Acad. Sci., USA*, **82**, 2565 (1985).
56. Wade, W.S.; Dervan, P.B., *J. Am. Chem. Soc*, **109**, 1574 (1987).
57. Griffin, J.H.; Dervan, P.B., *J. Am. Chem. Soc.*, **109**, 6840 (1987).
58. Strobel, S.A.; Moser, H.E.; Dervan, P.B., *J. Am. Chem. Soc.*, **110**, 7927 (1988).
59. Povsic, T.J.; Dervan, P.B., *J. Am. Chem. Soc.*, **111**, 3059 (1989).
60. Zieg, J.; Silverman, M.; Hilmen, M.; Simon, M., *Science*, **196**, 170 (1977).
61. Johnson, R.C.; Bruist, M.F.; Glaccum, M.B.; Simon, M.I., *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 751 (1984).
62. Glasgow, A.C.; Bruist, M.F.; Simon, M.I., *J. Biol. Chem.*, **264**, 10072 (1989).
63. Bruist, M.F.; Horvath, S.J.; Hood, L.E.; Steitz, T.A.; Simon, M.I., *Science*, 235, 777 (1987).
64. Merrifield, R.B., *Adv. Enzymol.*, **32**, 221 (1969).
65. Kent, S.B.H., *Ann. Rev. Biochem.*, **57**, 957 (1988).

66. Sarin, V.K.; Kent, S.B.H.; Tam, J.P.; Merridield, R.B., *Anal. Biochem.*, **117**, 147 (1981).
67. Tam, J.P.; Wong, T.W.; Riemen, M.W.; Tjoeny, F.S.; Merrifield, R.B., *Tetrahedron Lett.*, 4033 (1979).
68. Sluka, J.P.; Griffin, J.H.; Mack, D.P.; Dervan, P.B., *J. Am. Chem. Soc.*, in press.
69. Sluka, J.P.; Horvath, S.J.; Glasgow, A.C.; Simon, M.I.; Dervan, P.B., *Biochemistry*, **29**, 6551 (1990).
70. Boelenes, R.; Scheek, R.M.; van Boon, J.H.; Kaptein, R., *J. Mol. Biol.*, **193**, 213 (1987).
71. Sluka, J.P., Ph.D. thesis, California Institute of Technology, Pasadena, CA (1988).
72. Plasterk, R.H.A.; Brinkman, A.; van de Putte, P., *Proc. Nat. Acad. Sci., USA*, **80**, 5355 (1983).
73. Hiestand-Nauer, R.; Iida, S., *EMBO J.*, **2**, 1733 (1983).
74. Plasterk, R.H.; van de Putte, P., *EMBO J.*, **4**, 237 (1985).
75. Stewart, J.M.; Young, J.D., *Solid Phase Peptide Synthesis*, 2nd ed., Pierce Chemical Co., Rockford, Il (1984).
76. Mutter, M.; Vuilleumier, S., *Angew. Chem. Int. Ed. Eng.*, **28**, 535 (1989).
77. Hahn, K.W.; Klis, W.A.; Stewart, J.M., *Science*, **248**, 1544 (1990).
78. Barany, G.; Albericia, F., *J. Am. Chem. Soc.*, **107**, 4936 (1985).
79. Barltrop, J.A.; Plant, P.J.; Schofield, P., *Chem Comm.*, 822 (1966).
80. Patchornick, A.; Amit, B.; Woodward, R.B., *J. Am. Chem. Soc.*, **92**, 6333 (1970).
81. Amit, B.; Zehavi, U.; Patchornik, A., *J. Org. Chem.*,**39**, 192 (1974).
82. Pillai, V.N.R. *Synthesis*, 1 (1980).
83. Rich, D.H.; Gurwara, S.K., *J. Am. Chem. Soc.*, **97**, 1575 (1975).
84. Rich, D.H.; Gurwara, S.K., *Tetrahedron Lett.*, **301** (1975).
85. Rana, T.M.; Meares, C.F., *J. Am. Chem. Soc.*, **112**, 2457 (1990).
86. Shepartz, A.; Cuenond, B., *J. Am. Chem. Soc.*, **112**, 3247 (1990).
87. Hoyer, D.; Cho, H.; Schultz, P.G., *J. Am. Chem. Soc.*, **112**, 3249 (1990).
88. Mack, D.P.; Sluka, J.P.; Shin, J.A.; Griffin, J.H.; Simon, M.I.; Dervan, P.B., *Biochemistry*, **29**, 6561 (1990).
89. Kent, S.B.H.; Hood, L.E.; Belian, H.; Meister, S.; Geiser, T., in *Peptides* 1984, Proceedings of the 18th European Peptide Symposium, Djuroenaeset, Sweden, 1984 (Ragnarsson, U., ed.) pp. 185-188, Almqvist and Wiksell, Stockholm, Sweden (1984).
90. Kent, S.B.H.; Hood, L.E.; Belian, H.; Bridgham, J.; Marriott, M.; Meister, S.; Geiser, T., in *Peptide Chemistry* 1984, Proceedings of the Japanese Peptide

Symposium (Isumiya, N., Ed.) pp. 167-170, Protein Research Foundation, Osaka, Japan (1985).

91. Kent, S.B.H.; Clark-Lewis, I. in *Sythetic Peptides in Biology and Medicine* (Alitalo, K., Partanen, P., and Vaheri, A. eds.) pp. 29-57, Elsevier, Amsterdam, The Netherlands (1985).
92. Clark-Lewis, I.; Aebersold, R.; Ziltener, H.; Schrader, J.; Hod, L.E.; Kent, S.B.H., *Science*, **231**, 134 (1986).
93. Kent, S.B.H.; Parker, K.F.; Schiller, D.L.; Wood, D.D.-L.; Clark-Lewis, I.; Chait, B.T., in *Peptides: Chemistry and Biology*, Proceedings of the Tenth American Peptide Symposium, St. Louis, MO, 1987 (Marshall, R.R., ed.) pp. 173-178, ESCOM, Leiden, The Netherlands (1988).
94. Iverson, B.I.; Dervan, P.B., *Nucleic Acids Res.*, **15**, 7823 (1987).

Part II:

Design of a DNA Cleaving Protein Consisting of Only Natural Amino Acids

As was discussed in part I, the 52 residue peptide identical with the carboxy terminal domain of Hin recombinase (190 amino acids) has been shown to bind to Hin recombination half sites (13 bp) and to inhibit Hin activity.[1] Previous work in the group has described the conversion of this sequence-specific DNA-binding protein, Hin(139-190), into a sequence-specific DNA-cleaving protein by covalent attachment of an iron chelator, ethylenediaminetetraacetic acid (EDTA), to the amino-terminus has been described.[2-4] In the presence of Fe(II) and reducing agent, EDTA-Hin(139-190) oxidatively cleaves DNA at Hin binding sites, revealing the base position and minor groove location of the amino terminus of the peptide when bound to DNA. This work explores the issue of whether the unnatural amino acid, EDTA, could be replaced by a sequence of α-amino acids that bind transition metals capable of facilitating oxidative cleavage of DNA. This chapter describes studies of DNA cleavage employing the metal binding tripeptide GlyGlyHis with Cu(II) and Ni(II) when attached to the DNA binding domain of Hin recombinase.

Gly-Gly-His. The tripeptide, H-glyglyhis-OH (GGH), is a consensus sequence for the copper-binding domain of serum albumin across many species.[5,6] The tripeptide GGH binds Cu(II) in a 1:1 complex over the pH range 6.5-11 with a dissociation constant of 1.2 X 10^{-16} M^{-1}.[6] A crystal structure of Cu(II)•GGHa reveals square-planar complexation of the Cu(II) by an imidazole nitrogen, two deprotonated peptide nitrogens, and the terminal amino group (Figure 1).[7] It is further known that free Cu(II) at mM concentrations as well as Cu(II) in the presence of GGH[8] will degrade DNA randomly in the presence of reducing agents such as sodium ascorbate.

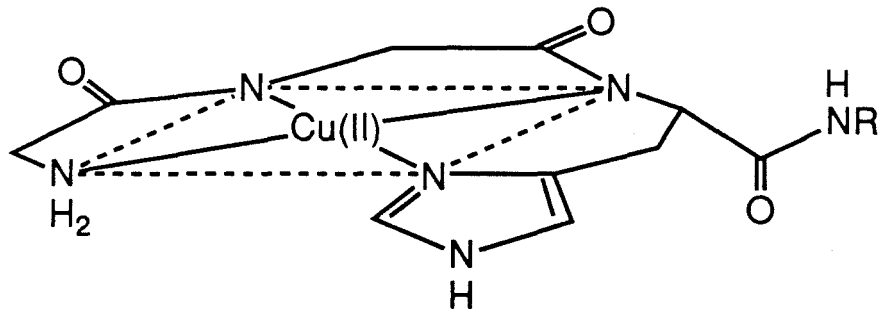


Figure 1 Model for GGH ligand bound to Cu(II) in a square planar complex.

As a first attempt at the design of a sequence-specific DNA-cleaving protein consisting wholly of naturally occurring α-amino acids, the tripeptide GGH was attached to the amino terminus of the DNA binding domain of Hin recombinase (residues139-190). This new 55 residue protein GGH(Hin 139-190) should have two structural domains each with distinct functions: sequence-specific recognition and cleavage of double helical DNA (Figure 2).

Synthesis. The 55-residue protein was synthesized by stepwise solid-phase methods with optimized manual protocols[9] on phenylacetamidomethyl (PAM) resin solid support with Boc protected amino acids. Each coupling cycle was monitored by quantitative ninhydrin analysis,[10] and couplings were repeated until the yield was maximal for each step (99.7% early in the synthesis decreasing to 99.0% at the end). After deprotection of the peptide-resin, the crude synthetic protein was purified by reverse phase preparative HPLC to afford the synthetic protein, GGH(Hin 139-190). Peptide sequence analysis by Edman degradation confirmed the identity of the purified synthetic protein.[11]

Cleavage by Cu(II)•GGH(Hin 139-190). MPE•Fe(II) footprinting (cleavage protection)[12] revealed that the synthetic GGH(Hin139-190) at 0.5 μM concentration (pH 7.5, 25 °C, 20 mM NaCl) binds all four Hin half sites, each 13 base pairs in length (Figures 3 and 4). The 55-residue protein alone or in the presence of Cu(II) does not cleave the DNA. However, GGH(Hin 139-190) at 5 μM concentration (pH 7.5, 25°C,

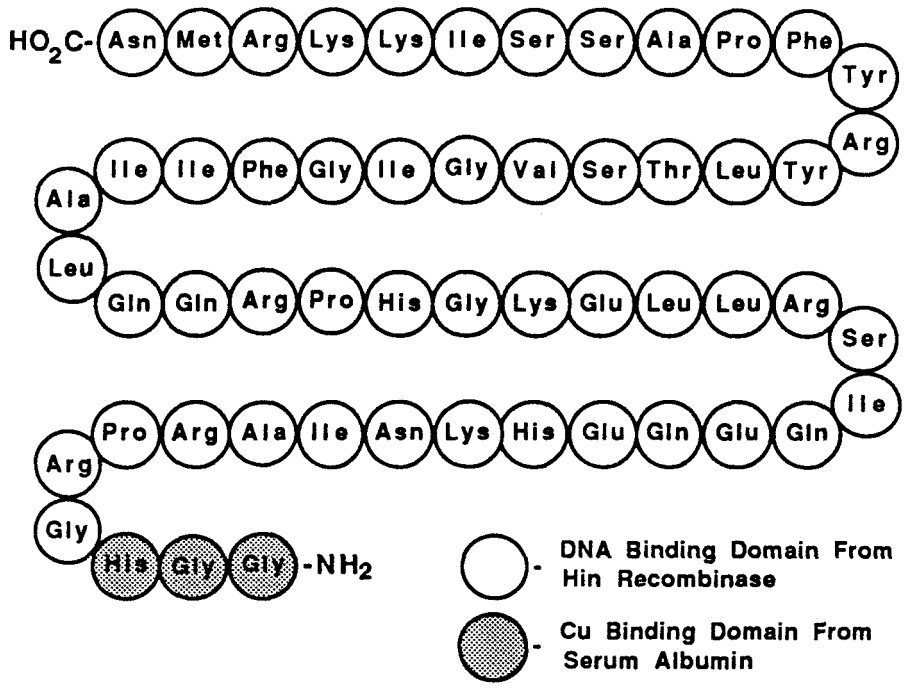


Figure 2 Sequence of the designed protein combining the DNA-binding domain of Hin recombinase (residues 139-190) with the copper binding domain (GGH) from serum albumin.

20 mM NaCl) in the presence of less than 1 equivalent of Cu(II) (2.5 μM) followed by addition of sodium ascorbate (1 mM) and hydrogen peroxide (1 mM), cleaved the DNA predominantly at one of the Hin half sites (Figures 3 and 4). Controls show that *both* sodium ascorbate and hydrogen peroxide are necessary for cleavage (Figure 3). The major cleavage pattern at the upper half of the secondary Hin site is located toward the center of the 13 bp site and covers two base pairs with maximal cleavage asymmetric to the 3' side. This is consistent with the Cu•GGH complex at the amino terminus of the protein being located toward the symmetry axis of the Hin site in the minor groove (Figure 5).

End Product Analysis. In order to determine the nature of the DNA cleavage reaction it was necessary to establish the identity of the DNA termini produced by the cleavage reaction. Since the electrophoretic mobility of a DNA fragment depends on its size, shape,

Figure 3 Autoradiogram of high resolution denaturing gel of MPE•Fe footprinting of GGH(Hin 139-190) and Cu•GGH(H 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 L. Final concentrations in control and footprinting lanes were MPE•Fe(II) at 10 μM, DTT at 5 mM and GGH(Hin 139-190) at 0.5 μM. Final concentrations in control and cleavage lanes were GGH(Hin 139-190) and Cu•GGH(Hin 139-190) at 5 μM; sodium ascorbate and hydrogen peroxide are at 1 mM. Reactions were run at 25 °C for 90 min for Cu•GGH(Hin 139-190) and controls and 12 min for MPE•Fe footprinting and controls, respectively. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea denaturing polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes are DNA labeled at the 5' and 3' end with ^{32}P, respectively. Lanes 1 and 2 are Maxam-Gilbert chemical sequencing G lanes. Lanes 3 and 4 are MPE•Fe controls. Lanes 5 and 6 are MPE•Fe footprints in the presence of GGH[a]Hin 139-190) (0.5 μM). Lanes 7 and 8 contain $CuCl_2$ (2.5 μM) and GGH(Hin 139-190) (5 μM) followed by sodium ascorbate (1 mM) and hydrogen peroxide (1 mM). Lanes 9 and 10 contain GGH(Hin 139-190). Lanes 11 and 12 and 13 and 14 contain GGH(Hin 139-190) (5 μM), sodium ascorbate (1 mM), and hydrogen peroxide (1 mM) without and with EDTA (1 mM), respectively. Lanes 15 and 16 contain $CuCl_2$ (2.5 μM), GGH(Hin 139-190) (5 μM), and sodium ascorbate (1 mM). Lanes 17 and 18 contain $CuCl_2$ (2.5 μM), GGH(Hin 139-190) (5 μM), and hydrogen peroxide (1 mM). Lanes 19 and 20 contain sodium ascorbate (1 mM) and hydrogen peroxide (1 mM). Lanes 21 and 22 are intact DNA.

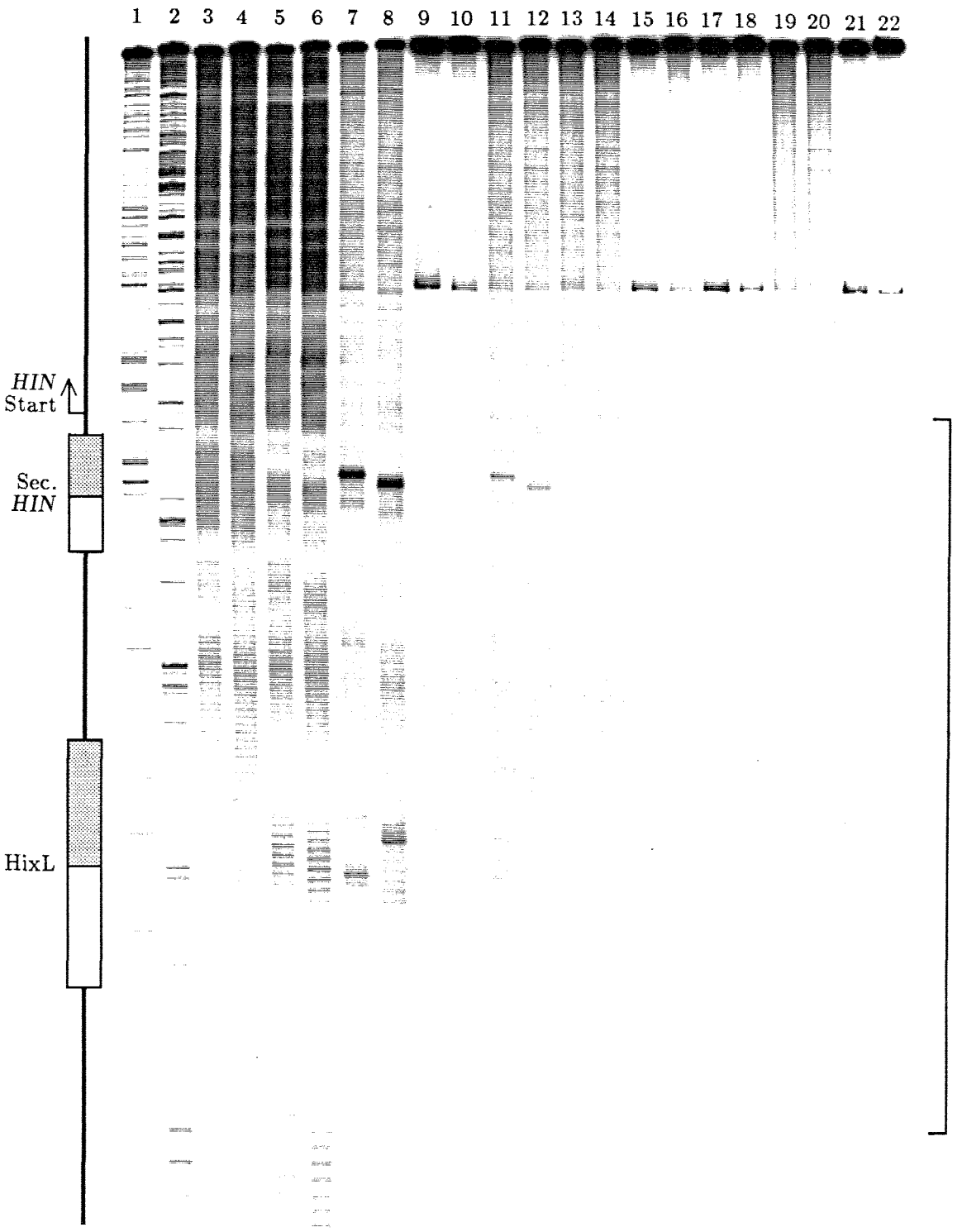
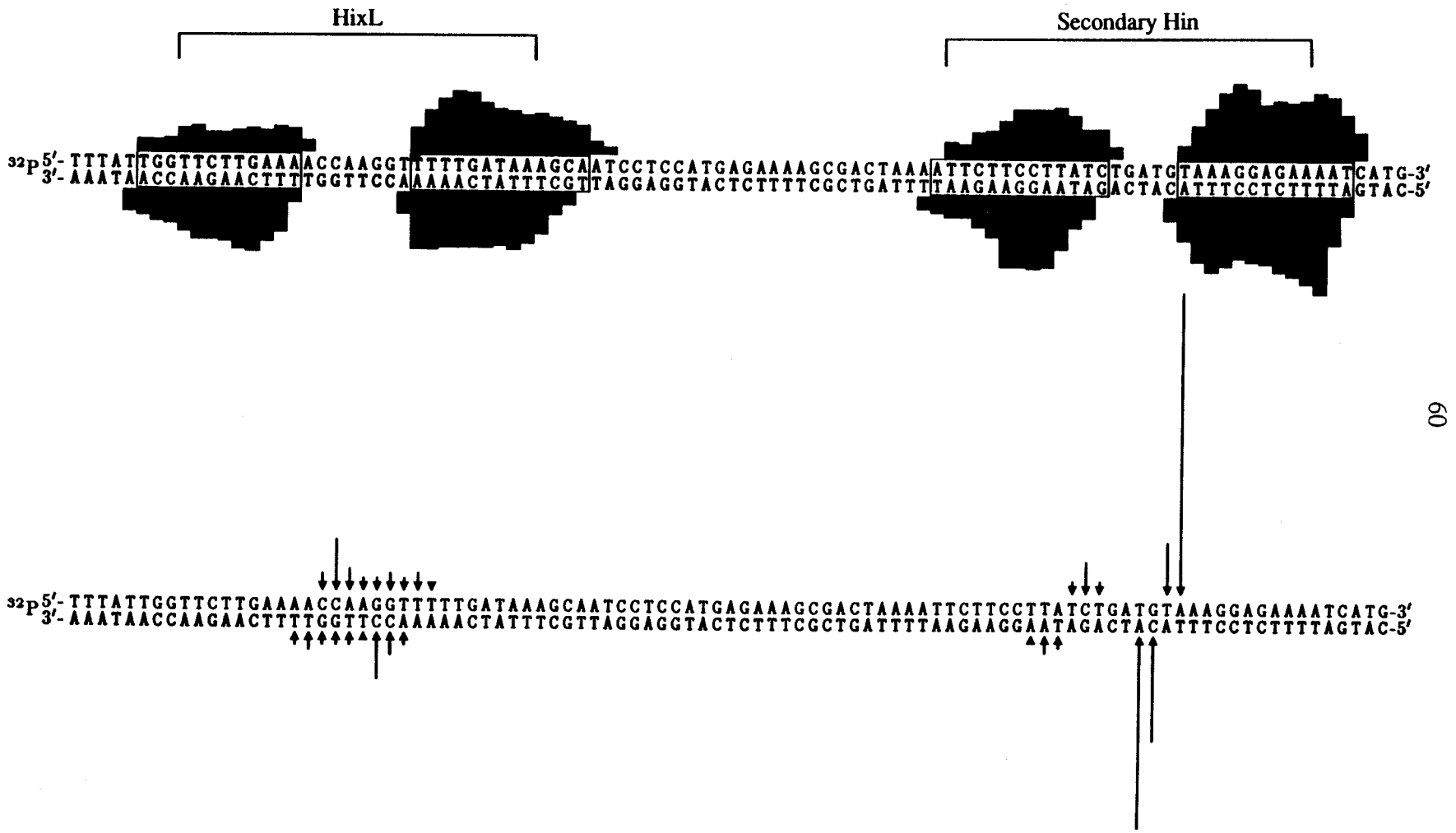

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
HIN
Start
Sec.
HIN
HixL

Figure 4 Histogram of the data in Figure 3. The sequence left to right represents the cleavage data from the bottom to the middle of the gel (bracket on the right) in Figure 3. Brackets show the hixL and secondary dimeric binding sites. Solid bars are histograms representing extent of protection from MPE•Fe cleavage in the presence of GGH(Hin 139-190) (0.5 μM) (Figure 3, lanes 5 and 6). Arrows represent extent of cleavage for Cu(II)•GGH(Hin 139-190) at 5 μM in the presence of sodium ascorbate and hydrogen peroxide (Figure 3, lanes 7 and 8). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.


HixL
Secondary Hin
32P 5′- TTTATTGGTTCTTGAAAACCAAGGTTTTTGATAAAGCAATCCTCCATGAGAAAAGCGACTAAAATTCTTCCTTATCTGATGTAAAGGAGAAAATCATG-3′
3′- AAATAACCAAGAACTTTTGGTTCCAAAAACTATTTCGTTAGGAGGTACTCTTTTCGCTGATTTTAAGAAGGAATAGACTACATTTCCTCTTTTAGTAC-5′
32P 5′- TTTATTGGTTCTTGAAAACCAAGGTTTTTGATAAAGCAATCCTCCATGAGAAAAGCGACTAAAATTCTTCCTTATCTGATGTAAAGGAGAAAATCATG-3′
3′- AAATAACCAAGAACTTTTGGTTCCAAAAACTATTTCGTTAGGAGGTACTCTTTTCGCTGATTTTAAGAAGGAATAGACTACATTTCCTCTTTTAGTAC-5′

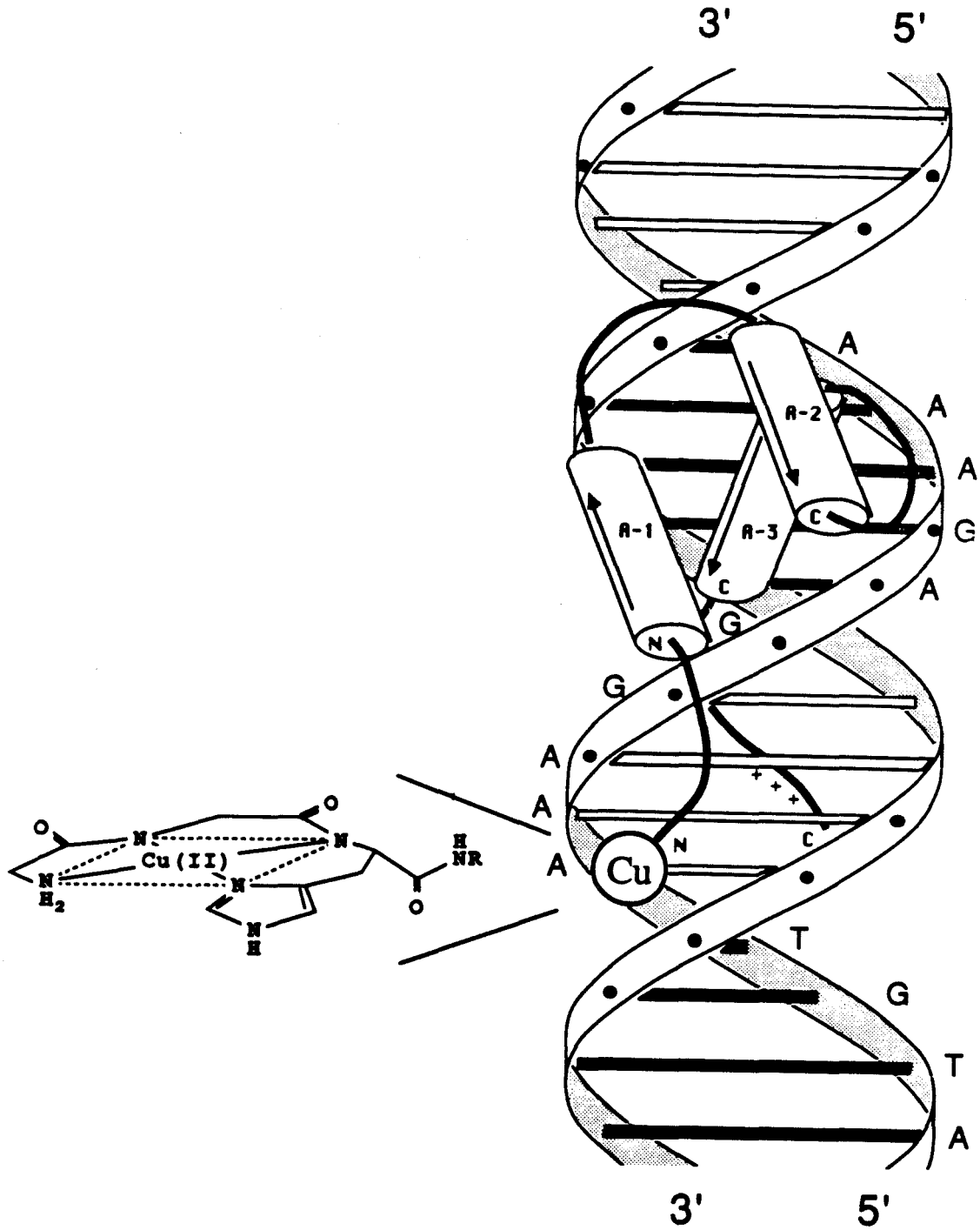


Figure 5 Schematic representation of a model for designed metalloprotein Cu(II)•GGH(Hin 139-190) binding to one Hin half site. Putative α helices are shown as cylinders with an arrow pointing from the NH_2 to COOH terminus.[51,69]

and charge, the nature of the termini of a cleavage fragment can be determined by using high resolution polyacrylamide gel electrophoresis if the site of cleavage is located in close proximity to the labeled end of the fragment.[13,14] For this purpose, the plasmid pDPM34 was constructed. Plasmid pDPM34 was constructed by deleting the DNA sequence in pBR322 between EcoRI and HindIII and replacing it with a sequence which contained the secondary Hin binding site (Figure 6). By linearizing the plasmid with either EcoRI or HindIII, end-labeling at the 5' or 3' end, and cutting again with restriction endonuclease RsaI, a 181 or 562 base pair restriction fragment could be obtained which had the secondary Hin binding site in very close proximity to the ^{32}P label.

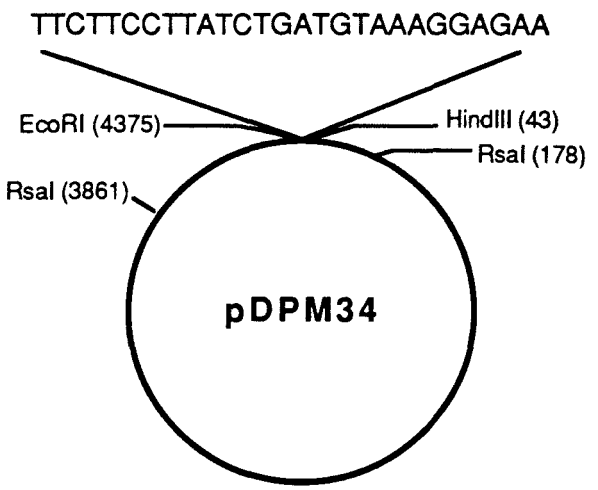


Figure 6 Schematic representation of plasmid pDPM34, indicating the sequence of the secondary Hin binding site which has been inserted between the EcoRI/HindIII sites of pBR322. The locations of the restriction sites for RsaI are also shown.

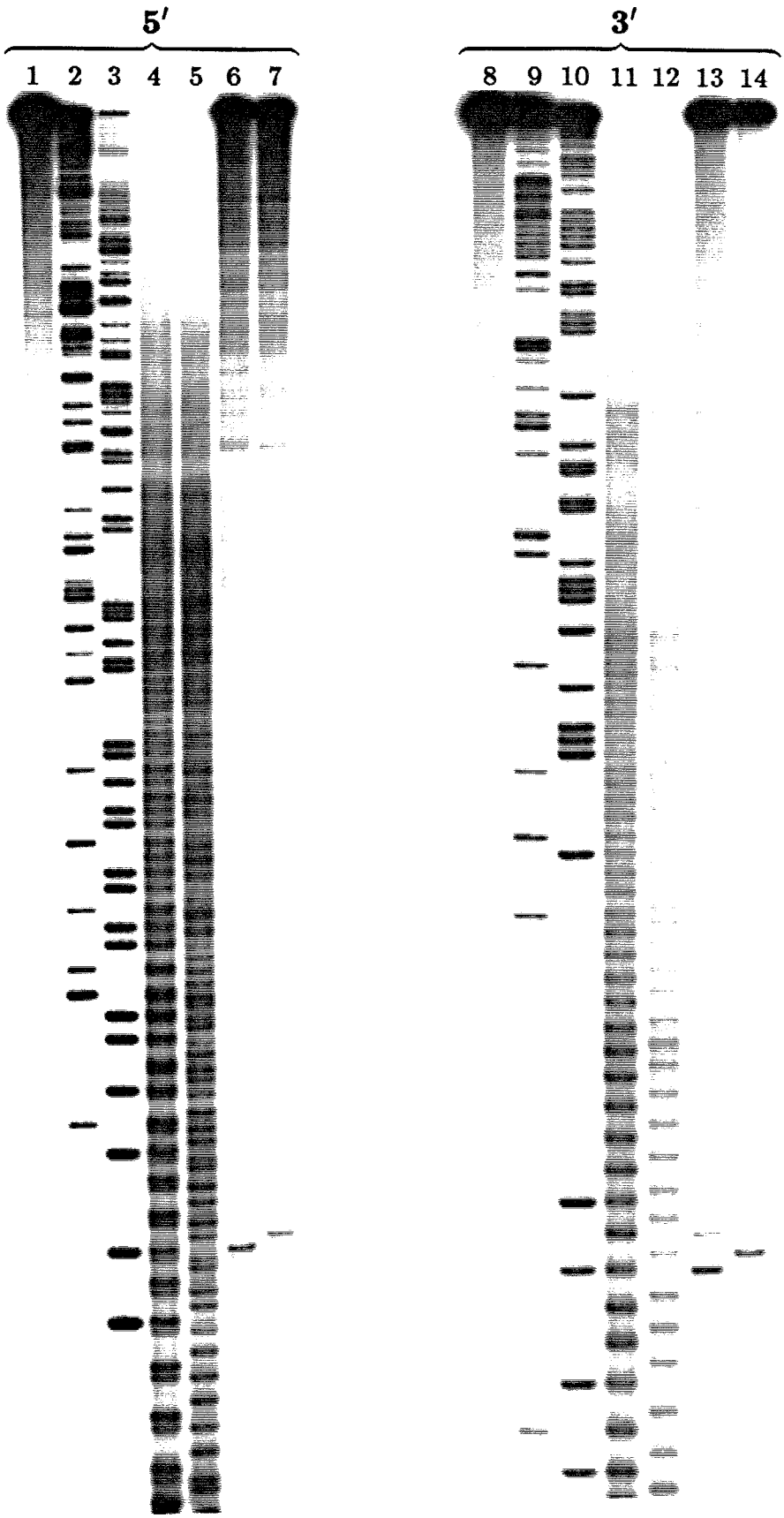
5'-End Products. The nature of the 5' termini of the cleavage reaction was investigated by use of a 3' labeled 181 bp fragment (EcoRI/RsaI) from pDPM34. The DNA restriction fragment was treated with Cu(II)•GGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide and the DNA cleavage products were analyzed by denaturing 20% polyacrylamide gel electrophoresis. The mobilities of the resulting DNA fragments were compared to those produced by the reaction of MPE•Fe(II) in the presence

of DTT. MPE•Fe(II) is known to produce 5' termini which are phosphorylated.[15] Figure 7 (lanes 11 and 13) demonstrates that Cu(II)•GGH(Hin 139-190) produced oligonucleotides which co-migrate with those from the MPE•Fe(II) reaction, indicating the presence of phosphate groups at the 5' termini.

Further evidence for the presence of phosphate groups is obtained by treatment of these oligonucleotide products with calf intestinal alkaline phosphatase, which removes 5' phosphate groups from DNA substrates.[16] Figure 7 (lane 12) shows that CAP treatment of a MPE•Fe(II) reaction results in a decrease in the electrophoretic mobility of the cleavage product due to the removal of a negatively charged phosphate group. CAP treatment of a Cu(II)•GGH(Hin 139-190) reaction (Figure 7, lane 14) produces the same shift in electrophoretic mobility, confirming that the Cu(II)•GGH(Hin 139-190) cleavage results in the production of 5' phosphate groups.

3'-End Products. In order to investigate the nature of the 3' termini, a 5' labeled 562 bp fragment (HindIII/RsaI) restriction fragment from pDPM34 was used. MPE•Fe(II) cleavage is known to produce a mixture of 5' phosphate and phosphoglycolate.[15] Treatment of these reaction products with T4 polynucleotide kinase under appropriate conditions leads to the removal of the 3' phosphate groups,[17] in analogy to the CAP reaction. These three reaction products, phosphate, phosphoglycolate, and hydroxyl group can be resolved by high resolution gel electrophoresis as shown in Figure 7 (lanes 4 and 5). Cleavage products from the reaction of Cu(II)•GGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide also produce a doublet which comigrates with the two bands produced by MPE•Fe(II) (Figure 7, lanes 4 and 6). The doublet for the Cu(II)•GGH(Hin 139-190) reaction however, is not of equal intensity. The upper band, tentatively identified as phosphate by analogy to the MPE•Fe(II) reaction, is in excess (at least 9:1) over the lower band, phosphoglycolate. Treatment of this reaction product with T4 polynucleotide kinase results in a slower moving band identified as a hydroxyl group, verifying the cleavage product as a 3' phosphate (Figure 7, lane 7).

Figure 7 Autoradiogram of high-resolution denaturing gel analyzing the 3' and 5' end-products of the Cu(II)•GGH(Hin 139-190)-sodium ascorbate-hydrogen peroxide DNA cleavage reaction. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in base pair) calf thymus DNA, and ≈15,000 cpm 5' or 3' end-labeled DNA (HindIII/RsaI or EcoRI/RsaI fragment from pDPM34, respectively) in a total volume of 15 μL. Reactions for MPE•Fe(II) (6.67 μM) were run for 17 min in the presence of 5 mM DTT. Reactions for Cu(II)•GGH(Hin 139-190) (5 μM) were run for 90 min in the presence of sodium ascorbate and hydrogen peroxide (1 mM each). DNA cleavage products were analyzed on a 20% 1:20 crosslinked 50% urea denaturing polyacrylamide gel 0.4 mm thick. Lanes 1-7 and lanes 8-14 contain 5' and 3' labeled DNA, respectively. Lanes 1 and 8 are intact DNA. Lanes 2 and 9 are Maxam-Gilbert chemical sequencing G lanes. Lanes 3 and 10 are A specific chemical sequencing lanes. Lanes 4 and 11 are MPE•Fe(II) reactions. Lanes 5 and 12 are MPE•Fe(II) reactions followed by treatment with T4 polynucleotide kinase and calf alkaline phosphatase, respectively. Lanes 6 and 13 are Cu(II)•GGH(Hin 139-190) cleavage reactions. Lanes 7 and 14 are Cu(II)•GGH(Hin 139-190) reactions followed by treatment with T4 polynucleotide kinase and calf alkaline phosphatase, respectively.


5′
3′
1 2 3 4 5 6 7
8 9 10 11 12 13 14

Identification of the termini at the cleavage site as 3' and 5' phosphate indicates that the cleavage reaction is oxidative degradation of the deoxyribose backbone. The range of the cleavage pattern (2 bp) is shorter than that usually seen with the diffusible oxidant generated by EDTA•Fe (typically 4-6 bp).[12] One interpretation of the restricted range of oxidative cleavage at the major site is that a nondiffusible oxidant is generated by the Cu•GGH complex in the minor groove. Two possibilities for the identity of the nondiffusible oxidant generated by the Cu•GGH complex are a metal bound oxygen or a peptide ligand-radical species.[18]

Discussion. Although it is not possible to rule out either of these oxidizing species, the presence of a ligand radical has been proposed in a number of peptide-Cu(II) systems.[18-20] The oxidation of peptides via their copper complexes has been a topic of interest for many years. The copper complexes of simple tetrapeptides, such as tetraglycine or tetra-L-alanine, undergo oxidative cleavage of the peptide chain at the third residue from the amino terminus.[18-20] Although not proven, a mechanism for the DNA cleavage, shown in Figure 8, has been proposed and is supported by a considerable amount of data. In the first step, the Cu(II) complex is oxidized to a Cu(III) complex. In the second step an electron is transferred from the peptide chain to the bound Cu(III). The formation of a carbon-centered radical is one of a number of possible third steps, depending on the exact system studied, all of which can lead to ligand oxidation. The presence of a ligand-centered radical positioned in the minor groove of DNA could lead to hydrogen atom abstraction and strand scission.

The cleavage pattern at the upper half of the secondary Hin site is strong and specific, while those at the other half-sites are weak and diffuse. Because the footprinting data suggest that all four half-sites are occupied, one explanation for the difference in cleavage patterns is that the structure of the DNA is different at each half-site and the oxidative pathway is sensitive to these differences. This analysis is not unreasonable since hixL and the secondary Hin sites have different biological roles[1] HixL is a recombination site while

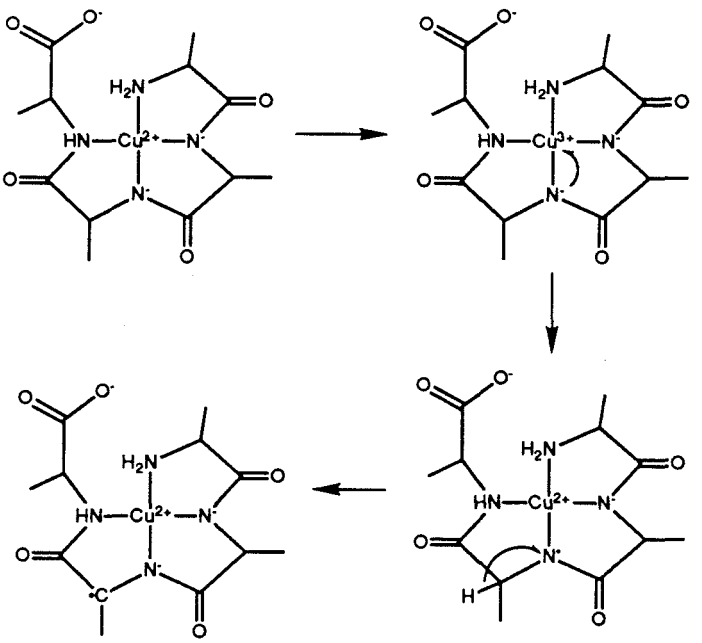


Figure 8 Mechanistic scheme for the formation of a ligand centered radical in the oxidation reaction of tetra l-alanine.[18-20]

secondary Hin is a binding site whose function is less well understood, but appears to be simply a binding site. Sequence-dependent structures at the two Hin sites might afford different accessibility of the Cu•GGH complex for oxidation of the deoxyribose ring in the minor groove at each half site.

Other Metal-Peptide Complexes. The ability of amino acids and small peptides to bind metals is well documented.[21-23] Small peptides, three to six amino acids in length are known to bind metals such as Pd(II), Cu(II), Ni(II), and Co(II). Most of these divalent metal ions form square-planar complexes with oligopeptides, or at least complexes in which the square-planar coordination sites are the dominant ones.[21] Preliminary results with a number of different metals indicated that the Ni(II) complex of GGH(Hin 139-190) was capable of cleaving DNA. Although a crystal structure of GGH•Ni(II) is not available, the Ni(II) complex of GGH is known and has been studied by other techniques.[24-27] Crystal structures of tetraglycine with Cu(II) or Ni(II) indicate that the metal ions are bound

by peptide ligands in a similar fashion.[28-30] Thus, as in the Cu(II) case, coordination of Ni(II) by GGH is believed to occur via the imidazole nitrogen, two deprotonated peptide nitrogens, and the terminal amino group.

Cleavage by Ni(II)•GGH(Hin 139-190). GGH(Hin 139-190) at 1.0 μM concentrations (pH 7.5, 25 °C, 20 mM NaCl) in the presence of $Ni(OAc)_2$ (1.0 μM) and monoperoxyphthalic acid (1.0 μM) cleaves DNA at all four sites upon base workup (0.1 M n-$BuNH_2$, 90 °C, 30 min) (Figures 9 and 10). The cleavage patterns observed at the hixL sites are strong and occur predominantly on one strand of each DNA site with single-base specificity, while cleavage at the secondary sites is modest and covers one to two base positions on both DNA strands (Figures 9 and 10). Maximal cleavage on opposite strands of the DNA is asymmetric to the 3' side, consistent with the known location of the NH_2-terminus of Hin(139-190) in the minor groove.[2,3]

Ni(II)•GGH(Hin 139-190) will also cleave the DNA in the presence of a number of other oxygen atom sources, such as hydrogen peroxide and iodosylbenzene. Figure 11 shows the DNA cleavage pattern obtained with GGH(Hin 139-190) (5 μM), $Ni(OAC)_2$ (5 μM), and hydrogen peroxide (5 mM) (lanes 5 thru 8), monoperoxyphthalic acid (2 μM) (lanes 9 thru 12), or iodosylbenzene (5 μM) (lanes 13 thru 16). The cleavage pattern obtained with or without base workup is identical in all cases except in the case of iodosylbenzene where cleavage at the tertiary site is lost. The similarity in cleavage pattern implies that a common intermediate is involved in each case.

End Product Analysis. The nature of the DNA cleavage reaction was again examined by high resolution gel electrophoresis to determine the exact nature of the DNA termini upon cleavage. Plasmid pDPM12 was used for the Ni(II)•GGH(Hin 139-190) reaction. Plasmid pDPM12 was constructed in a similar fashion to pDPM34, but in this case the hixL site was placed between the EcoRI and HindIII sites of pBR322 (Figure 12). By linearizing the plasmid at either the EcoRI or HindIII restriction sites, end-labeling at the 3' or 5' end, and cutting again with restriction endonuclease RsaI, a 180 or 561 base pair

Figure 9 Autoradiogram of high-resolution denaturing gel of Cu•GGH(Hin 139-190) and Ni(II)•GGH(Hin 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Cu•GGH(Hin 139-190) and Ni•GGH(Hin 139-190) were run for 90 and 15 min, respectively (25 °C). Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 contain $CuCl_2$ (2.5 μM) and GGH(Hin 139-190) (5 μM) followed by sodium ascorbate (1 mM) and hydrogen peroxide (1 mM). Lanes 7 and 8 are identical with lanes 5 and 6 except that they contain $CuCl_2$ (0.5 μM) and GGH(Hin 139-190) (1 μM). Lanes 9 and 10 contain $Ni(OAc)_2$ (5 μM) and GGH(Hin 139-190) (5 μM) followed by monoperoxyphthalic acid, magnesium salt (5 μM). Lanes 11 and 12 are identical to lanes 9 and 10 except they contain $Ni(OAc)_2$ (1 μ M), GGH(Hin 139-190) (1 μM), and monoperoxyphthalic acid (1 μM).

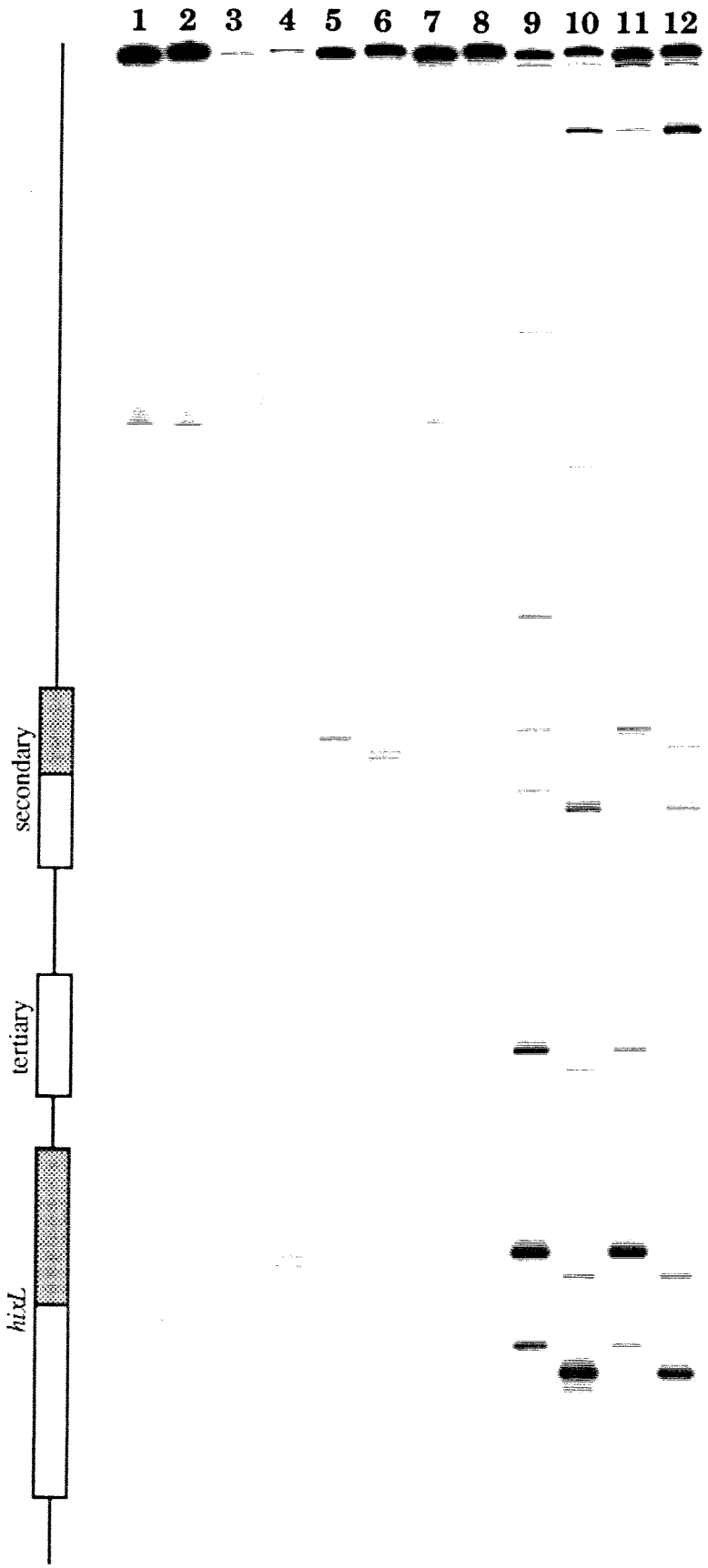

1 2 3 4 5 6 7 8 9 10 11 12
secondary
tertiary
hixL

Figure 10 Histograms of the data in Figure 9. The sequence left to right represents the cleavage data from the bottom to the middle of the gel in Figure 9. Boxes indicate the dimeric hixL and secondary Hin binding sites. Arrows represent location and extent of cleavage. Top: Cu(II)•GGH(Hin 139-190) (5 μM) in the presence of sodium ascorbate (mM) and hydrogen peroxide (mM) (Figure 9 lanes, 11 and 12). Bottom: Ni(II)•GGH(Hin 139-190) (1 μM) in the presence of monoperoxyphthalic acid (1 μM). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.

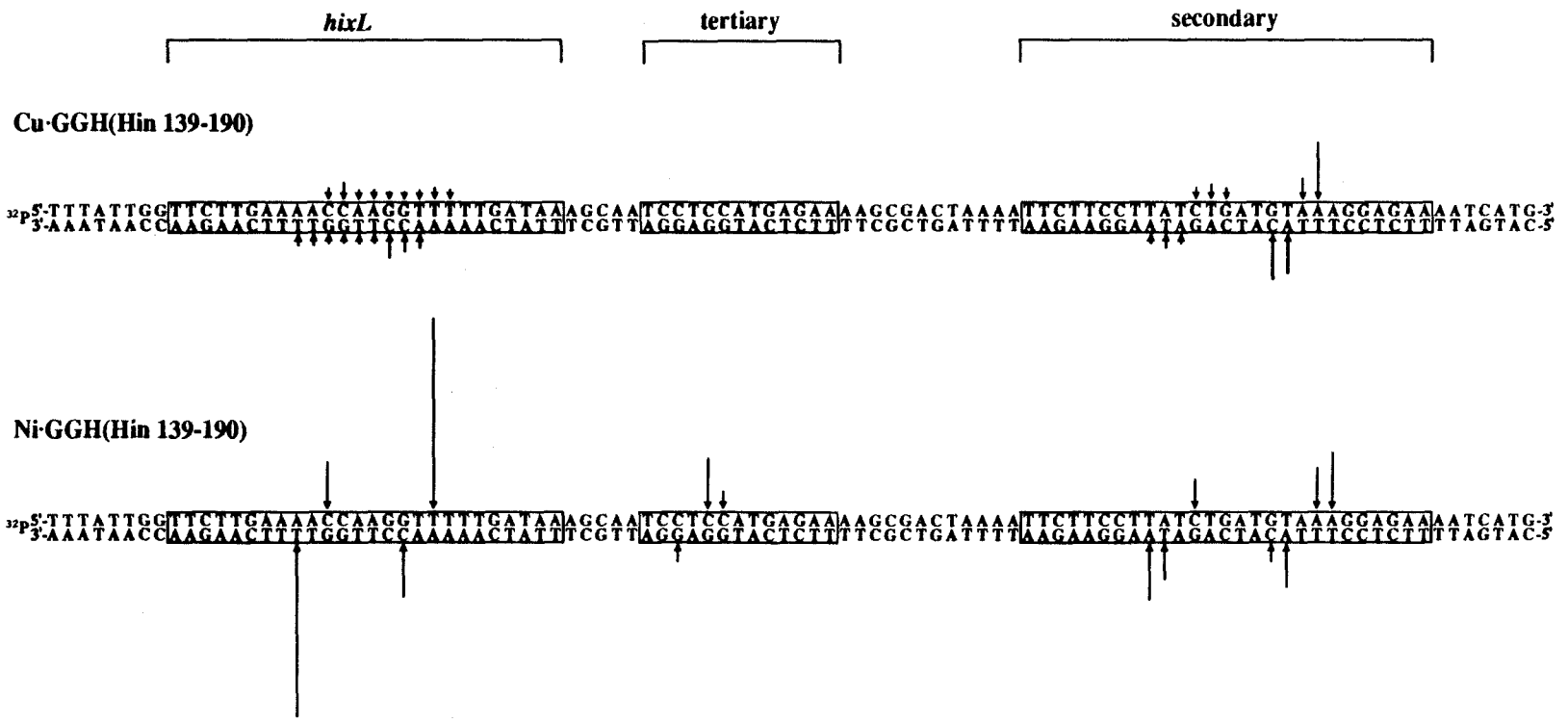
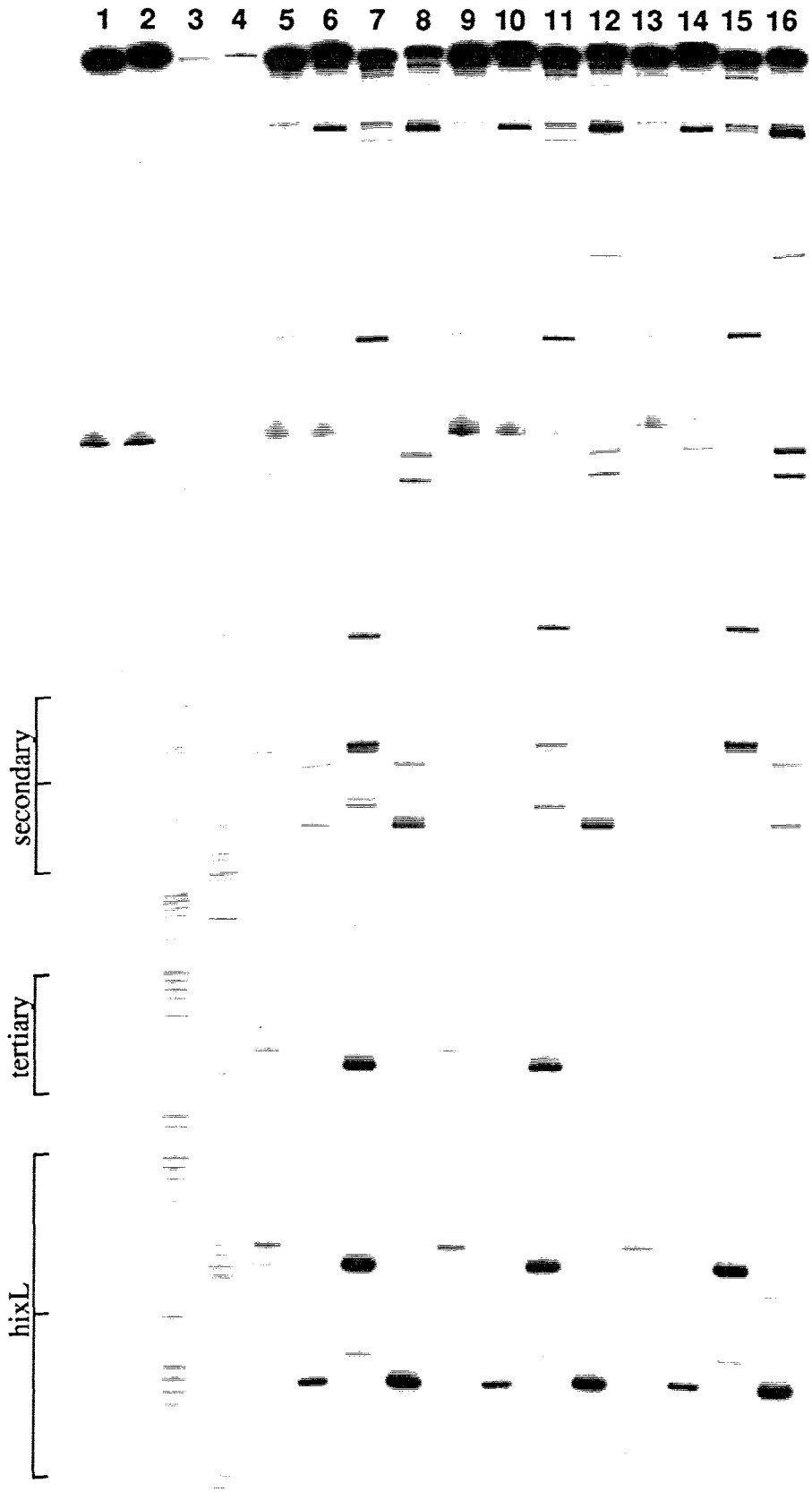

hixL
tertiary
secondary
Cu·GGH(Hin 139-190)
Ni·GGH(Hin 139-190)

Figure 11 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190) cleavage using hydrogen peroxide, monoperoxyphthalic acid, and iodosylbenzene of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 5 μM $Ni(OAc)_2$, 5 μM GGH(Hin 139-190), 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Ni•GGH(Hin 139-190) were run for 15 min (25 °C). Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 and 7 and 8 contain hydrogen peroxide (5 mM) without and with base workup, respectively. Lanes 9 and 10 and 11 and 12 contain monoperoxyphthalic acid (2 μM) without and with base workup respectively. Lanes 13 and 14 and 15 and 16 contain iodosylbenzene (5 μM) without and with base workup respectively.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
secondary
tertiary
hixL

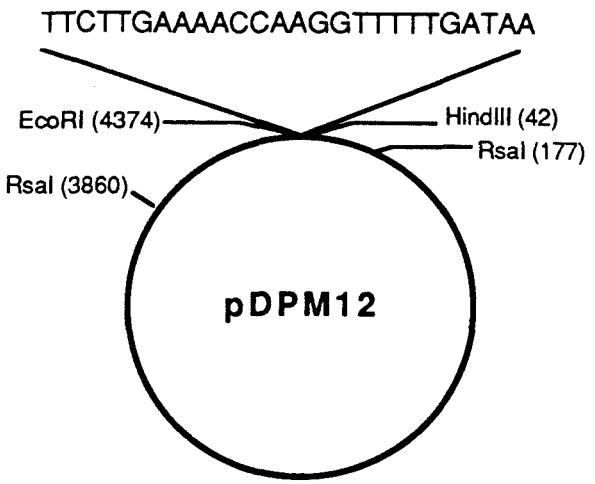
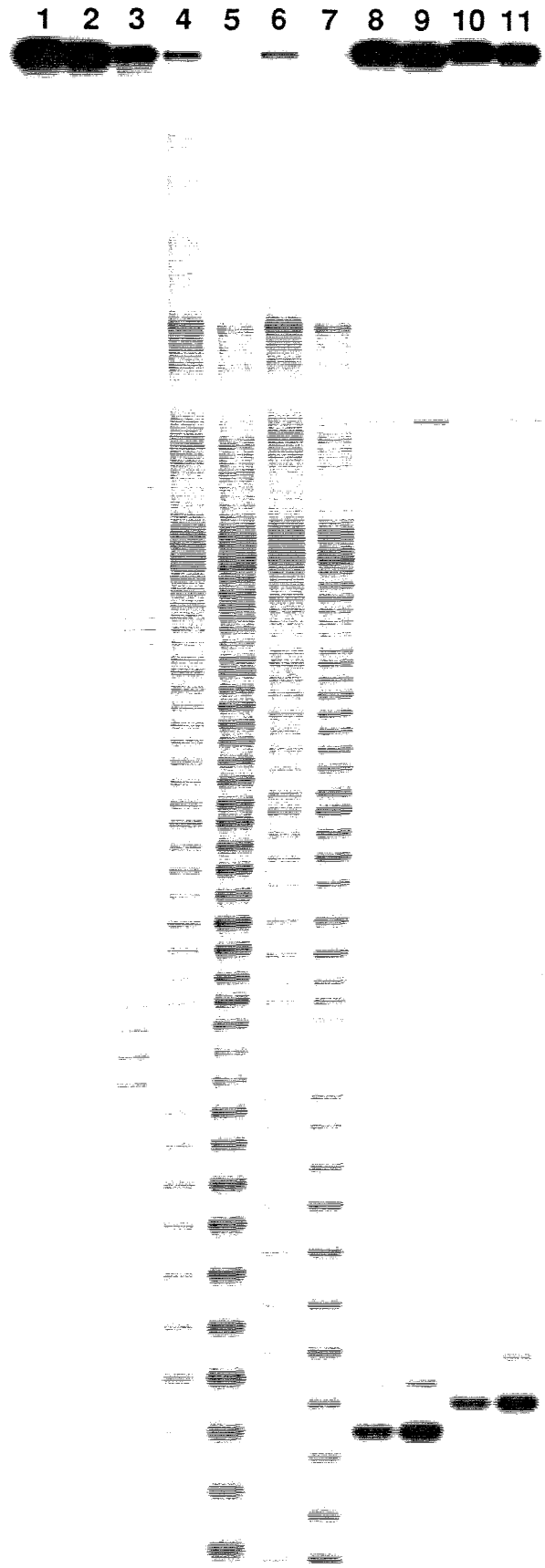


Figure 12 Schematic representation of plasmid pDPM12, indicating the sequence of the hixL binding site which has been inserted between the EcoRI/HindIII sites of pBR322. The locations of the restriction sites for RsaI are also shown.

restriction fragment could be obtained which has the hixL site in very close proximity to the ^{32}P label.

5'-End Products. The nature of the 5' termini of the cleavage reaction was investigated by use of a 3' end-labeled 180 bp fragment (EcoRI/RsaI) from pDPM12. The DNA restriction fragment was treated with Ni(II)•GGH(Hin 139-190) in the presence of monoperoxyphthalic acid and the DNA cleavage products analyzed by denaturing 20% polyacrylamide gel electrophoresis with and without base workup. The mobilities of the resulting fragments were compared to those produced by MPE•Fe(II), which is known to produce 5' termini with a phosphate group.[15] Figure 13 (lanes 4, 5, 8, and 9) demonstrates that Ni(II)•GGH(Hin 139-190) produces oligonucleotides that co-migrate with those from the MPE•Fe(II) reaction, indicating the presence of phosphate groups at the 5' termini. The identity of the product was confirmed by treatment of the reaction products with CAP to remove the 5' phosphate group[16] and create a fragment that should have decreased electrophoretic mobility. Figure 13 (lanes 6,7,10, and 11) shows that treatment with CAP produces the same shift in the mobility of the products of Ni(II)•GGH(Hin 139-190) as for

Figure 13 Autoradiogram of high-resolution denaturing gel analyzing the 5' end-products of the Ni(II)•GGH(Hin 139-190)-monoperoxyphthalic acid DNA cleavage reaction. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in base pair) calf thymus DNA, and ≈15,000 cpm 3' end-labeled DNA (EcoRI/RsaI fragment from pDPM12) in a total volume of 20 μL. Reactions for MPE•Fe(II) (10 μM) were run for 15 min in the presence of 5 mM DTT. Reactions for Ni(II)•GGH(Hin 139-190) (5 μM) were run for 15 min in the presence of monoperoxyphthalic acid (5 μM). Nickel-mediated cleavage reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. DNA cleavage products were analyzed on a 20% 1:20 crosslinked 50% urea denaturing polyacrylamide gel 0.4 mm thick. Lanes 1 and 2 are intact DNA without and with base workup, respectively. Lane 2 is a Maxam-Gilbert chemical sequencing G lanes. Lanes 4 and 5 are MPE•Fe(II) reactions without and with base workup, respectively. Lanes 6 and 7 are MPE•Fe(II) reactions followed by treatment with calf alkaline phosphatase, without and with base workup, respectively. Lanes 8 and 9 are Ni(II)•GGH(Hin 139-190) cleavage reactions without and with base workup, respectively. Lanes 10 and 11 are Ni(II)•GGH(Hin 139-190) reactions followed by treatment with calf alkaline phosphatase, without and with base workup, respectively.


1 2 3 4 5 6 7 8 9 10 11

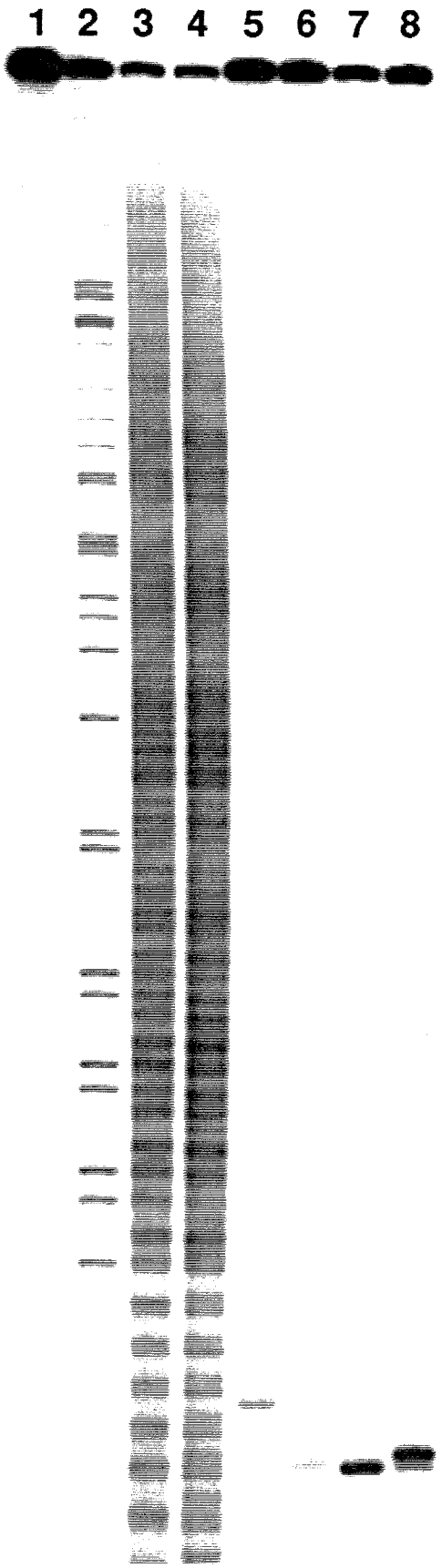
the products of MPE•Fe(II), confirming the identity of the cleavage product as 5' phosphate with or without base workup.

3'-End Products. The nature of the 3' termini was investigated on the 5' end-labeled 180 bp fragment (EcoRI/RsaI) from pDPM12. This DNA fragment was also treated with Ni(II)•GGH(Hin 139-190) in the presence of monoperoxyphthalic acid and the cleavage products analyzed by 20% denaturing polyacrylamide gel electrophoresis with and without base workup. The mobilities of the resulting fragments were compared to the known products produced by an MPE•Fe(II) reaction, 5' phosphate and phosphoglycolate.[15] Figure 14 (lanes 3 and 7) shows that the cleavage product produced by Ni(II)•GGH(Hin 139-190) upon base workup comigrates with the upper band of the doublet produced by MPE•Fe(II), identified as a 3' phosphate. The identity of this product was confirmed by treatment with T4 polynucleotide kinase to remove the 3' phosphate groups.[17] Figure 14 (lanes 4 and 8) shows that treatment with kinase produces the same shift in the mobility of the cleavage product of Ni(II)•GGH(Hin 139-190) as with the upper band from the MPE•Fe(II) reaction, indicating that the end-product is a 3' phosphate after base workup.

Discussion. Upon base workup, the DNA termini at the cleavage sites are 3' and 5' phosphate, consistent with oxidative degradation of the deoxyribose backbone. The small number of cleavage sites within each binding site implies that a nondiffusible oxidant is generated by both the Ni(II)•GGH(Hin 139-190)-peracid and Cu(II)•GGH(Hin 139-190)-H_2O_2-ascorbate reactions. However, differences in cleavage specificity are evident (Figure 9). Cu(II)•GGH(Hin 139-190) affords modest cleavage at one-half of the secondary site and weak cleavage at the other three sites. Ni(II)•GGH(Hin 139-190) affords strong cleavage at the hixL sites and modest cleavage at the secondary sites. Footprinting data presented earlier, however, suggests that all four sites are occupied at these concentrations by GGH(Hin 139-190).

One possible explanation for the difference in cleavage for the two metalloproteins is that the structure of the DNA is different at each site, and the reactive moieties on the

Figure 14 Autoradiogram of high-resolution denaturing gel analyzing the 3' end-products of the Ni(II)•GGH(Hin 139-190)-monoperoxyphthalic acid DNA cleavage reaction. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in base pair) calf thymus DNA, and ≈15,000 cpm 5' end-labeled DNA (EcoRI/RsaI fragment from pDPM12) in a total volume of 20 μL. Reactions for MPE•Fe(II) (10 μM) were run for 15 min in the presence of 5 mM DTT. Reactions for Ni(II)•GGH(Hin 139-190) (5 μM) were run for 15 min in the presence of monoperoxyphthalic acid (5 μM). Nickel-mediated cleavage reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. DNA cleavage products were analyzed on a 20% 1:20 crosslinked 50% urea denaturing polyacrylamide gel 0.4 mm thick. Lane 1 is intact DNA. Lane 2 is an A specific chemical sequencing lane. Lane 3 is an MPE•Fe(II) reaction. Lane 4 is an MPE•Fe(II) reaction followed by treatment with T4 polynucleotide kinase. Lanes 5 is a Ni(II)•GGH(Hin 139-190) cleavage reaction without base workup. Lanes 6 is a Ni(II)•GGH(Hin 139-190) cleavage reaction without base workup followed by treatment with T4 polynucleotide kinase. Lanes 7 is a Ni(II)•GGH(Hin 139-190) cleavage reaction with base workup. Lanes 8 is a Ni(II)•GGH(Hin 139-190) cleavage reaction with base workup followed by treatment with T4 polynucleotide kinase.


1 2 3 4 5 6 7 8

metalloproteins are sensitive to these sequence-dependent conformational differences. As was discussed earlier this is a reasonable possibility since the two binding sites, hixL and secondary Hin, have different sequences and different biological functions. HixL is one of the recombination sites for the native protein, while secondary Hin is proposed to be a simple binding site, responsible for autoregulation of the Hin gene.[1] The strong specific cleavage observed at hixL for Ni(II)•GGH(Hin 139-190) implies that a nondiffusible oxidant is held in a very rigid position and able to attack only a single base position. The cleavage patterns observed at secondary Hin for both Cu(II) and Ni(II)-GGH(Hin 139-190) imply that the oxidizing species is not held in a rigid a fashion and has some mobility along the DNA. The lack of cleavage by Cu(II)•GGH(Hin 139-190) at hixL is likely due to the fact that the protein complex is held in a rigid fashion that does not position the oxidative species, which is likely different from that in the Ni(II) system, to be in a position to react with the DNA.

There are also significant differences in the reactivities of the Cu(II)•GGH-H_2O_2-ascorbate and Ni(II)•GGH-peracid systems toward the DNA. The nickel-mediated reaction in the presence of 1 equivalent of peracid is complete in 1 min (25 °C) (data not shown). For copper-mediated cleavage in the presence of a 500-fold excess of ascorbate and hydrogen peroxide, 1.5 hour reaction times are necessary (25 °C). Moreover, the extent of cleavage obtained from the Ni(II)•GGH-peracid reaction is significantly greater than that for the Cu(II)•GGH-H_2O_2-ascorbate reaction. Cu(II)•GGH(Hin 139-190) will also cleave the DNA in the presence of monoperoxyphthalic acid to give a cleavage pattern identical to that of the Ni(II) complex, however longer reaction times and significantly higher oxidant concentrations are required (data not shown).

The highly reactive nature of the oxidizing species, the requirement for an oxygen atom source, the precision of the cleavage reaction, and the identity of the DNA termini created upon treatment with butylamine[31] imply that the reactive may be a high valent nickel bound oxygen that abstracts a specific hydrogen atom on the deoxyribose backbone.

Although a mechanisitic discussion will be deferred until later, these results with Cu(II) and Ni(II)-GGH(Hin 139-190) have implications for design of protein catalysts. The tripeptide GGH is a metal-specific structural domain consisting of naturally occurring amino acids that could be incorporated at the NH_2 terminus of a variety of recombinant proteins (such as DNA binding proteins, receptors, or antiibodies) with the function of precise efficient substrate-directed oxidation, activated in the presence of Ni(II) and peracid.

DNA Double Strand Cleavage. The ability of EDTA modified DNA binding molecules to cause double strand cleavage is well documented,[2,32,33] and the assay used to observe this cleavage is shown in Figure 15. In this assay, supercoiled plasmid is linearized and specifically labeled at one end with ^{32}P. After the cleavage reaction the DNA fragments are separated on an agarose gel and visualized by autoradiography. The size of the fragments (± 20 bp) is then determined by comparison with the appropriate molecular weight markers produced by restriction enzymes. The cleavage patterns produced by Ni(II)•GGH(Hin 139-190) on DNA restriction fragments indicate that double strand cleavage should also be possible by this protein. The increased cleavage observed for Ni(II)•GGH(Hin 139-190) at some sites and decreased cleavage at others indicates that the cleavage pattern should be different than that obtained with Fe•EDTA-Hin(139-190). This difference may be further enhanced by the fact that Ni(II)•GGH(Hin 139-190) shows a strand preference at some sites (hixL and tertiary), indicating that a head-to-head orientation of the proteins may be necessary for maximal double strand cleavage. Double strand cleavage of linearized plasmid pDPM12 containing the hixL site, by Ni(II)•GGH(Hin 139-190)/peracid and Fe•EDTA-Hin(139-190)/DTT is shown in Figure 16. A number of cleavage sites are seen for each protein, and each maps to sites on pDPM12 with varying degrees of homology to various hix half sites (Figures 16 and 17). The sequence of each cleavage site is shown in Table 1.

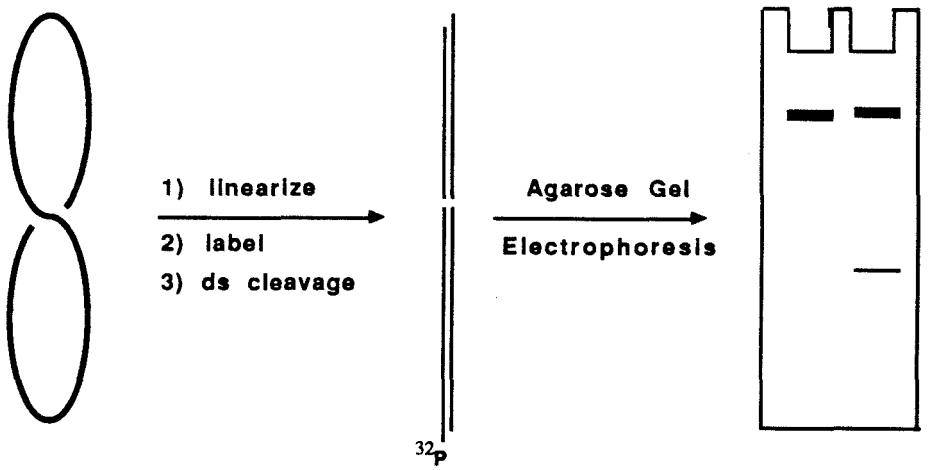
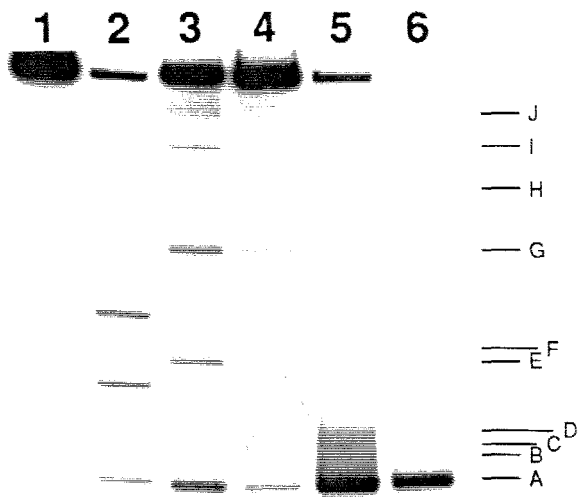


Figure 15 Double stranded cleavage assay. A circular plasmid is linearized with a restriction enzyme that gives differentiable termini. Labeling at a single end and treatment with specific cleavage agents gives a set of double stranded fragment (only one shown) which can be separated on an agarose gel.

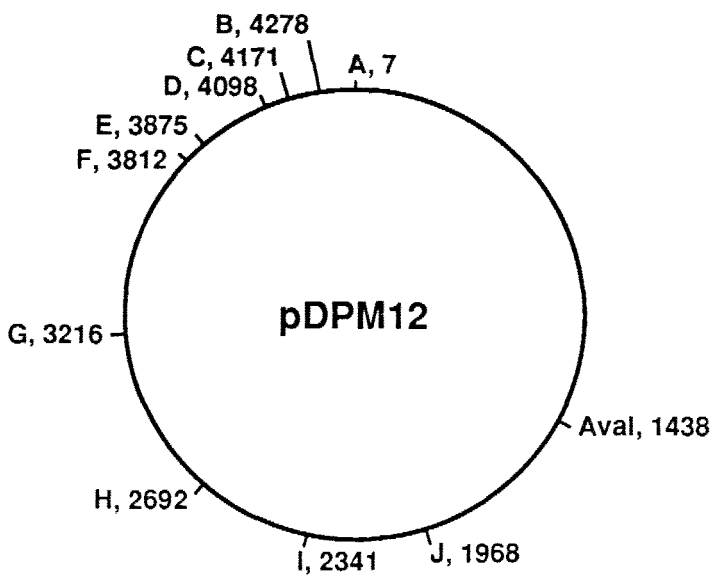
Histograms for the cleavage sites in Ni(II)•GGH(Hin 139-190) and Fe•EDTA-Hin(139-190) are shown in Figure 17. The strongest cleavage site for each protein is located at the hixL site as expected. All of the sites cut by Ni(II)•GGH(Hin 139-190) are also cut by Fe•EDTA-Hin(139-190); however, the Ni-protein does not cut at all of the sites cut by the EDTA-protein, nor is the cleavage intensity at a given site the same. These differences are expected given the cleavage patterns observed at different sites on restriction fragments by Ni(II)•GGH(Hin 139-190) as discussed earlier. It is interesting to note that the strongest cleavage by Ni(II)•GGH(Hin 139-190) occurs at the hixL site where there are two half-sites oriented such that the proteins are in a head-to-head configuration. These results again show the dramatic cleavage efficiency attained by the Ni(II)•GGH-peracid system. Both proteins were studied at 2.5 μM concentrations; however, the cleavage by Fe•EDTA-Hin(139-190) required 60 minutes and 2000 equivalents of DTT, while Ni(II)•GGH(Hin 139-190) required less than 15 minutes and only 1 equivalent of monoperoxyphthalic acid to produce stronger cleavage.

Figure 16 Double strand cleavage by Fe(II)•EDTA-Hin(139-190)/DTT and Ni(II)•GGH(Hin 139-190)/monoperoxyphthalic acid reactions. A) Autoradiogram of a nondenaturing agarose gel of Fe(II)•EDTA-Hin(139-190)/DTT and Ni(II)•GGH(Hin 139-190)/monoperoxyphthalic acid cleavage reactions on linearized pDPM12. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in base pair) calf thymus DNA, and ≈6000 cpm end-labeled linearized (AvaI) pDPM12. Reactions for Fe(II)•EDTA-Hin(139-190) and Ni(II)•GGH(Hin 139-190) were run for 60 and 15 min, respectively. Nickel-mediated reactions were treated with 0.1% piperidine in the presence of 80 mM NaCl at 55 °C for 15-18 hours. Cleavage products were analyzed on a 1% agarose gel (25 cm long X 4 mm thick). Lane 1 is intact DNA. Lane 2 is molecular weight markers. Lane 3 contains Fe(II)•EDTA-Hin(139-190) (2.5 μM) and DTT (5 mM). Lanes 3,4, and 5 contain Ni(II)•GGH(Hin 139-190) (2.5 μM) and 0.25, 1, or 4 equivalents of monoperoxyphthalic acid, respectively. B) Location of the cleavage sites (A-J) with EcoRI defined as the origin. The AvaI site of labeling is also shown.

A.


1
2
3
4
5
6
J
I
H
G
F
E
D
C
B
A

B.


B, 4278
C, 4171
D, 4098
A, 7
E, 3875
F, 3812
pDPM12
G, 3216
AvaI, 1438
H, 2692
I, 2341
J, 1968

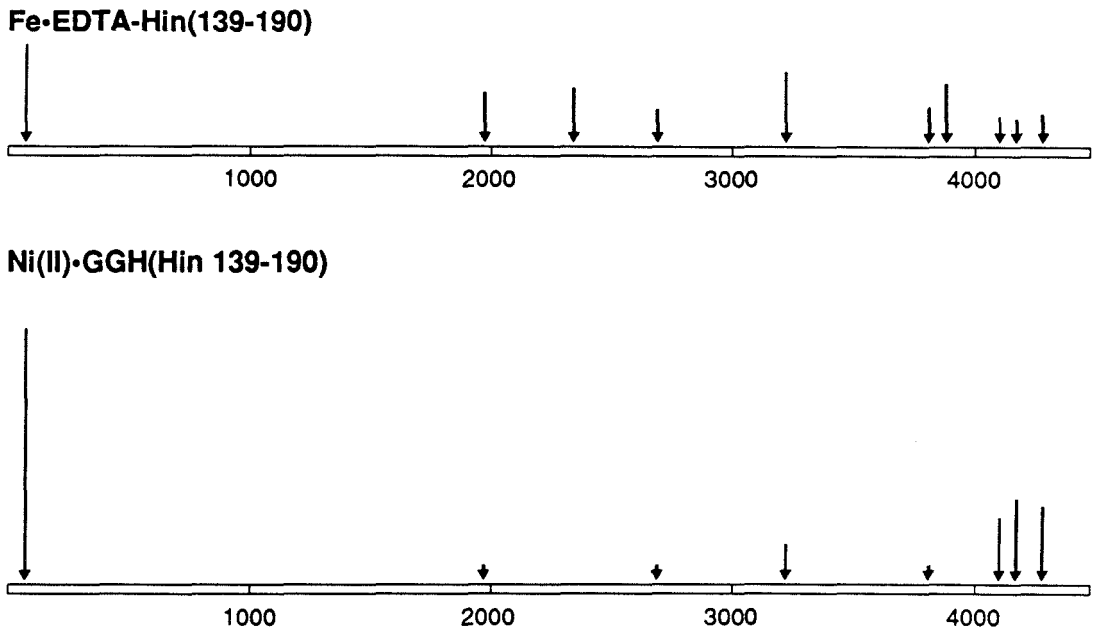


Figure 17 Cleavage sites on pDPM12 for Fe•EDTA-Hin (139-190)/DTT and Ni(II)•GGH(Hin 139-190)/peracid reactions. Arrows represent the extent of cleavage for lanes 3 and 5 in Figure 15.

Table 1 Cleavage Sites		
A	7	TTCTTGAAAA
		TTATCAAAAA
B	4278	TTCCCCGAAA
C	4171	TCTTCCTTTT
D	4098	TGAGCAAAAA
E	3875	TTCTCAGAAT
F	3812	TTCTCTTACT
G	3216	TTATCAAAAA
H	2692	TTACCGGATA
I	2341	TTTTCTCCTTA
J	1968	TTCTGATAAA

Cleavage Conditions

As has been shown, a large number of possible reaction conditions (time, concentration, nature of oxidizing or reducing agent, pH, and metal equivalents) can all be important for determining the exact nature and extent of the DNA cleavage reaction for both Cu(II) and Ni(II) complexes of GGH(Hin 139-190). For this reason studies were

undertaken to determine the exact requirements for the cleavage reaction of both Cu(II) and Ni(II)-GGH(Hin 139-190). Gel autoradiograms (appendix) were analyzed by laser densitometry and the extent of cleavage measured as peak area.

Time Courses. The time course of the DNA cleavage reactions by Ni(II)•GGH(Hin 139-190) with monoperoxyphthalic acid and hydrogen peroxide were studied along with those of Cu(II)•GGH(Hin 139-190) with monoperoxyphthalic acid and sodium ascorbate-hydrogen peroxide (Figure 18). Reactions involving iodosylbenzene were not studied because of errors involved in determining the exact concentrations in water due to solubility problems. The DNA cleavage reaction of Ni•GGH(Hin 139-190) mediated by monoperoxyphthalic acid was complete within the initial 15 minutes (Figure 18A). Further studies indicated that the reaction is complete in less than 1 minute (data not shown). The cleavage reaction mediated by H_2O_2 however appeared to be still increasing even after 3 hours (Figure 18B). Although the extent of cleavage for the H_2O_2 reaction approached that of the peracid reaction, the H_2O_2 was used in a 1000 fold excess as compared to the peracid which was in an equimolar amount compared to protein; in addition, the cleavage reaction was >100 times slower.

The DNA cleavage reaction of Cu(II)•GGH(Hin 139-190) with monoperoxyphthalic acid also appeared to be complete within the initial time period examined, 30 minutes (Figure 18C). The cleavage efficiency as compared to the Ni(II) reaction is considerably less despite a 200 fold increase in oxidant concentration. This reaction then appears to be quite facile as in the case of the Ni(II) complex, but the overall efficiency of the DNA cleavage reaction is less, indicating that possibly the final oxidizing species is not as reactive. The reaction for the Cu(II)•GGH(Hin 139-190) complex with a mixture of hydrogen peroxide and sodium ascorbate, in comparison to the peracid system, appeared to continue cutting after 3 hours (Figure 18D).

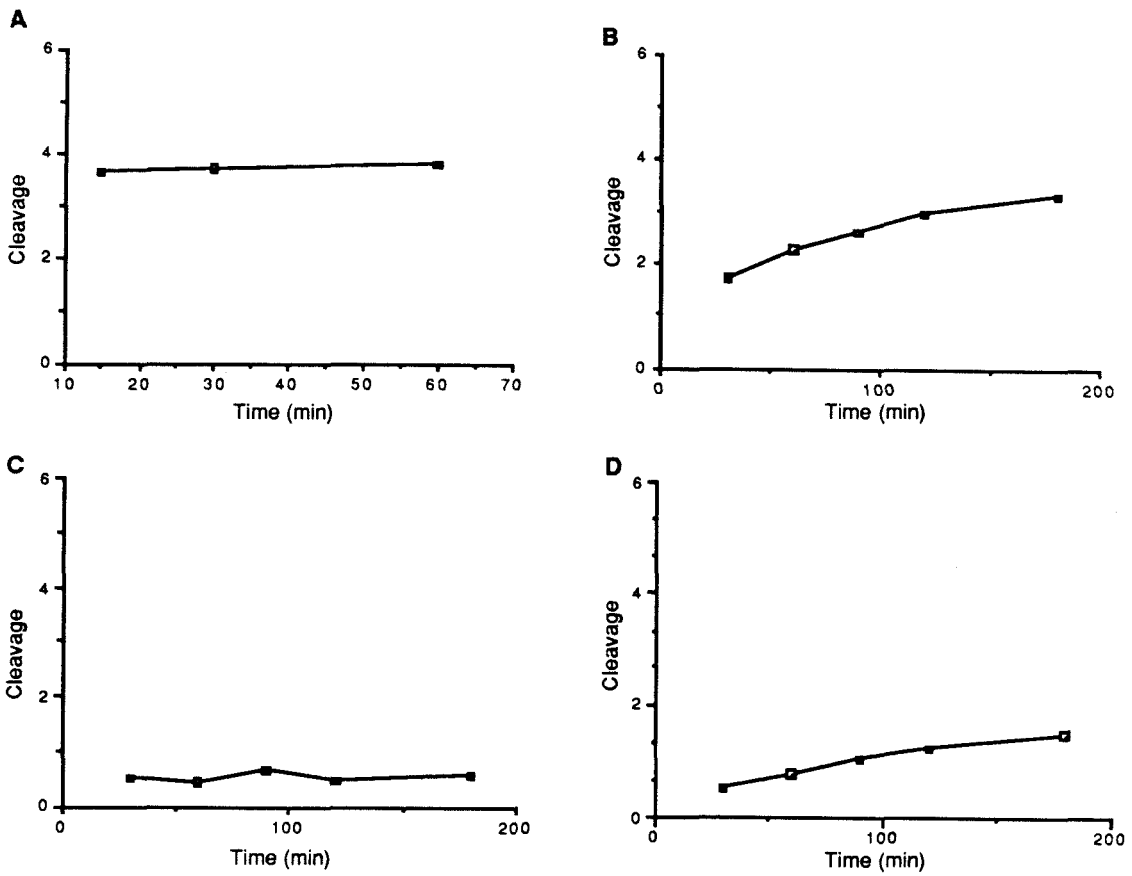


Figure 18 Time course of DNA cleavage reactions by A) Ni(II)•GGH(Hin 139-190)/monoperoxyphthalic acid, B) Ni(II)•GGH(Hin 139-190)/H_2O_2, C) Cu(II)•GGH(Hin 139-190)/monoperoxyphthalic acid, and D) Cu(II)•GGH(Hin 139-190)/H_2O_2-sodium ascorbate. Vertical axis represents an integrated area from laser densitometry.

Cu(II)•GGH(Hin139-190)-H_2O_2-Sodium Ascorbate Cleavage. The extent of DNA cleavage for the Cu(II)•GGH(Hin 139-190)-H_2O_2-ascorbate system was studied as a function of H_2O_2 and ascorbate concentrations as well as a function of pH (Figure 19). By varying both the concentrations of H_2O_2 and ascorbate independently, it can be seen that the DNA cleavage reaction is much more sensitive to the concentration of H_2O_2 than to the concentration of ascorbate (Figure 19A). The pH profile for the cleavage reaction is shown in Figure 19B. The sigmoidal nature of the curve indicates that the protonation state of some species is important for the cleavage reaction. Although the Cu(II)•GGH complex

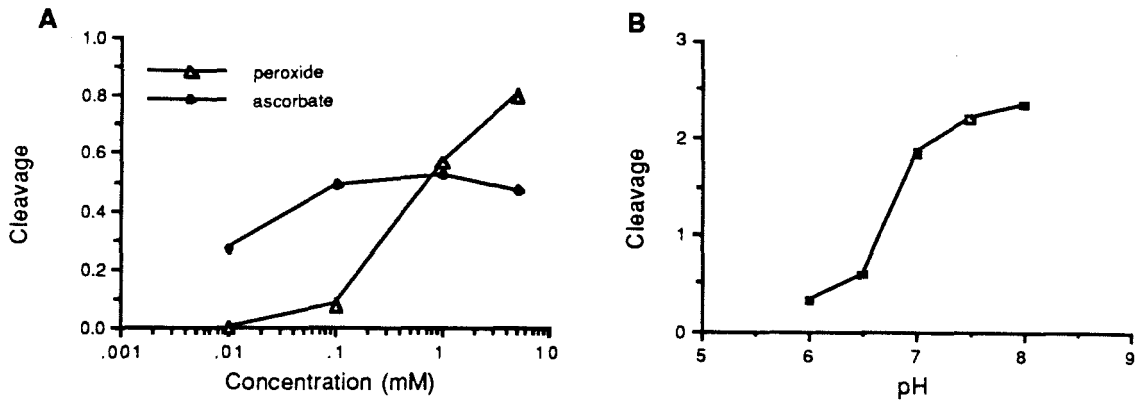


Figure 19 Plots of data from A) H_2O_2 and sodium ascorbate concentration study of Cu(II)•GGH(Hin 139-190) reaction and B) pH study of Cu(II)•GGH(Hin 139-190) reaction. Vertical axis represents an integrated are from laser densitometry of the gel autoradiograms (scales are not comparable since the data comes from two independent autoradiograms).

is known to be stable over this pH range,[6] it is likely that the exact protonation state of the complex is considerably different at the low pH values based on the known pK values for the deprotonation of the complex.[24] A change in protonation state of the ligand will alter the coordination about the metal center, possibly changing its redox properties and thus inhibiting the cleavage reaction.

Cu(II)•GGH(Hin 139-190)-Peracid Cleavage. The DNA cleavage reaction of Cu(II)•GGH(Hin 139-190) with monoperoxyphthalic acid was studied as a function of peracid concentration (Figure 20). Examination of the data plot shows that the cleavage efficiency initially went up with peracid concentration but then decreased as the peracid concentration continued to increase. The cleavage reaction, however, never reached the yield obtained for the Ni(II) complex using the same oxidant. In fact, maximal cleavage in the Cu(II) system occurred with ca. 10 equivalents of peracid, and the amount of cleavage did not reach that obtained in the Ni(II) system with a single equivalent of oxidant (see appendix for gel autoradiograms).

Ni(II)•GGH(Hin 139-190)-H_2O_2 Cleavage. The DNA cleavage reaction of Ni(II)•GGH(Hin 139-190) mediated by hydrogen peroxide was studied as a function of

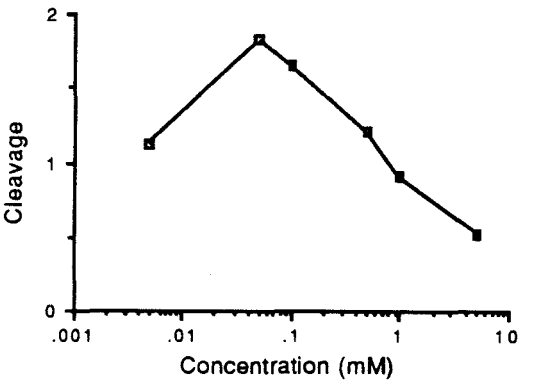


Figure 20 Plot of the data for the Cu(II)•GGH(Hin 139-190) reaction with varying concentrations of monoperoxyphthalic acid. Vertical axis represents an integrated area from laser densitometry of the gel autoradiogram.

oxidant concentration (Figure 21). The concentration range studied for this reaction was large, four orders of magnitude, due to the slower nature of the cleavage reaction and the necessity to choose a reasonable time of reaction (90 min). Quantitation revealed that the extent of cleavage reached a maximum of 100% at a 20,000 fold excess of H_2O_2 over protein complex. These concentrations of H_2O_2, however, randomly damage the DNA. Control reactions which contained no metal-protein complex, but a 100 mM concentration of H_2O_2 without and with base workup indicate that a large amount of base labile lesions

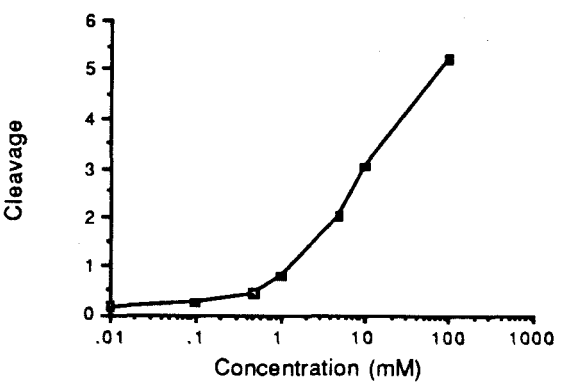


Figure 21 Plot of the data for the Ni(II)•GGH(Hin 139-190) reaction with varying concentrations of H_2O_2. Vertical axis represents an integrated area from laser densitometry of the gel autoradiogram.

had been generated in the DNA at these high concentrations of oxidant as there is a complete absence of intact DNA in the control lane that had been treated with base (see appendix for gel autoradiogram).

Ni(II)•GGH(Hin 139-190)-Peracid Cleavage. The DNA cleavage reaction of Ni(II)•GGH(Hin 139-190) with monoperoxyphthalic acid was studied as a function of oxidant concentration, pH, and equivalents of metal (Figure 22). The range of peracid concentrations studied varied from 1 to 1000 equivalents (Figure 22A). The cleavage efficiency increased from a low value at 1 equivalent to a high of 100% cleavage at 50 equivalents; after that point, increasing the peracid concentration further resulted in a decrease in the amount of cleavage. The reaction that produces 100% cleavage is further remarkable in that every radiolabeled strand was cut at the hixL site, and that this reaction was quite fast, occurring within 15 minutes. The observation that the total cleavage reaches a maximum and then falls off with increasing oxidant concentration is interesting. Two possible explanations for this observation are that a further oxidation of the nickel species, which inhibits cleavage is taking place in competition with reaction with the DNA, or that the oxidant itself could be a substrate for the nickel species and at high concentrations begins to compete with the DNA.

The extent of cleavage by Ni(II)•GGH(Hin 139-190) with peracid was also studied as a function of pH. The pH profile is again, as in the Cu(II)-H_2O_2-ascorbate system, a sigmoidal curve, indicating that the protonation state of some species is important for the cleavage reaction (Figure 22B). It is again likely that at the lower pH values studied here the protonation state of the ligand and hence the coordination about the metal will be different than at the higher pH values.[24] This change in coordination could indeed have large effects on the cleavage chemistry by changing the redox properties of the metal.

Due to the small amount of purified protein available and the unstable nature of the metal-protein complex, it is not possible to determine by standard spectroscopic techniques how much, if any, of a given metal is bound to the protein. As an indirect approach to

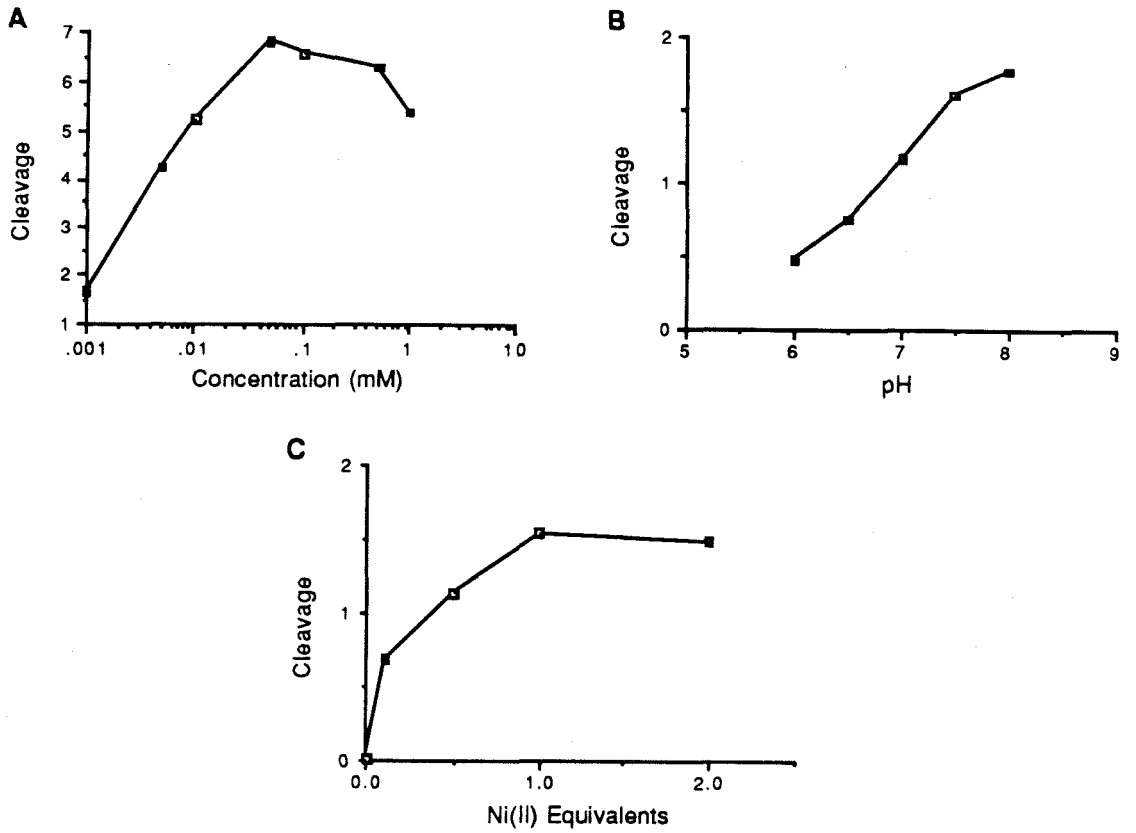


Figure 22 Plots of the data for A) monoperoxyphthalic acid concentration study of the Ni(II)•GGH(Hin 139-190) reaction, B) pH study of the Ni(II)•GGH(Hin 139-190)/peracid reaction, and C) Ni equivalents study of the Ni(II)•GGH(Hin 139-190)/peracid reaction. Vertical axes represent an integrated area from laser densitometry of the gel autoradiograms (scales are not comparable since the data comes from three independent autoradiograms).

answer this question the DNA cleavage reaction of Ni(II)•GGH(Hin 139-190) with peracid was studied as a function of the amount of nickel loaded into the protein before reaction with the DNA. The amount of metal loaded into the protein varied from 0.1 equivalents to 2 equivalents (Figure 22C). The amount of cleavage increases up to one equivalent and then levels off, not increasing upon addition of more metal. Although the curve is not linear as might be expected for this type of experiment, the results imply that at 1 equivalent of Ni(II) all of the proteins capable of binding a metal are occupied. The nonlinearity in the curve can be explained if the DNA cleavage reaction is catalytic, thus more cleavage is observed than expected at low metal concentrations due to catalytic turnover.

Modifications to the GGH Ligand System

MGGH(Hin 139-190) and AcGGH(Hin 139-190). As has been shown, the DNA cleavage efficiency for the Ni(II)•GGH(Hin 139-190) metalloprotein in the presence of an oxygen atom donor (e.g., monoperoxyphthalic acid) is quite remarkable. Since the metal binding domain of the protein is composed entirely of naturally occurring α-amino acids the question then arises whether this sequence can be coded for by the proper DNA sequence and expressed functionally in a living cell. The initial design of the 55 residue protein took into account that the tripeptide, GGH, was a consensus sequence for the amino terminal residues of serum albumin and that the protein would be produced by synthetic methodology.[34] However, nature synthesizes proteins from amino to carboxy terminus, which is opposite that of synthetic procedures; furthermore, in a natural system, a start codon is necessary to begin a given coding region on the DNA, which always begins with an amino terminal methionine.

In order to explore whether or not the GGH sequence could be used in the middle of a protein chain or at the amino terminus of a natural protein, two new proteins were synthesized, AcGGH(Hin 139-190) and MGGH(Hin 139-190), which contain an acylated amino terminus and an amino terminal methionine residue, respectively. Both of these proteins contain an amide bond two amino acids away from the histidine residue, as would be found at an internal position along the protein, instead of a free amino group. The acetylated protein was chosen since it produces the smallest possible steric change in the protein, while still altering the functionality of the terminal amino group. The acylated protein will therefore hopefully answer the question as to what effect a third amide group will have on the chemistry of the metal center without altering the DNA binding properties of the protein. The amino terminal methionine protein was chosen in order to test whether a functional protein could be synthesized by a cell with the smallest number of possible changes between synthetic and natural methodologies (addition of a single Met residue).

The proteins AcGGH(Hin 139-190) and MGGH(Hin 139-190) were tested under standard conditions with both the Cu(II)-H_2O_2-sodium ascorbate and Ni(II)-monoperoxyphthalic acid systems and compared to the parent protein GGH(Hin 139-190) (Figures 23 and 24). Histograms for the copper reactions are shown in Figure 25. It can be seen that only minimal changes occur at the secondary binding site by the addition of either an acetyl or a methionine residue to the amino terminus. The cleavage pattern at hixL however, changes dramatically in the case of an added methionine, producing moderate cleavage where there was weak diffuse cleavage in the GGH protein. Examination of the autoradiogram for the nickel reactions (Figure 24) shows that upon addition of either an acetyl or a methionine residue to the amino terminus, DNA cleavage is completely abolished. The results for the copper system, at least at the secondary site, indicate that the metal binding domains for all three proteins are in approximately the same location on the DNA. The lack of cleavage observed in the nickel system therefore implies that the electronic nature of the metal center or binding of the metal itself has been altered in these new proteins, negatively affecting the cleavage chemistry.

These results for the Cu(II) chemistry imply that the nature of the oxidizing species has not been altered by the addition of extra residues to the amino terminus, and that it should therefore be possible to place this amino acid sequence at any position along a protein chain and still have an active oxidizing species provided the polypeptide chain can adopt the required conformation. The results for the Ni(II) chemistry indicate the necessity for the amino terminal residue to be part of the ligand. This requirement does not diminish the utility of this functionality since the amino terminal methionine can be removed by normal processing in the organisms, and it is also possible using standard molecular biology techniques to produce a fusion protein which can be cleaved at any particular position liberating any amino terminal residue desired.[35,36] Another interesting sequence to try chemically which would not have any of these problems would be Met-Gly-His.

Figure 23 Autoradiogram of high-resolution denaturing gel of Cu(II)•GGH(Hin 139-190) Cu(II)AcGGH(Hin 139-190) and Cu(II)MGGH(Hin 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Cu•GGH(Hin 139-190) were run for 90 (25 °C). Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 contain $CuCl_2$ (2.5 μM) and GGH(Hin 139-190) (5 μM) followed by sodium ascorbate (1 mM) and hydrogen peroxide (1 mM). Lanes 7 and 8 are identical with lanes 5 and 6 except that they contain $CuCl_2$ (2.5 μM) and AcGGH(Hin 139-190) (5 μM). Lanes 9 and 10 are identical with lanes 5 and 6 except that they contain $CuCl_2$ (2.5 μM) and MGGH(Hin 139-190) (5 μM).

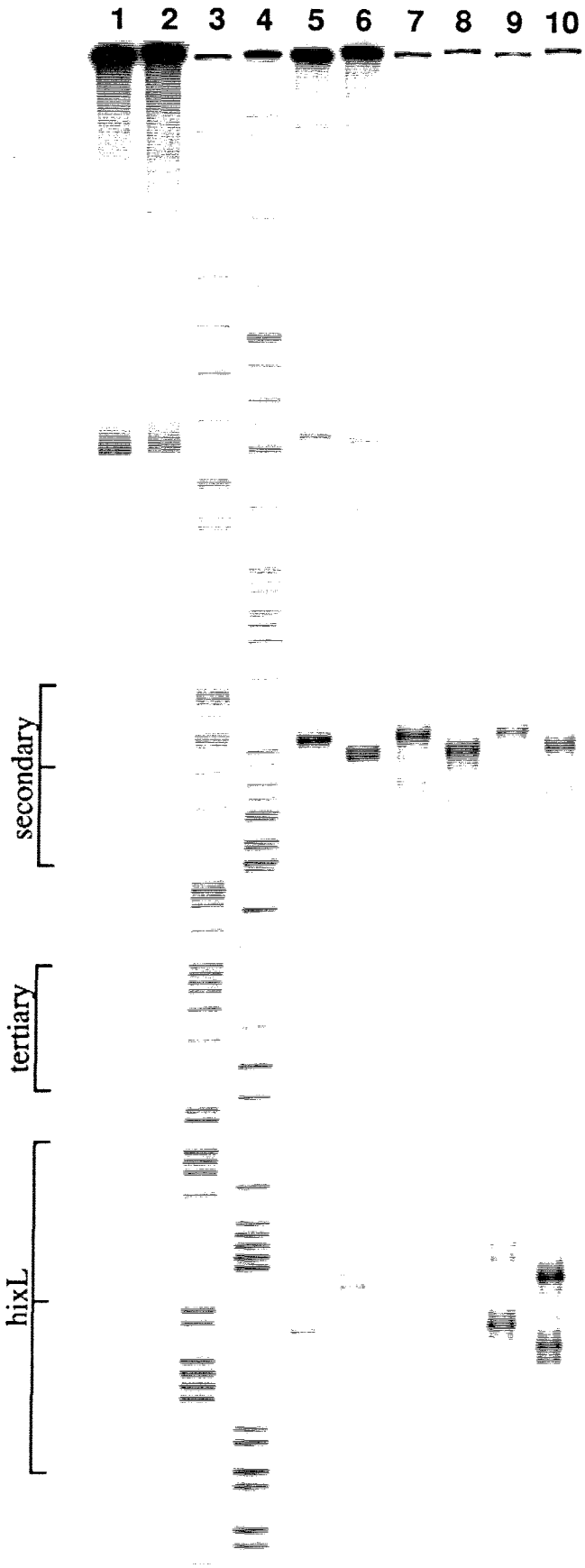
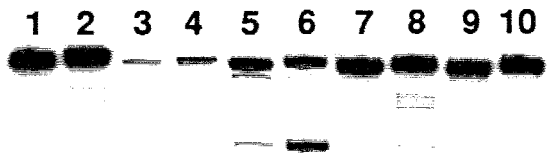
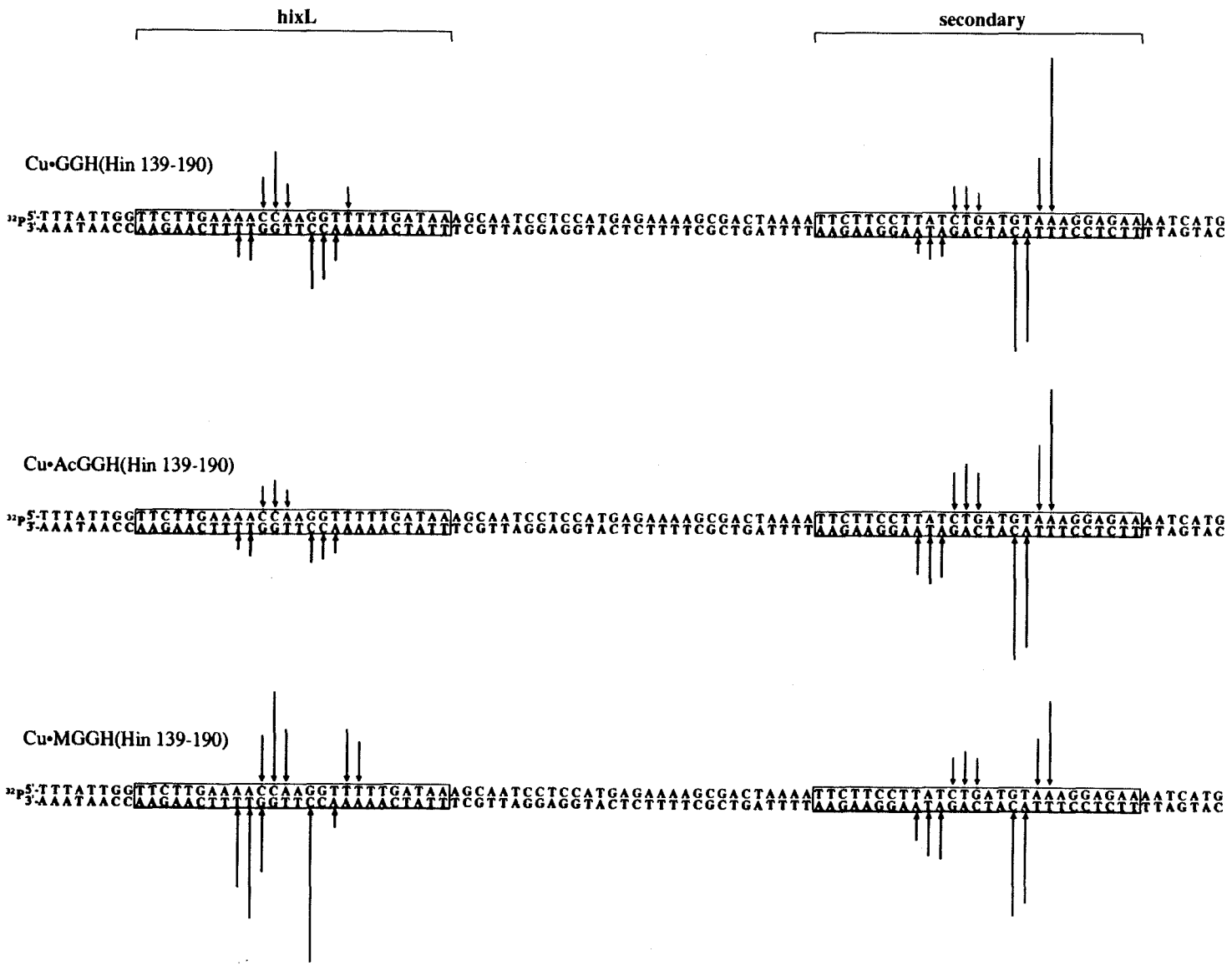

1 2 3 4 5 6 7 8 9 10
secondary
tertiary
hixL

Figure 24 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190), Ni(II)•AcGGH(Hin 139-190), andNi(II)•MGGH(Hin 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Ni•GGH(Hin 139-190) were run for 15 min (25 °C). Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 contain $Ni(OAc)_2$ (5 μM) and GGH(Hin 139-190) (5 μM) followed by monoperoxyphthalic acid, magnesium salt (5 μM). Lanes 7 and 8 are identical to lanes 5 and 6 except they contain $Ni(OAc)_2$ (5 μ M) and AcGGH(Hin 139-190) (5 μM). Lanes 9 and 10 are identical to lanes 5 and 6 except they contain $Ni(OAc)_2$ (5 μ M) and MGGH(Hin 139-190) (5 μM).


1 2 3 4 5 6 7 8 9 10


secondary
tertiary
hixL

Figure 25 Histograms of the data from Figure 23. The sequence left to right represents the data from the bottom to the middle of the gel. Boxes indicate the dimeric hixL and secondary Hin binding sites. Arrows represent the location and extent of cleavage. Top: Cu(II)•GGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 23 lanes, 5 and 6). Middle: Cu(II)•AcGGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 23 lanes, 7 and 8). Bottom: Cu(II)•MGGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 23 lanes, 9 and 10). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.


hixL
secondary
Cu•GGH(Hin 139-190)
Cu•AcGGH(Hin 139-190)
Cu•MGGH(Hin 139-190)

This sequence does not require any special processing, but the Met side chain may interfere with metal binding or cleavage chemistry.

GG(d)H(Hin 139-190). The preferential cleavage observed on one strand of the DNA at each hixL half-site suggests some sort of diastereofacial selectivity in the DNA cleavage reaction. Since the only stereocenter in the metal binding domain of the protein is at the histidine residue, the protein was synthesized with the single change of a d-histidine residue for the normal l-histidine residue in order to to determine whether the cleavage selectivity could be changed. The new GG(d)H(Hin 139-190) was tested under both the Ni(II)-peracid and Cu(II)-H_2O_2-ascorbate reaction conditions and compared to the l-histidine containing protein (Figure 26).

Histograms of the DNA cleavage reactions for the Ni(II)-peracid system and Cu(II)-H_2O_2-ascorbate system are shown in Figures 27 and 28 respectively. In the Ni(II)-peracid system, a number of general observations can be made. First, at all sites studied (hixL, secondary, and tertiary), the same base positions are cleaved for the l and d proteins, and, second, at all sites, the positions of major cleavage have changed from one strand to the other with the change of l to d histidine in the protein. In the Cu(II)-H_2O_2-ascorbate system the differences between l and d histidine are more subtle, but at all sites the location of maximal cleavage has moved out away from the center of the dimeric binding sites. In the Cu(II) system, this change in position of cleavage pattern could simply mean that the Cu(II)•GGH moiety has changed position along the DNA strand. Examination of models of Cu(II)•GGH indicates that if the position of the amide bond between the NH_2 terminal glycine of Hin(139-190) and the histidine residue is held fixed in space, as would be expected in the protein-DNA complex, a change of l to d histidine, holding the coordination geometry constant, changes the directionality that the peptide wraps around the metal atom. This change in directionality would be expected to change the position of a ligand radical species as well as that of a metal-oxo species with respect the DNA.

Figure 26 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190), Ni(II)•GG(d)H(Hin 139-190), Cu(II)•GGH(Hin 139-190), and Cu(II)•GG(d)H(Hin 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Cu•GGH(Hin 139-190) and Ni•GGH(Hin 139-190) were run for 90 and 15 min, respectively (25 °C). Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 contain $Ni(OAc)_2$ (5 μM) and GGH(Hin 139-190) (5 μM) followed by monoperoxyphthalic acid, magnesium salt (5 μM). Lanes 7 and 8 are identical to lanes 5 and 6 except they contain $Ni(OAc)_2$ (5 μ M), GG(d)H(Hin 139-190) (5 μM). Lanes 9 and 10 contain $CuCl_2$ (2.5 μM) and GGH(Hin 139-190) (5 μM) followed by sodium ascorbate (1 mM) and hydrogen peroxide (1 mM). Lanes 11 and 12 are identical with lanes 9 and 10 except that they contain $CuCl_2$ (2.5 μM) and GG(d)H(Hin 139-190) (5 μM).

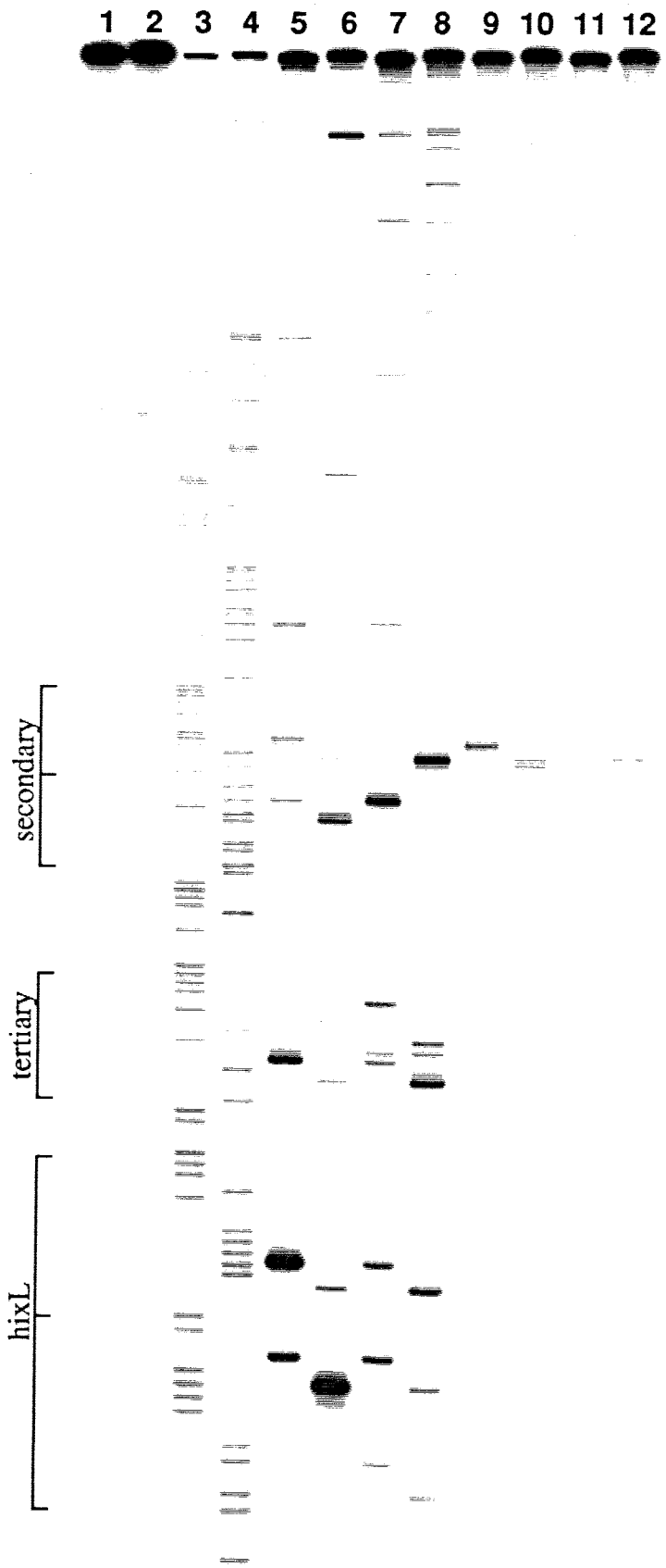

1 2 3 4 5 6 7 8 9 10 11 12
secondary
tertiary
hixL

Figure 27 Histograms of the data from nickel mediated cleavage in Figure 26. The sequence left to right represents the data from the bottom to the middle of the gel. Boxes indicate the dimeric hixL and secondary Hin binding sites. Arrows represent the location and extent of cleavage. Top: Ni(II)•GGH(Hin 139-190) in the presence of monoperoxyphthalic acid (Figure 26 lanes, 5 and 6). Bottom: Ni(II)•GG(d)H(Hin 139-190) in the presence of monoperoxyphthalic acid (Figure 26 lanes, 7 and 8). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.

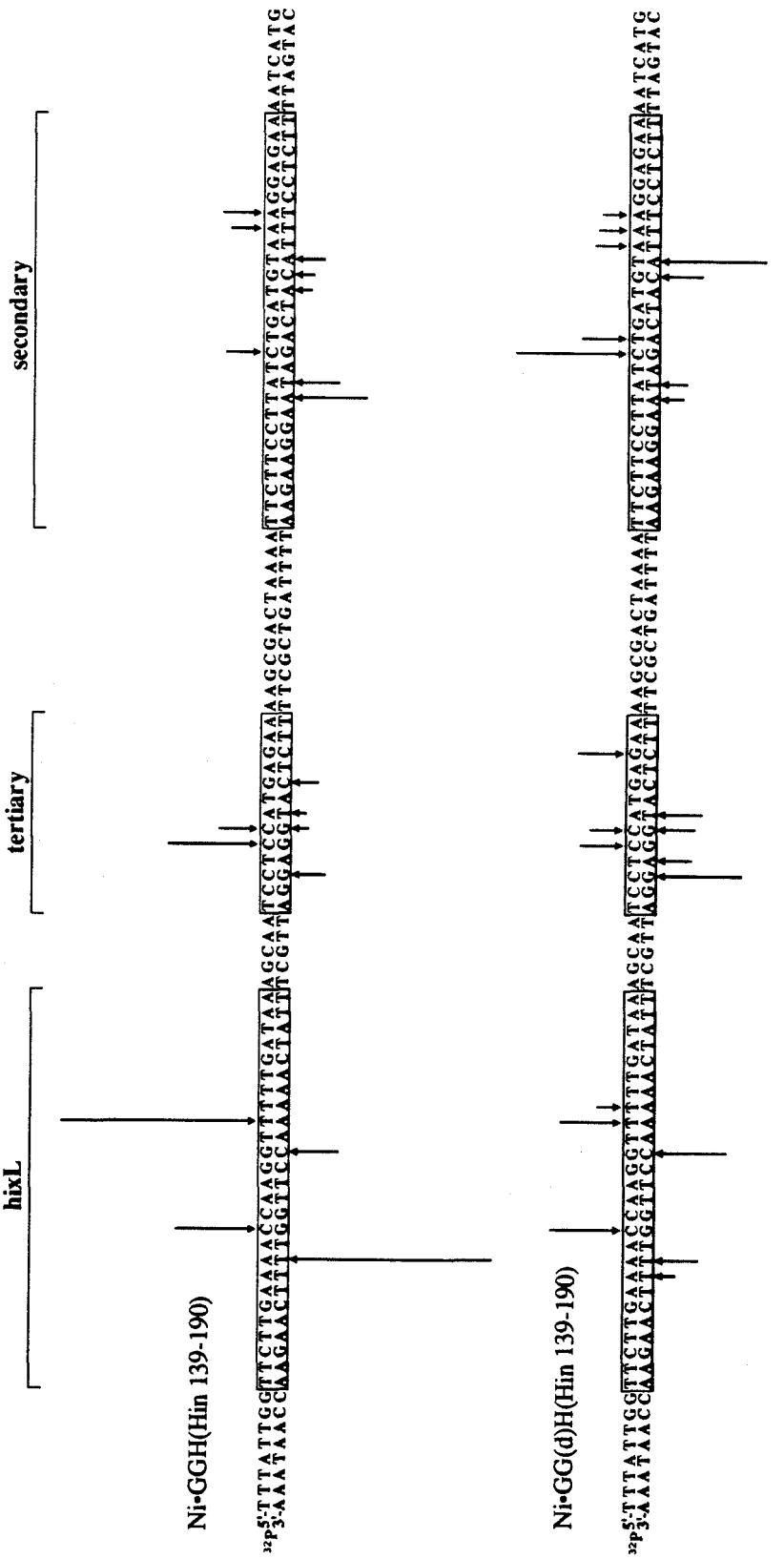
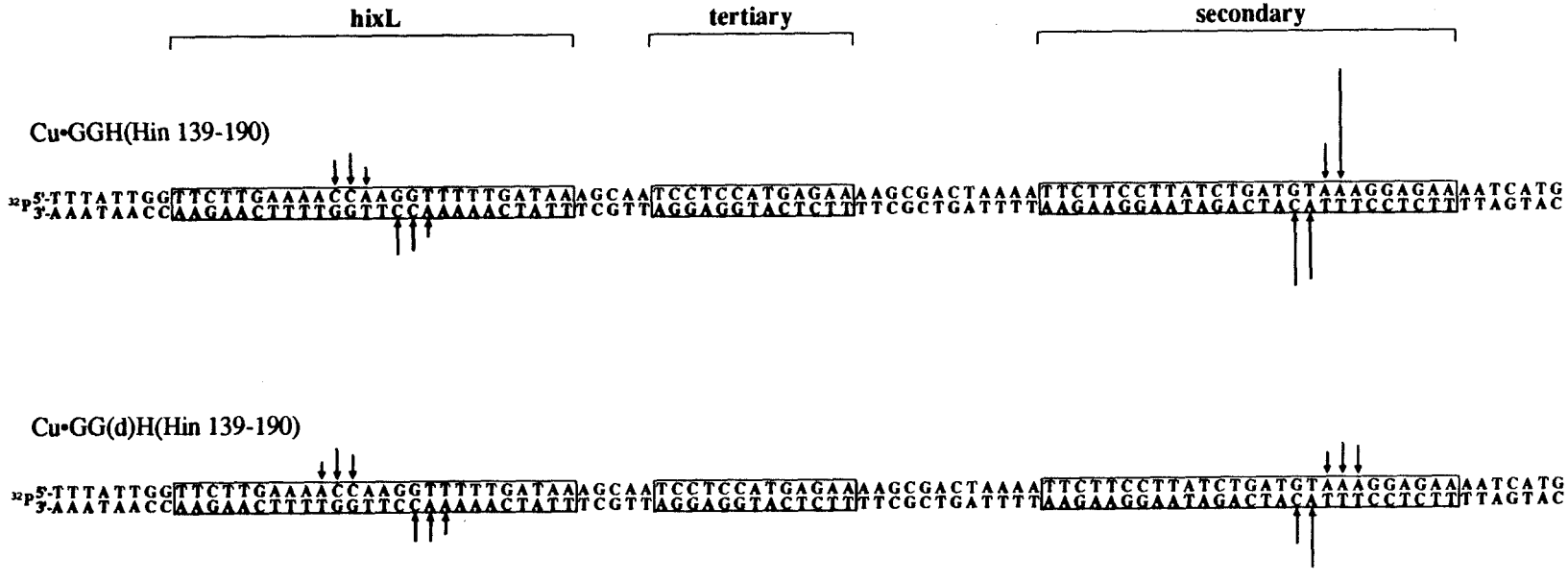

hixL
tertiary
secondary
Ni•GGH(Hin 139-190)
Ni•GG(d)H(Hin 139-190)

Figure 28 Histograms of the copper mediated cleavage data from Figure 26. The sequence left to right represents the data from the bottom to the middle of the gel. Boxes indicate the dimeric hixL and secondary Hin binding sites. Arrows represent the location and extent of cleavage. Top: Cu(II)•GGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 26 lanes, 9 and 10). Bottom: Cu(II)•GG(d)H(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 26 lanes, 11 and 12). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.


hixL
tertiary
secondary
Cu•GGH(Hin 139-190)
Cu•GG(d)H(Hin 139-190)

If the interpretation of a change in position of the metal-ligand species is correct and the nature of the oxidizing species in the nickel system is indeed a metal-oxo, then in order to account for the data seen with Ni(II)•GGH(Hin 139-190) it must be possible to form an oxo species from both faces of the Ni(II)•GGH moiety. This reasoning suggests that the nature of the strand specificity of the cleavage reaction arises because the nickel-oxo species is held rigidly in the groove of DNA and in close proximity to a particular deoxyribose on one strand. The change in position upon going from l to d histidine then alters the complex, such that the nickel-oxo species on the other face is closer to the deoxyribose ring on the opposite strand while the original nickel-oxo species is now farther away from the deoxyribose originally cleaved. This explanation for the change in cleavage pattern seen is only a possibility, and others, such as the metal complex being pushed to one side or the other of the groove on changing from l to d histidine are also reasonable.

AibAibH(Hin 139-190). As an initial attempt to examine the mechanisms by which these metalloproteins cleave the DNA, the protein AibAibH(Hin 139-190), where Aib is α-aminoisobutyric acid, was synthesized. The replacement of α carbon hydrogens with methyl groups enhances the stability of the Cu(III) oxidation state.[37-40] This stabilization can be accounted for in two ways; first the electron donating nature of the α methyl groups is expected to increase the donor strength of the peptide nitrogens, and second, the methyl groups cause a steric block to axial coordination, which is favored for the Cu(II)-d^9 complex and not the Cu(III)-d^8 complex. Furthermore, by replacing the α carbon hydrogen with methyl groups, the pathways for ligand oxidations which require the removal of a α carbon hydrogen are blocked. This combination of stabilizing factors renders the Cu(III) complex of Aib_3 stable for months and in fact, its X-ray crystal structure has been solved.[39] Thus the stabilizing of Cu(III) in the protein AibAibH(Hin 139-190) was studied to examine its effects on the DNA cleavage reaction. The protein AibAibH(Hin 139-190) was tested using the Cu(II)-H_2O_2-ascorbate and Ni(II)-peracid systems with GGH(Hin 139-190) as a standard (Figure 29).

Figure 29 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190), Ni(II)•AibAibH(Hin 139-190), Cu(II)•GGH(Hin 139-190), and Cu(II)AibAibH(Hin 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Cu•GGH(Hin 139-190) and Ni•GGH(Hin 139-190) were run for 90 and 15 min, respectively (25 °C). Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 contain $Ni(OAc)_2$ (5 μM) and GGH(Hin 139-190) (5 μM) followed by monoperoxyphthalic acid, magnesium salt (5 μM). Lanes 7 and 8 are identical to lanes 5 and 6 except they contain $Ni(OAc)_2$ (5 μ M), AibAibH(Hin 139-190) (5 μM). Lanes 9 and 10 contain $CuCl_2$ (2.5 μM) and GGH(Hin 139-190) (5 μM) followed by sodium ascorbate (1 mM) and hydrogen peroxide (1 mM). Lanes 11 and12 are identical with lanes 9 and 10 except that they contain $CuCl_2$ (2.5 μM) and AibAibH(Hin 139-190) (5 μM).

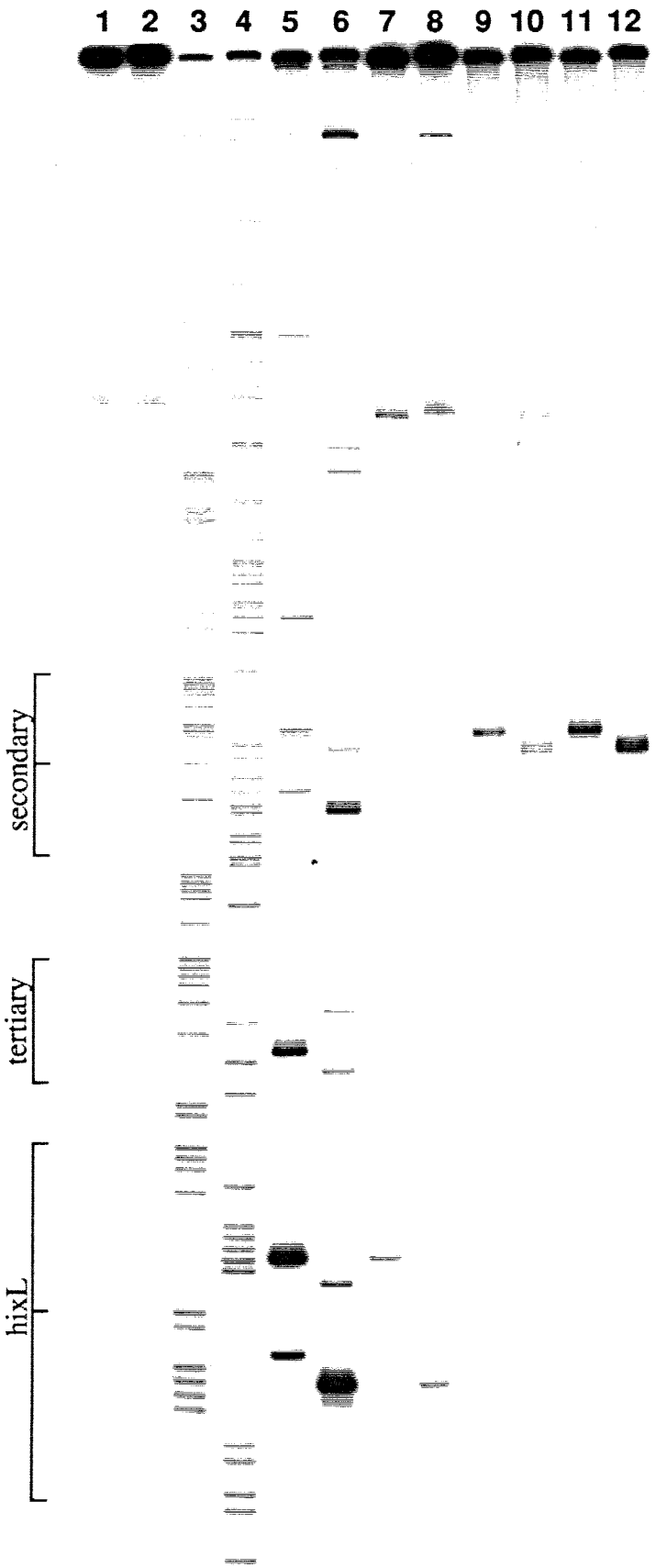

1 2 3 4 5 6 7 8 9 10 11 12
secondary
tertiary
hixL

Figure 30 Histograms of the data from nickel mediated cleavage in Figure 29. The sequence left to right represents the data from the bottom to the middle of the gel. Boxes indicate the dimeric hixL and secondary Hin binding sites. Arrows represent the location and extent of cleavage. Top: Ni(II)•GGH(Hin 139-190) in the presence of monoperoxyphthalic acid (Figure 29 lanes, 5 and 6). Bottom: Ni(II)•GGAibAibH(Hin 139-190) in the presence of monoperoxyphthalic acid (Figure 29 lanes, 7 and 8). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.

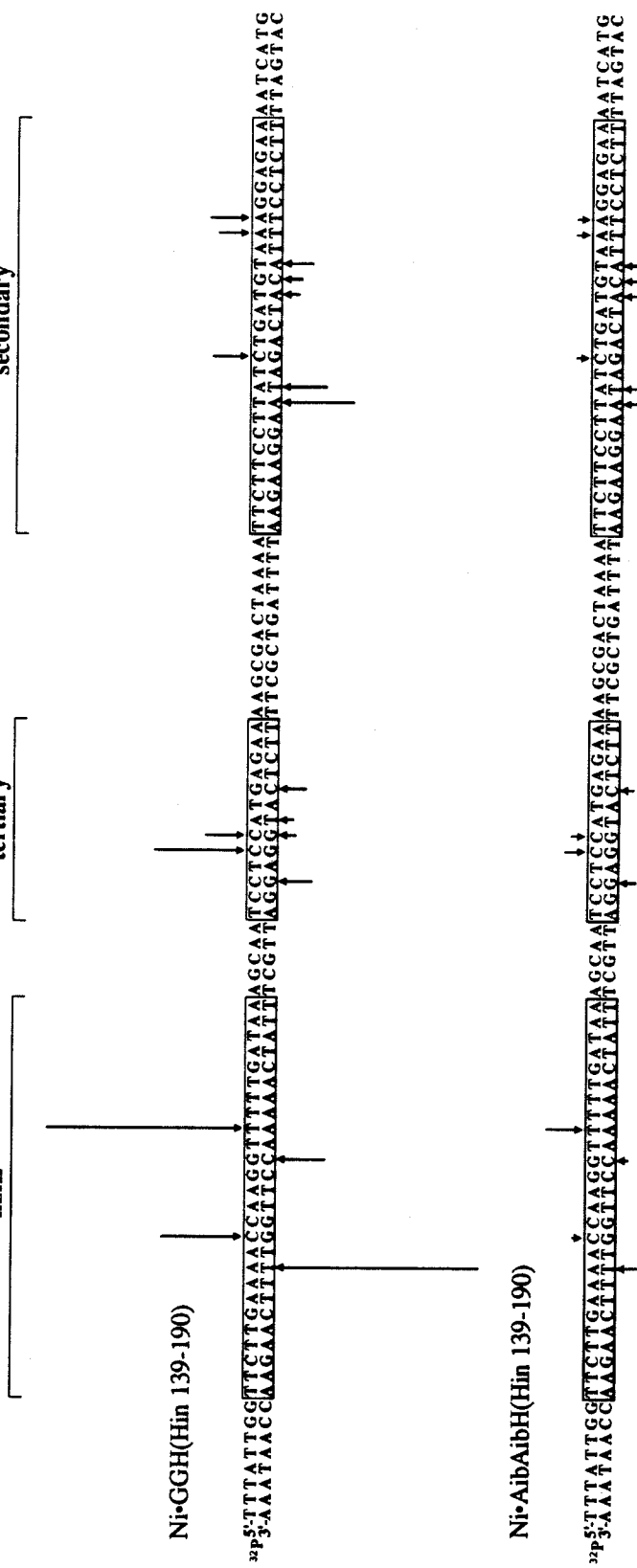
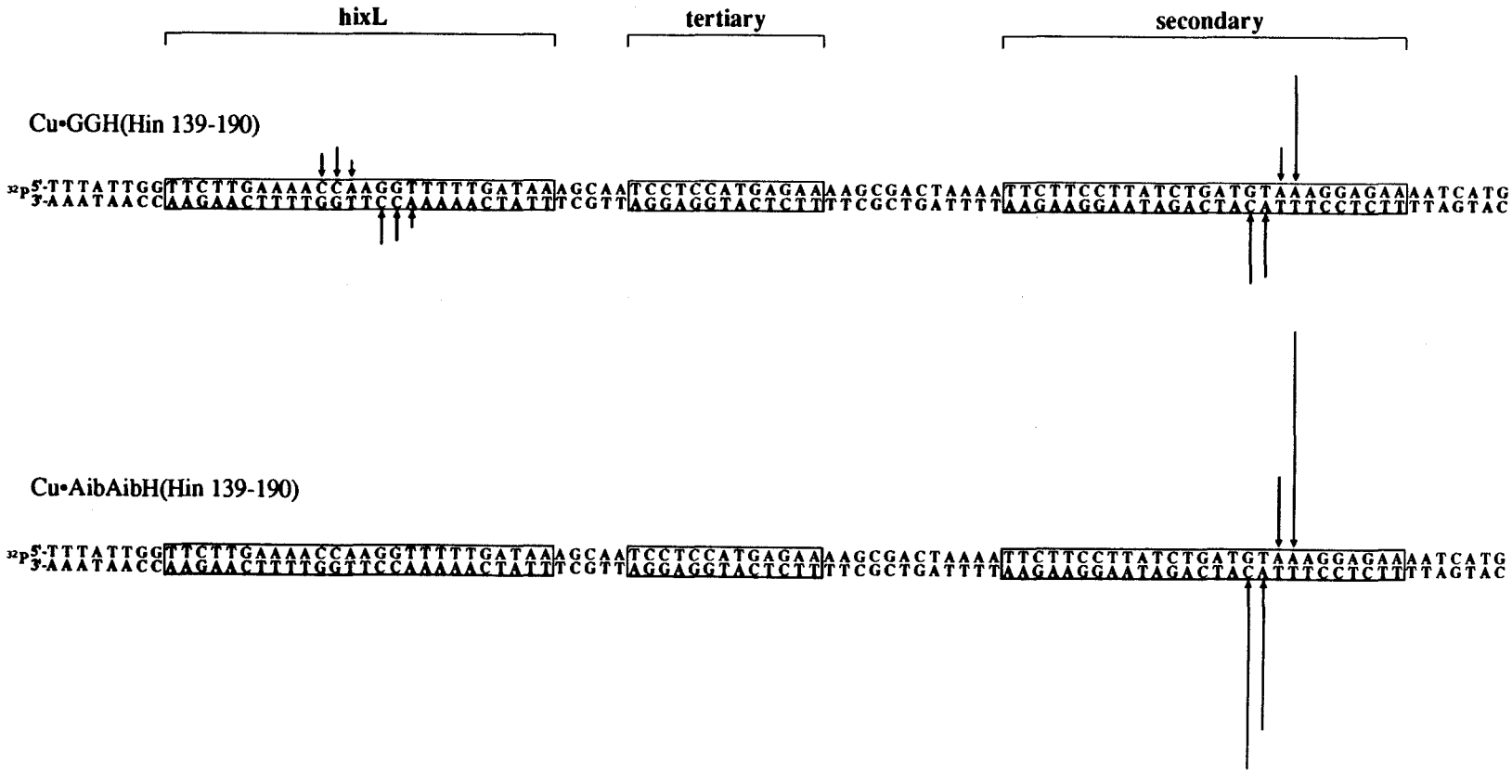

hixL
tertiary
secondary
Ni•GGH(Hin 139-190)
Ni•AibAibH(Hin 139-190)

Figure 31 Histograms of the copper mediated cleavage data from Figure 29. The sequence left to right represents the data from the bottom to the middle of the gel. Boxes indicate the dimeric hixL and secondary Hin binding sites. Arrows represent the location and extent of cleavage. Top: Cu(II)•GGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 29 lanes, 9 and 10). Bottom: Cu(II)•AibAibH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 29 lanes, 11 and 12). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.


hixL
tertiary
secondary
Cu•GGH(Hin 139-190)
Cu•AibAibH(Hin 139-190)

Histograms of the cleavage for the Cu(II)-H_2O_2-ascorbate and Ni(II)-peracid cleavage reactions are shown in Figures 30 and 31, respectively. As is immediately apparent from the autoradiogram (Figure 29, lanes, 5 thru 8), the DNA cleavage reaction of Ni(II)AibAibH(Hin 139-190) has been almost completely stopped. Examination of the histogram (Figure 30) shows that the cleavage patten is identical but significantly reduced in intensity. As can also be seen from the autoradiogram (Figure 29, lanes 9 thru 12) the DNA cleavage reaction of Cu(II)AibAibH(Hin 139-190) is more efficient than that of the glycine containing protein. The histogram (Figure 31) again shows that the cleavage pattern for the Aib and glycine containing proteins are identical, except for a significant increase in the intensity of the cleavage by the Aib containing protein.

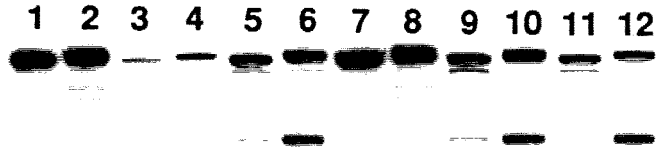
The increased cleavage observed for Cu(II)AibAibH(Hin 139-190) versus Cu(II)GGH(Hin 139-190) is in agreement with that expected from the above discussion, wherein, by replacing the α carbon hydrogens by methyl groups, Cu(III) has been stabilized relative to Cu(II). This data can then be taken as evidence for the intermediacy of a Cu(III) species for the DNA cleavage reaction. An interesting experiment would then be to synthesize a protein with an amino terminal AibAibH where the histidine α carbon hydrogen had been replaced by a methyl group. If the ligand radical mechanism is correct, this ligand should no longer cleave due to the absence of any α carbon hydrogens. The arguments used as to why the copper reaction is better can also be used to explain why the nickel reaction is worse with Aib as compared to glycine. Since the ligand system stabilizes square planar d^8 configurations with no axial coordination,[37-40] in the nickel system this would be stabilizing the Ni(II) state instead of a higher oxidation state, which could slow down the reaction. Another implication of blocking the axial coordination sites is that it may not be possible for the oxygen atom donors to get close enough to deliver the oxygen, or the blockade is so great as to disfavor the formation of an oxo species. Both of these explanations for the lack of cleavage by the Ni(II)-Aib protein could possibly be involved in stopping the cleavage reaction.

AAH(Hin 139-190) and (d)A(d)AH(Hin 130-190). If the two methyl groups in the Aib residues are really responsible for blocking access to the face of the metal complex and stopping formation of an oxo species, then it should be possible to block only a single face of the complex and make the reaction more facially selective. To this end, two proteins were synthesized AAH(Hin 139-190) and (d)A(d)AH(Hin 139-190) each of which will have methyl groups on opposite faces of the metal complex. These proteins were tested both with the Ni(II)-peracid and Cu(II)-H_2O_2-ascorbate cleavage systems (Figures 32 and 33). As can be clearly seen for the Ni(II)-peracid system (Figure 32) either alanine or d-alanine produces cleavage identical to the original glycine containing protein. It seems clear then that the methyl groups do not cause a steric blockade of the metal centers. Histograms of the cleavage by the Cu(II) proteins are shown in Figure 34. In the case of the alanine containing protein, the cleavage is enhanced on one strand, and in the case of d-alanine the cleavage seems diminished on both strands. These results are not easily interpreted and it is clear that a more detailed three-dimensional picture of the protein-DNA complex is necessary.

Mechanistic Studies of Ni(II)•GGH(Hin 139-190)

End Product Analysis. Due to the cleaner and more efficient nature of the Ni(II)•GGH-peracid reaction attempts were made to gain a more detailed mechanistic understanding of the reaction. As was shown earlier, the DNA cleavage reaction mediated by Ni(II)•GGH(Hin 139-190) and an oxygen atom source (e.g., hydrogen peroxide, monoperoxyphthalic acid, or iodosylbenzene), produced a large amount of base labile lesions. Because the DNA cleavage reaction has also been shown to be oxidative degradation of the deoxyribose backbone producing 3' and 5' phosphate ends upon workup with n-$BuNH_2$, it seemed likely that one or more specific hydrogen atoms on the ribose ring were being abstracted. Of the hydrogen atoms available for abstraction in the minor groove (C-1', C-2', C-4', or C-5'), little is known about abstraction at C-2',[41] and

Figure 32 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190), Ni(II)•AibAibH(Hin 139-190), Ni(II)•AAH(Hin 139-190), and Ni(II)•(d)A(d)AH(Hin 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Ni•GGH(Hin 139-190) were run for 15 min (25 °C). Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 contain $Ni(OAc)_2$ (5 μM) and GGH(Hin 139-190) (5 μM) followed by monoperoxyphthalic acid, magnesium salt (5 μM). Lanes 7 and 8 are identical to lanes 5 and 6 except they contain $Ni(OAc)_2$ (5 μ M) and AibAibH(Hin 139-190) (5 μM). Lanes 9 and 10 are identical to lanes 5 and 6 except they contain $Ni(OAc)_2$ (5 μ M) and AAH(Hin 139-190) (5 μM). Lanes 11 and 12 are identical to lanes 5 and 6 except they contain $Ni(OAc)_2$ (5 μ M) and (d)A(d)AH(Hin 139-190) (5 μM).


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

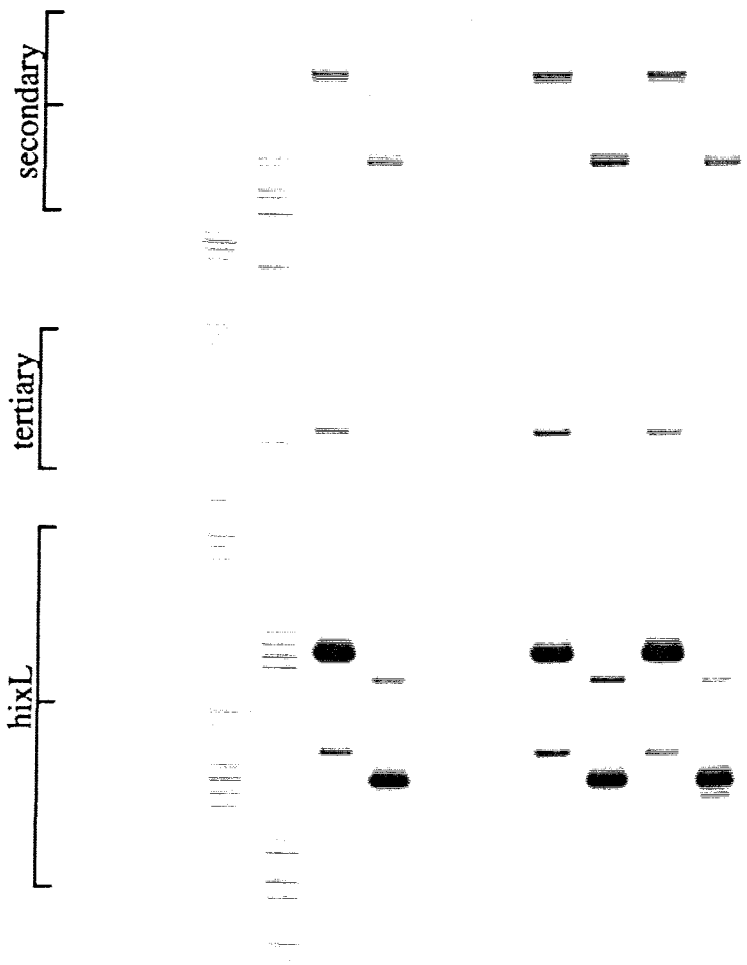

secondary
tertiary
hixL

Figure 33 Autoradiogram of high-resolution denaturing gel of Cu(II)•GGH(Hin 139-190) Cu(II)AibAibH(Hin 139-190) Cu(II)AAH(Hin 139-190), and Cu(II)•(d)A(d)AH(Hin 139-190) cleavage of a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions for Cu•GGH(Hin 139-190) were run for 90 (25 °C). Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Odd-numbered lanes and even-numbered lanes contain 5' and 3' end-labeled DNA, respectively. Lanes 1 and 2 are DNA control lanes. Lanes 3 and 4 are A specific sequencing reactions. Lanes 5 and 6 contain $CuCl_2$ (2.5 μM) and GGH(Hin 139-190) (5 μM) followed by sodium ascorbate (1 mM) and hydrogen peroxide (1 mM). Lanes 7 and 8 are identical with lanes 5 and 6 except that they contain $CuCl_2$ (2.5 μM) and AibAibH(Hin 139-190) (5 μM). Lanes 9 and 10 are identical with lanes 5 and 6 except that they contain $CuCl_2$ (2.5 μM) and AAH(Hin 139-190) (5 μM). Lanes 11 and 12 are identical with lanes 5 and 6 except that they contain $CuCl_2$ (2.5 μM) and (d)A(d)AH(Hin 139-190) (5 μM).

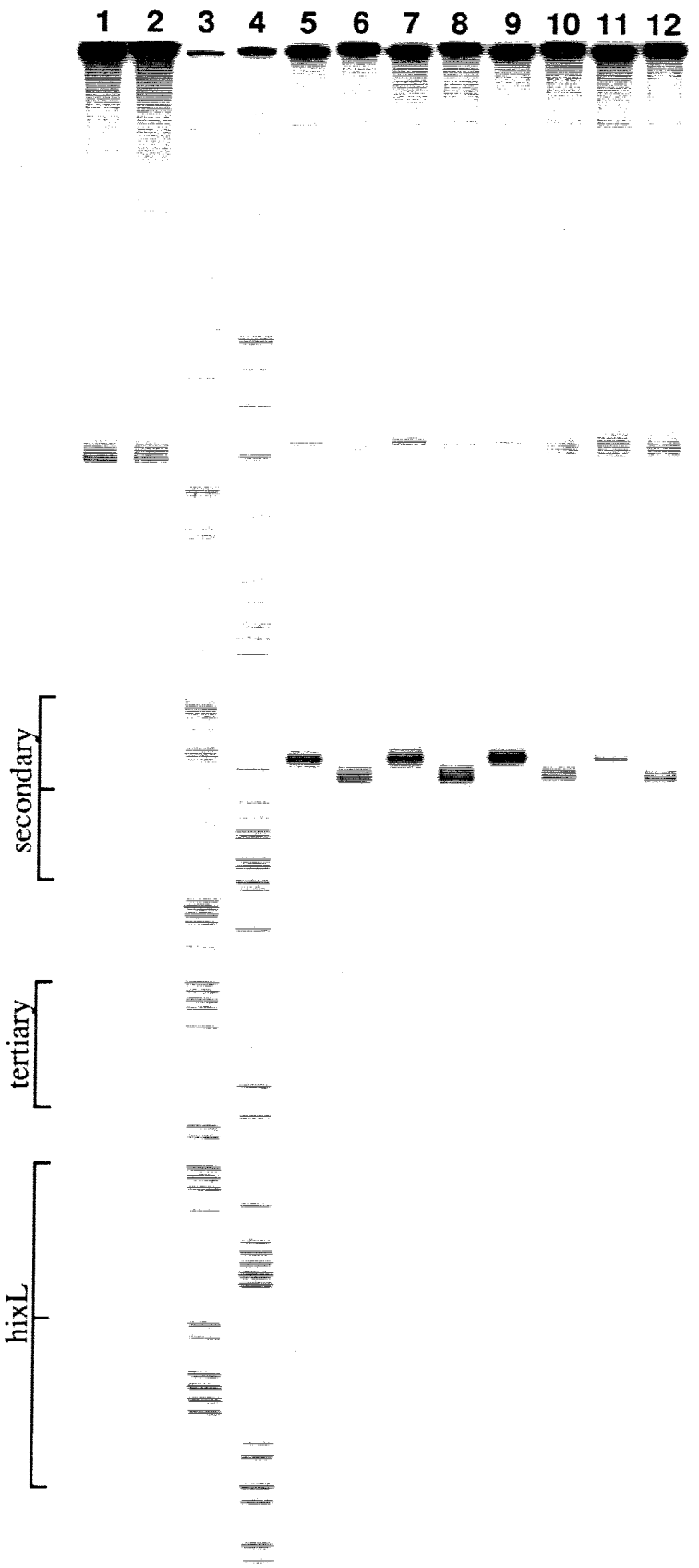
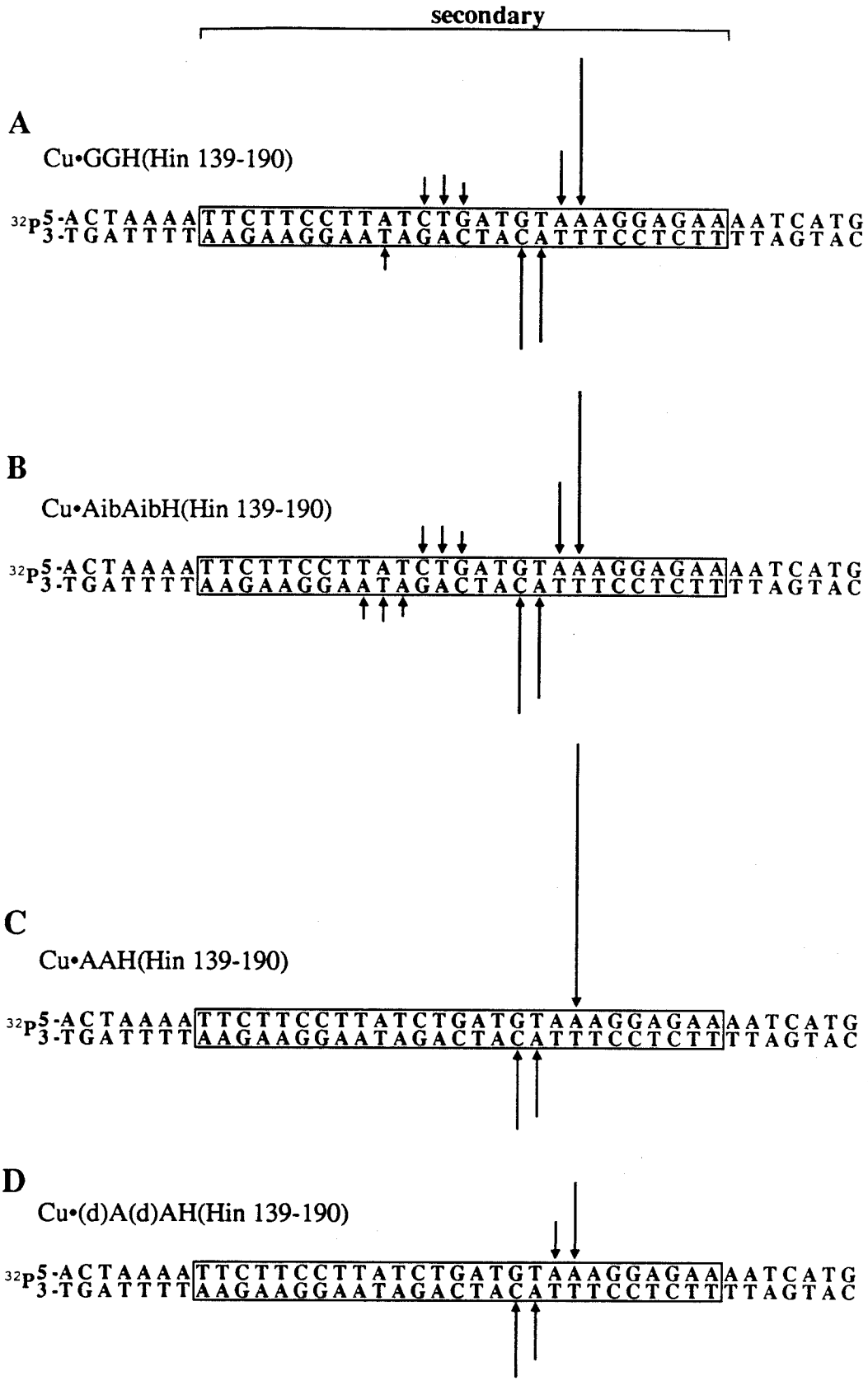

1 2 3 4 5 6 7 8 9 10 11 12
secondary
tertiary
hixL

Figure 34 Histograms of the copper mediated cleavage data from Figure 33. The sequence left to right represents the data from the bottom to the middle of the gel. Boxes indicate the dimeric hixL and secondary Hin binding sites. Arrows represent the location and extent of cleavage. A) Cu(II)•GGH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 33 lanes, 5 and 6). B) Cu(II)•AibAibH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 33 lanes, 7 and 8). C) Cu(II)•AAH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 33 lanes, 9 and 10). B) Cu(II)•(d)A(d)AH(Hin 139-190) in the presence of sodium ascorbate and hydrogen peroxide (Figure 33 lanes, 11 and 12). Extent of cleavage was determined by densitometric analysis of the gel autoradiogram.


secondary
A
Cu•GGH(Hin 139-190)
32P5-ACTAAAATTCTTCCTTATCTGATGTAAAGGAGAAAATCATG
3-TGATTTTAAGAAGGAATAGACTACATTTCCTCTTTTAGTAC
B
Cu•AibAibH(Hin 139-190)
32P5-ACTAAAATTCTTCCTTATCTGATGTAAAGGAGAAAATCATG
3-TGATTTTAAGAAGGAATAGACTACATTTCCTCTTTTAGTAC
C
Cu•AAH(Hin 139-190)
32P5-ACTAAAATTCTTCCTTATCTGATGTAAAGGAGAAAATCATG
3-TGATTTTAAGAAGGAATAGACTACATTTCCTCTTTTAGTAC
D
Cu•(d)A(d)AH(Hin 139-190)
32P5-ACTAAAATTCTTCCTTATCTGATGTAAAGGAGAAAATCATG
3-TGATTTTAAGAAGGAATAGACTACATTTCCTCTTTTAGTAC

reaction at C-1' does not produce base labile lesions.[42] Of the two remaining possibilities, C-4' or C-5', either would produce base labile lesions,[31,43] however the C-5' abstraction leads to strand scission with a sugar product at the 5' end prior to base workup.[44] That is, C-5' abstraction leads to strand scission immediately upon abstraction with only the end-products changing upon base workup.

The hydrogen atom abstraction chemistry at the C-4' position of the deoxyribose ring is the best studied of any of the positions. The naturally occurring antibiotic Fe•Bleomycin is known to abstract the C-4' hydrogen atom from duplex DNA leading to two types of DNA lesions.[45,46] In the presence of oxygen the DNA is cut at neutral pH, and in the absence of oxygen strong base is required for strand scission. Furthermore, the nature of the base used in the workup for anaerobic conditions can have an effect on the identity of the DNA end-products.[31] Schemes for both pathways are shown in Figure 35. In the aerobic pathway (a), the 3' DNA end-product is a phosphoglycolate group. In the anaerobic pathway (b), the 3' end-product upon base workup depends upon the base used. With NaOH two products are observed, one containing the modified sugar fragment, and the other a phosphate, with n-$BuNH_2$, only the phosphate end product is seen.

The hydrogen abstraction chemistry at the C-5' position is also well characterized as it is one of the sites of attack of the neocarzinostatin chromophore.[43,44] A scheme for the aerobic cleavage reaction following abstraction of a C-5' hydrogen atom is shown in Figure 36. In this pathway trapping of the carbon centered radical by oxygen leads to a 5' ribose product and a 3' phosphate. It has been shown that this 5' ribose product migrates on high resolution polyacrylamide gels slower than the corresponding 5' phosphate product. Treatment of the ribose product with base then liberates the 5' phosphate and a ribose fragment.

The 3'-DNA end-products for the Fe-Bleomycin mediated cleavage of DNA in the presence of oxygen were analyzed on a 20% polyacrylamide gel and are shown in Figure 37. Lane 3 shows the DNA cleavage observed with no base workup, lane 4 shows the

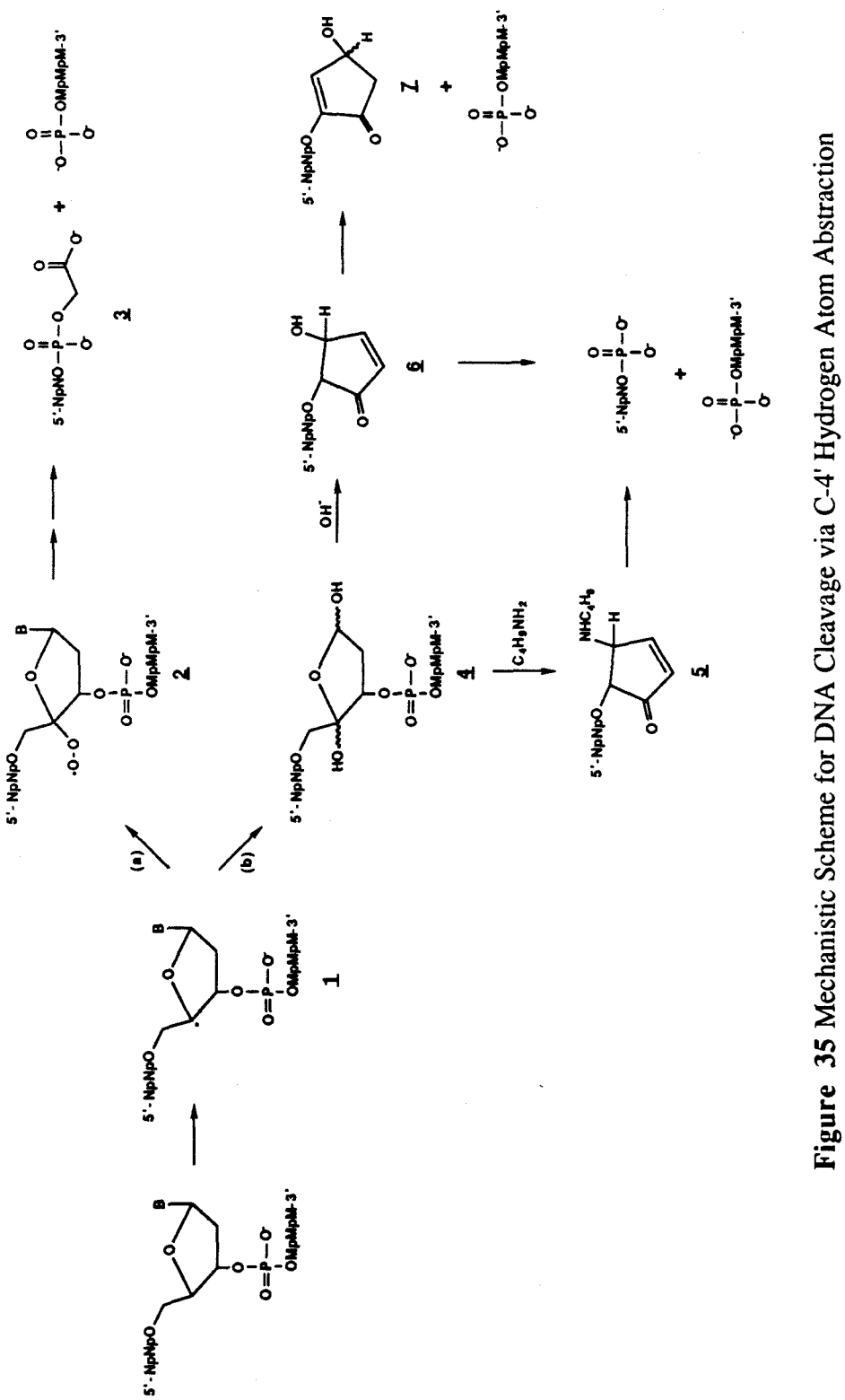


Figure 35 Mechanistic Scheme for DNA Cleavage via C-4' Hydrogen Atom Abstraction

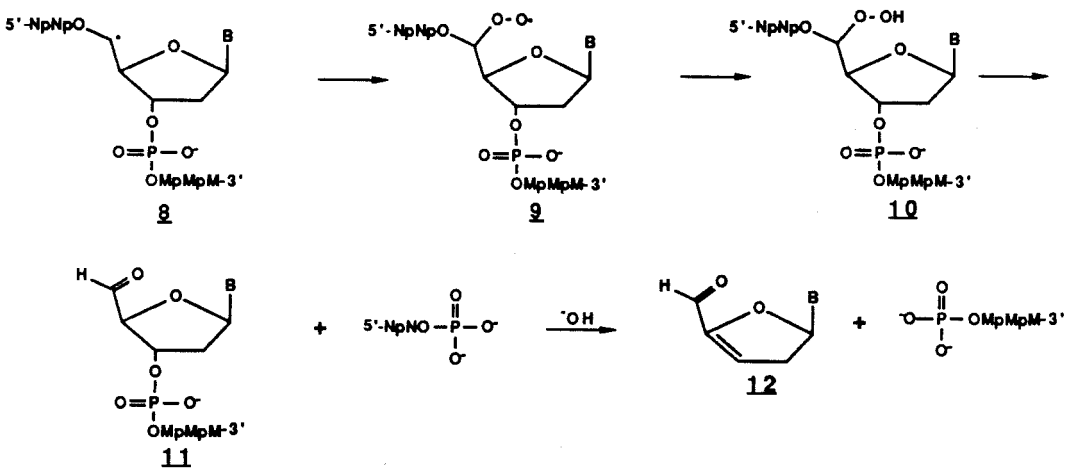


Figure 36 Mechanistic scheme for DNA cleavage via a C-5' hydrogen atom abstraction.

same cleavage reaction when worked up with NaOH, and lane 5 shows the reaction when worked up with n-$BuNH_2$. The band in lane 3 is identified as a 3' phosphoglycolate as expected for the aerobic cleavage reaction of bleomycin. Upon NaOH workup (lane 4) two new bands appear, both with slower electrophoretic mobility than the original product. The band directly above the phosphoglycolate is identified as phosphate in analogy to the MPE•Fe reacion[15] shown earlier, and the slowest band is identified as the ribose product 7 expected from Figure 35. Upon workup with n-$BuNH_2$ (lane 5), as expected from the mechanism only one new product is seen as compared to the reaction that had not been treated with base and can be assigned as 3' phosphate in analogy to lane 4.

The 3' DNA end-products of the Ni(II)•GGH(Hin 139-190) reaction mediated with monoperoxyphthalic acid with no workup, NaOH, and n-$BuNH_2$ treatment were also analyzed on a 20% polyacrylamide gel and compared to an MPE•Fe(II) reaction (Figure 38). As can be seen, without any base workup very little cleavage is observed (Figure 38, lanes 3 and 4). That a small amount of product is seen, however is interesting. The faster moving band is identified as 3' phosphate by comparison with the MPE•Fe(II) reaction.[15] The identity of the slower moving band is not known, but in analogy to the products expected for C-5' abstraction, it is likely to be the expected ribose product **11** (Figure 36). Upon treatment with NaOH, two strong bands are produced. The lower of the two bands

Figure 37 Autoradiogram of a high resolution gel of Fe•Bleomycin cleavage of a 5' end-labeled oligonucleotide (see Figure 40 for oligonucleotide sequence). Reactions conditions were 5 μM Fe•Bleomycin, 1 mM DTT, 20 mM NaCl, 10 mM Tris•HCl, pH 7.8, 100 μM (in base pair) calf thymus DNA, and ≈15,000 cpm of ^{32}P end-labeled oligonucleotide.Reactions were run at 25 °C for 60 minutes. Base treatment consisted of heating at 90 °C for 30 min in the presence of 0.1 M of the appropriate base. Cleavage products were analyzed on a 20%, 1:20 crosslinked, 50% urea denaturing polyacrylamide gel, 0.4 mm thick. Lane 1 is intact DNA. Lane 2 is a Maxam-Gilbert chemical sequencing G lane. Lanes 3,4, and 5 are Fe•Bleomycin reactions which have been treated with no base, NaOH, and n-$BuNH_2$, respectively.

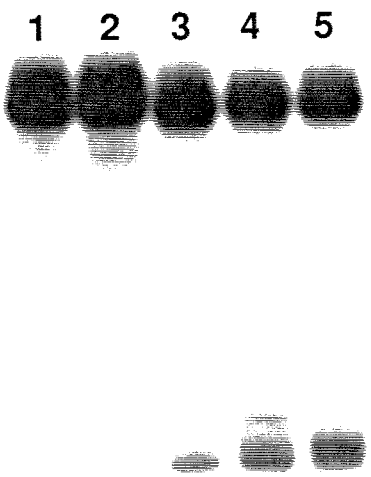
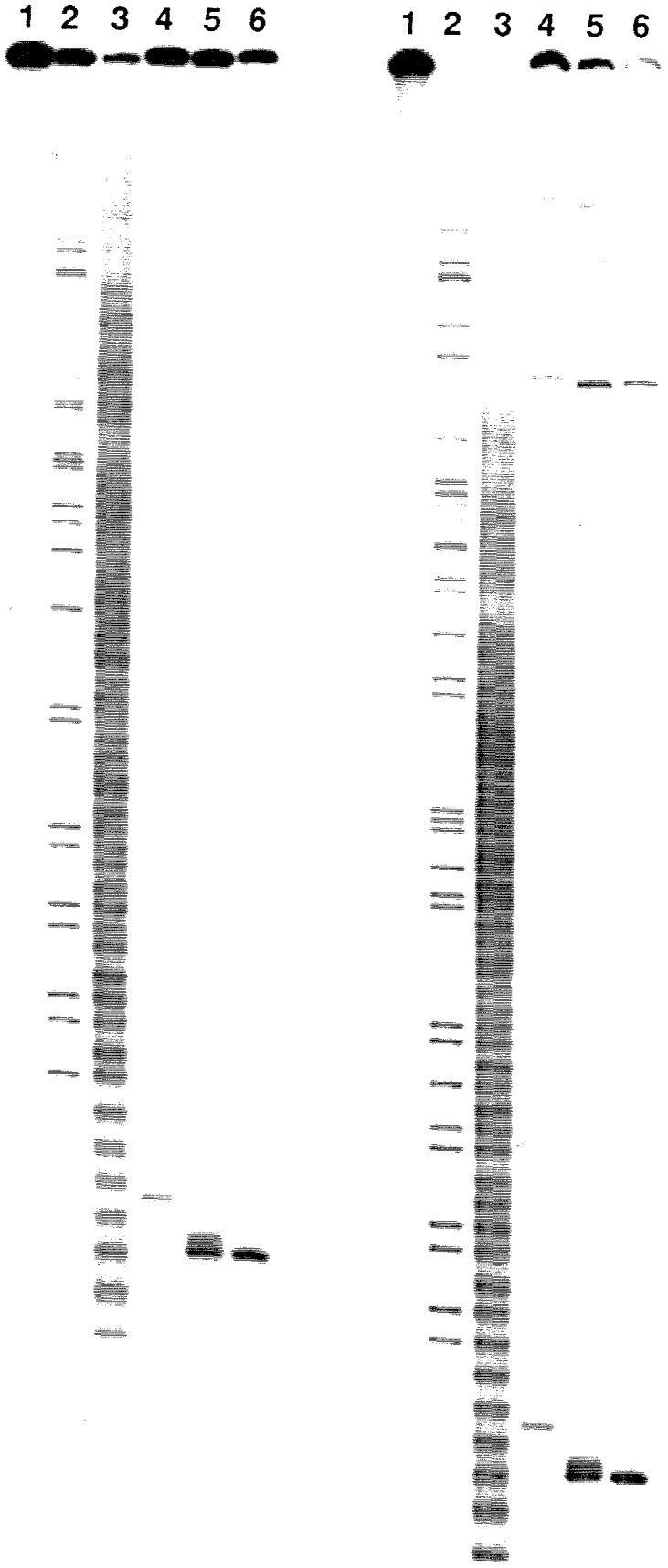

1
2
3
4
5

Figure 38 Autoradiogram of a high resolution gel of Ni(II)•GGH(Hin 139-190) cleavage of 5' end-labeled restriction fragments (EcoRI/RsaI and HindIII/RsaI fragments from pDPM12). Reactions conditions were 5 μM Ni(II)•GGH(Hin 139-190), 5 μM monoperoxyphthalic acid, 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in base pair) calf thymus DNA, and ≈15,000 cpm of ^{32}P end-labeled restriction fragments. Reactions were run at 25 °C for 15 minutes. Base treatment consisted of heating at 90 °C for 30 min in the presence of 0.1 M of the appropriate base. Cleavage products were analyzed on a 20%, 1:20 crosslinked, 50% urea denaturing polyacrylamide gel, 0.4 mm thick. Lane 1 of each fragment is intact DNA. Lane 2 of each fragment is a Maxam-Gilbert chemical sequencing G lane. Lanes 3,4, and 5 of each fragment are Ni(II)•GGH(Hin 139-190) reactions which have been treated with no base, NaOH, and n-$BuNH_2$, respectively.


1 2 3 4 5 6
1 2 3 4 5 6

comigrates with the phosphate band, and the upper band is tentatively identified as the ribose product 7 from Figure 35. Treatment with n-$BuNH_2$ produced only a single band that has previously been identified as phosphate. The similarity in the products seen between the Ni(II)•GGH(Hin139-190)-peracid chemistry and the known chemistry of anaerobic Fe•Bleomycin reaction, along with the absence of any 3' phosphoglycolate, strongly imply that the Ni(II)•GGH(Hin139-190) mediated strand scission occurs predominantly by hydroxylation of the 4' position of the deoxyribose ring, although a small percentage could also be occurring via C-5' abstraction.

Although the above evidence strongly implies that the major site of attack on the deoxyribose ring is the C-4' position, it does not rule attack at the other positions, nor is it direct evidence for C-4' hydroxylation. It is interesting to note that no 3' phosphoglycolate is observed in the Ni(II)•GGH(Hin139-190)-peracid chemistry, even though the reactions are done under aerobic conditions. Because the end-products imply that the sugar modification is a formal hydroxylation at the 4' position, one possible explanation for this result is a radical recombination mechanism (Figure 39). In this mechanism, a Ni-oxo species would abstract a hydrogen atom from the 4' position of the deoxyribose ring forming a radical pair, and before the carbon centered radical could be trapped with oxygen from the solution a recombination event would occur, thus leading to 4' hydroxylation. Another possible mechanism for the formation of a 4' hydroxylated position is that after formation of the radical pair an electron transfer from the carbon to the high valent nickel would occur producing a carbocation which could then be trapped by solvent to produce the hydroxylated product. These mechanisms are identical to those proposed for Fe•Bleomycin under anaerobic conditions, wherein an Fe-oxo species effects the chemistry.[45,46]

Olefin Epoxidation. The postulation of a high valent nickel-oxygen intermediate responsible for hydrogen atom abstraction is not unreasonable in the light of recent results from other laboratories. Ni(II) complexes of ligand systems such as cyclam and salen

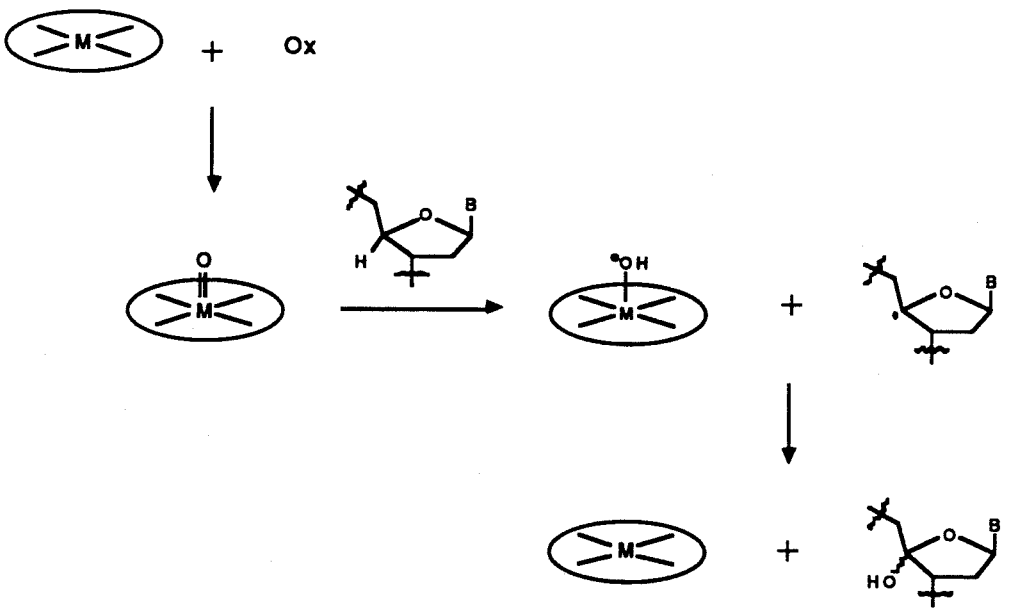


Figure 39 Possible mechanistic scheme for the hydroxylation of the C-4' position of the deoxyribose ring by a metal oxo species.

(Figure 39) have recently been shown to be capable of epoxidizing olefins and abstracting hydrogen atoms from alkanes using iodosylbenzene as the oxygen source.[47-50] Evidence in these studies has been presented for the intermediacy of a high valent Ni(III)-oxo species which is responsible for this chemistry. As an initial test of whether the Ni(II)GGH(Hin 139-190) oxidizing species is a metal-oxo, the tripeptide, GGHa, was synthesized and examined for its ability to epoxidize olefins as the Ni(II) complex. Preliminary results

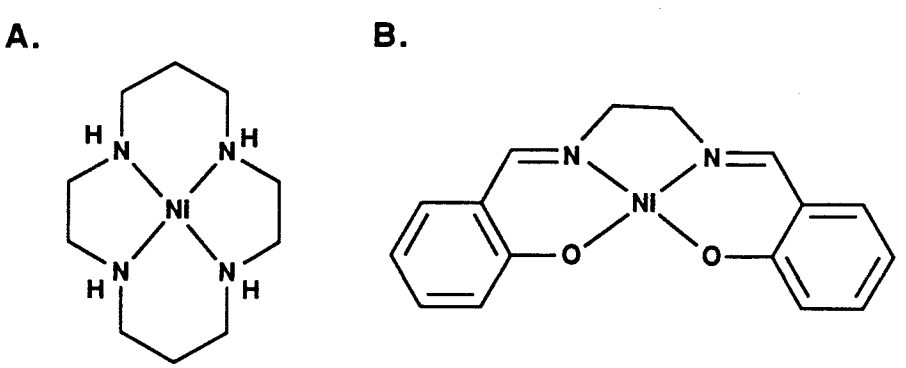


Figure 40 Nickel complexes of A) cyclam and B) salen.

cate that the reaction of Ni(II)•GGHa with cis-stilbene or norbornene using iodosylbenzene as the oxidant produced epoxide, as well as other oxidation products (data not shown). These results suggest that a nickel-oxo species is likely involved in the DNA cleavage chemistry observed for Ni(II)•GGH(Hin 139-190).

Deuterium Isotope Effects. In the mechanistic scheme discussed above, wherein a nickel-oxo species abstracts the C-4' hydrogen atom from the deoxyribose ring and a C-4' hydroxylated species is produced, the hydrogen atom abstraction step is likely to be rate limiting.[46] In an attempt to locate the exact positions on the deoxyribose ring that are attacked by the Ni(II)•GGH(Hin 139-190) oxidizing species and to verify that indeed the C-4' position is attacked, deuterium isotope effect studies were undertaken.

In order to obtain an isotope effect for a DNA cleavage reaction, two design criteria must be met; first, it must be possible to incorporate a specifically deuterated nucleotide at a specific position along the DNA strand, and second, a second cleavage band must be present on the same DNA strand which does not have a deuterium atom present to be used as a reference. In the case of Ni(II)•GGH(Hin 139-190) reacting at the hixL site the second criteria is not met. The sites of strong cleavage are located on opposite strands of the DNA due to the head to head orientation of the peptides. In order to overcome this problem an oligonucleotide was designed which consisted of two hixL half sites positioned head to tail. The use of a synthetic oligonucleotide also meets the first criteria, since a synthetic nucleotide can be incorporated specifically at any one or more positions. The exact sequence of the oligonucleotide chosen along with the position of incorporation of the deuterated nucleotide is shown in Figure 41. Nucleotides containing a deuterium atom at the C-1', C-2', C-4', and C-5' positions were each incorporated into an oligonucleotide of this composition (Figure 41).[51]

Each of the 5 oligonucleotides (containing all proteo, or C-1', C-2', C-4', or C-5' deuterated thymine at the appropriate position) was specifically labeled with ^{32}P at the 5' end and hybridized with an equimolar amount of the corresponding opposite strand to

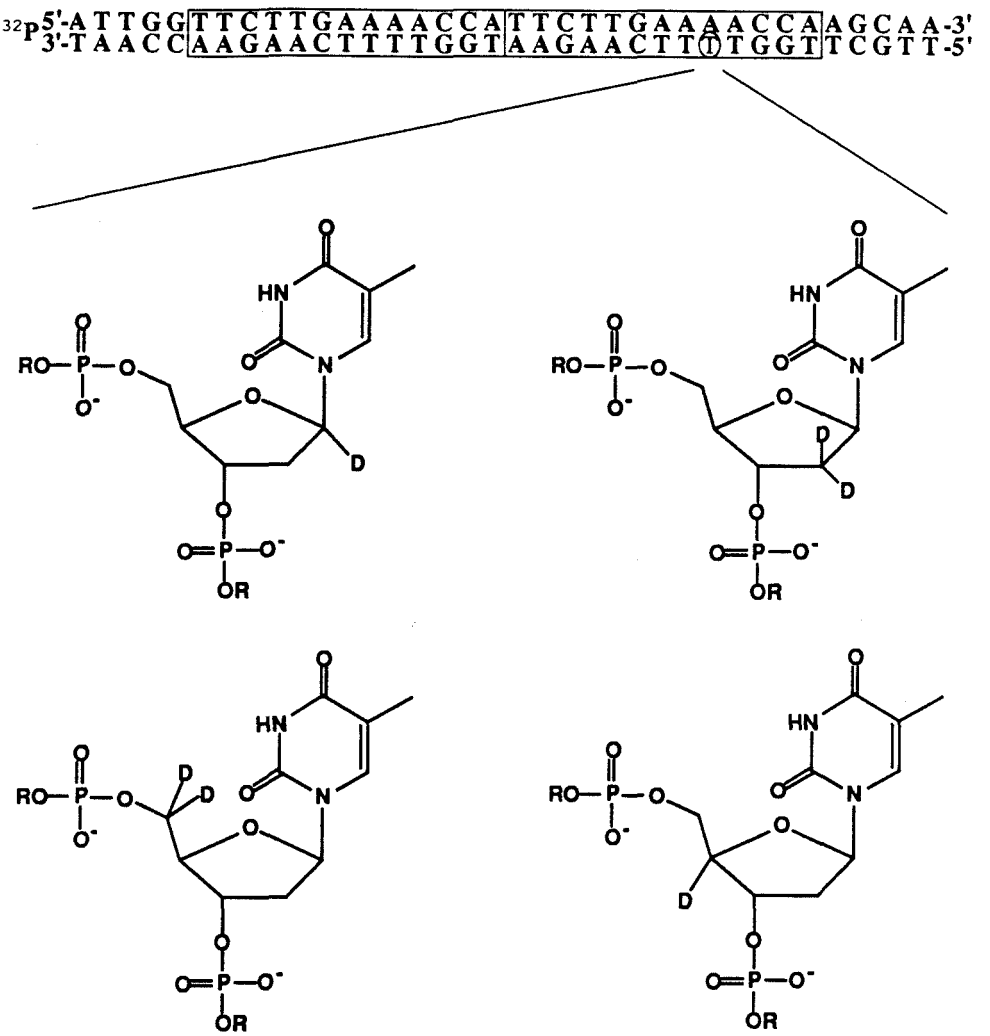


Figure 41 Oligonucleotide used in the deuterium isotope experiments. The oligonucleotide contains the 13 base pair left side of hixL (5'-TTCTTGAAAACCA-3') repeated twice in a head to tail fashion. In the bottom of the Figure are shown the 4 different deuterated thymine residues that were inserted at the highlighted position.

produce the duplex DNA for the isotope experiment. Ni(II)•GGH(Hin 139-190) in the presence of monoperoxyphthalic acid was allowed to react with the DNA to produce a total of 3-5% strand scission upon workup with n-$BuNH_2$. This low conversion prevents any single strand from being cut more than one time, and thus changing the ratio of the two cleavage sites. An example of the cleavage observed with each different oligonucleotides is shown in Figure 42. From the densitometric traces (Figure 43), a clear isotope effect can

Figure 42 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190) cleavage of 5' ^{32}P labeled oligonucleotides containing all proteo, C-1', C-2', C-4', or C-5' deuterated thymines at the sites of attack for the nickel-protein (see Figure 40 for sequence). Reaction conditions were 5 μM Ni(II)•GGH(Hin 139-190), 1 μM monoperoxyphthalic acid, 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in base pair) calf thymus DNA, and ≈15,000 cpm of end-labeled oligonucleotide. Cleavage reactions were run for 15 min at 25 °C and worked up by treatment with 0.1 M n-$BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on a 20%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Lane 1 is intact DNA (all proteo). Lane 2 is a Maxam-Gilbert chemical sequencing G lane. Lanes 3, 4, 5, 6, and 7 contain all proteo, C-1', C-2', C-4', and C-5' denatured thymine at the cleavage site for the nickel reaction, respectively (see Figure 40 for sequence).

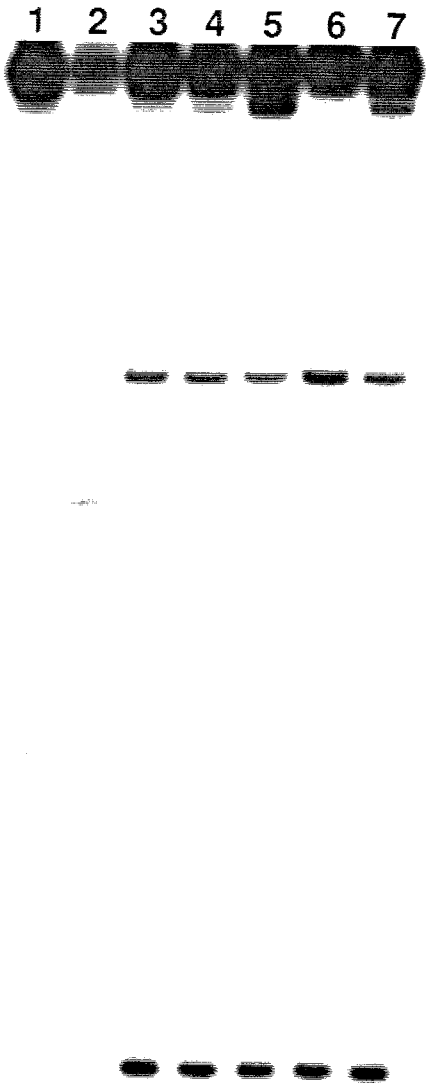
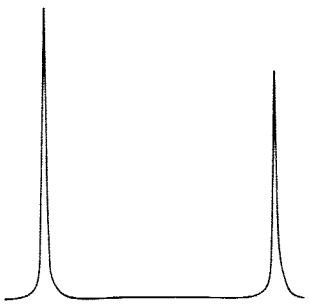
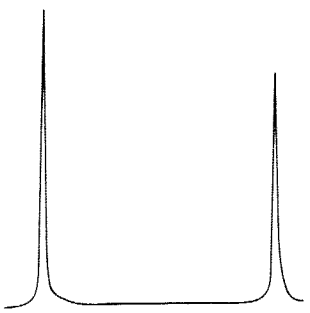

1 2 3 4 5 6 7

Figure 43 Densitometric traces of the gel autoradiogram in Figure 42. The densitometric traces left to right represent the cleavage data from the bottom to the top of the gel in Figure 42 and have been smoothed. Panels A-E correspond to lanes 3-7 in Figure 42. The site of deuterium incorporation is located at the lower site of cleavage in the gel (the left peak in the traces) Panel A contains the all proteo oligonucleotide, panel B contains the C-1', panel C the C-2', panel D the C-4', and panel E the C-5' deuterated oligonucleotides (see Figure 41 for the sequence).

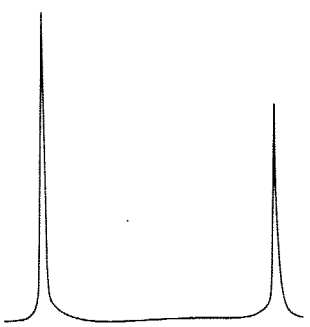
A



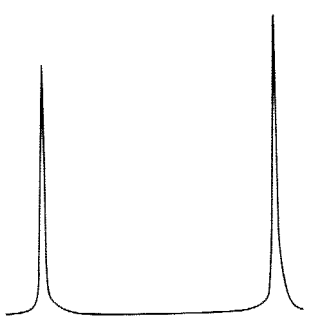
B



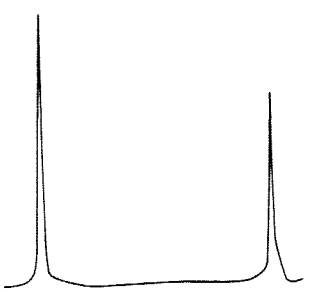
C



D



E



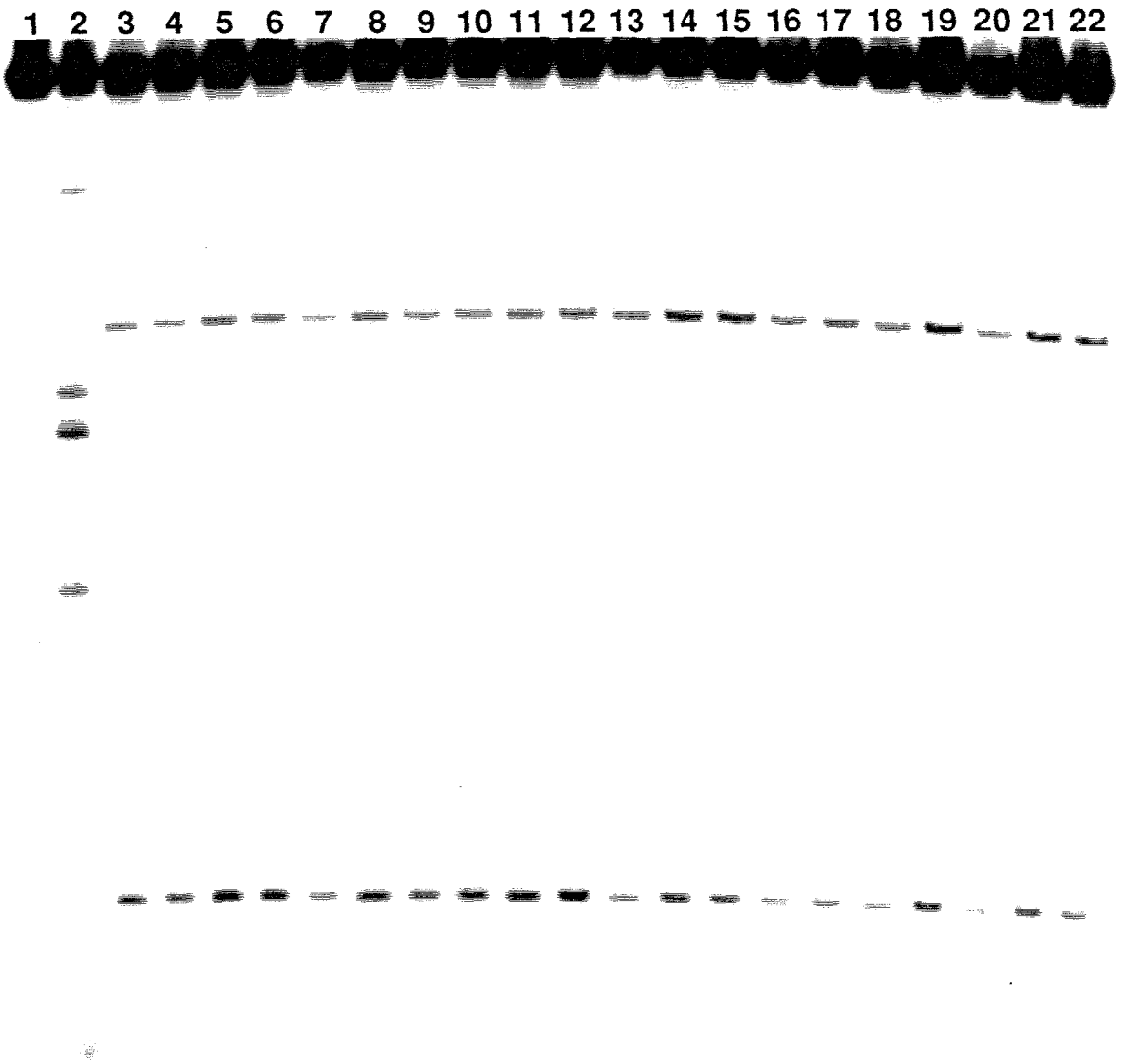
be seen at the C-4' position, where the amount of cleavage at the lower site compared to the upper site has dramatically shifted compared to the control DNA (Figure 43 panels A and D).

For quantitation purposes, each experimental determination of an isotope effect was done on a single gel with 10 lanes each of control DNA, which contained no deuterium atoms at either cleavage site, and deuterated DNA, which contained one of the specifically deuterated nucleotides at the lower cleavage site (Figure 44). The cleavage bands were quantitated on a phosphorimager to give the total radioactivity in the band. The ratio of the lower band to the upper band in the control lanes was averaged as well as the lower band to the upper band in the lanes containing the deuterated position. The control ratio was then divided by the deuterated ratio to give k_H/k_D.[52] Each gel was considered an idependent measurement due to day to day fluctuations in the exact reaction conditions. The data was analyzed for all four positions and shown in Table 2.

Table 2 k_H/k_D	
C-1'	1.02 ± .04 1.04 ± .03
C-2'	0.96 ± .04 1.04 ± .05
C-4'	1.81 ± .12 1.90 ± .07
C-5'	0.81 ± .06 0.86 ± .04

The data in Table 2 indicate that there is no isotope effect at either the C-1' or C-2' positions; however, at the C-4' position there is a large effect and at C-5' there is a small inverse effect.[53] The isotope effect at C-4' of approx 1.8 is likely a primary isotope effect and is consistent with values (2-4) obtained with Fe•bleomycin.[54] This data then implies that the C-4' hydrogen atom is important in the rate limiting steps leading to strand scission

Figure 44 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190) cleavage of 5' ^{32}P labeled oligonucleotides containing all proteo or C-4' deuterated thymine at the sites of attack for the nickel-protein (see Figure 40 for sequence). Reaction conditions were 5 μM Ni(II)•GGH(Hin 139-190), 1 μM monoperoxyphthalic acid, 20 mM NaCl, 20 mM phosphate, pH 7.5, 100 μM (in base pair) calf thymus DNA, and ≈15,000 cpm of end-labeled oligonucleotide. Cleavage reactions were run for 15 min at 25 °C and worked up by treatment with 0.1 M n-$BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on a 20%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. Lane 1 is intact DNA (all proteo). Lane 2 is a Maxam-Gilbert chemical sequencing G lane. Lanes 3-12 contain all proteo oligonucleotide. Lanes 13-22 contain the C-4' deuterated oligonucleotide.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

produced by Ni(II)•GGH(Hin 139-190) in the presence of monoperoxyphthalic acid, consistent with the postulated mechanism.

The isotope effect observed at C-5' also implies that some cleavage is occurring at the C-5' position as was expected from the end-product analysis. This small inverse effect is interesting and may give us some insight into the hydrogen atom abstraction reaction. An isotope effect occurs because the difference in zero-point energy for the C-H and C-D bond in the ground state is not manifested in the transition state energy assuming a linear symmetric transition state. This difference in the zero-point energy of the ground state then becomes a difference in the overall activation energy for the reaction causing the C-D bond breaking reaction to have an overall higher energy. Small or inverse isotope effects are known [55-57] and can arise for a number a reasons. First, the transition state for the hydrogen atom transfer may be bent. In a bent transition state, the bending modes, and thus the energy of the transition states will be different in the presence of a deuterium or a hydrogen atom. This difference in energy in the transition state will compensate for the difference in energy in the ground state rendering the overall transition state energies similar for the two reactions, and thereby decreasing the observed isotope effect. Second, the transition state for the reaction may not be the transition state for the hydrogen atom transfer, that is, the transition state may have more product or reactant bond character. This difference in bond character will again affect the energy of the transition state containing a deuterium or a hydrogen atom and lead to a decrease in the the magnitude of the isotope effect. Either one of these possibilities may be valid for the case of C-5' hydrogen abstraction. If the nickel-oxo species is held rigidly in the minor groove in close proximity to the C-4' hydrogen atom, this could alter the angle of attack on the 5' position as well as affect the nature of the bonding in the transition state for the C-5' abstraction. The explanation for the inverse effect at the C-5' position arising from a primary isotope effect is not the only one, however, and others include α-secondary and β-secondary effects. In order to distinguish these possibilities other experiments will be needed.

Direct Evidence for C4' Hydroxylation. The deuterium isotope effects observed at the 4' and 5' positions of the deoxyribose ring indicate that hydrogen atom abstraction is likely occurring at these positions; however, the lack of an isotope effect at the 1' and 2' positions does not rule out reactivity at these centers. If our mechanistic theory of C-4' hydroxylation is correct, it will be necessary to have a direct quantitative measure of this cleavage pathway. A scheme by which the presence of a 4'-hydroxylated position can be detected has been worked out by Sugiyama *et.al.*[58] The chemistry used in this scheme is identical to that developed to identify the DNA lesions produced by bleomycin,[31] but has been modified to examine large pieces of DNA. The experimental scheme for this assay is shown in Figure 45. Treatment of DNA containing a 4'-hydroxylated position with hydrazine produces strand scission, creating a 3'-(3-pyridazinylmethyl) phosphate end. Enzymatic digestion of the DNA fragments with snake venom phosphodiesterase and alkaline phosphatase produces the 4 different nucleosides and the 4 different 3'-(3-pyridazylmethyl)-nucleosides. The eight possible products (4 nucleosides and 4 pyridazine nucleotides) can then be analyzed by reverse phase HPLC yielding quantitative information on the presence of 4'-hydroxylated sites.

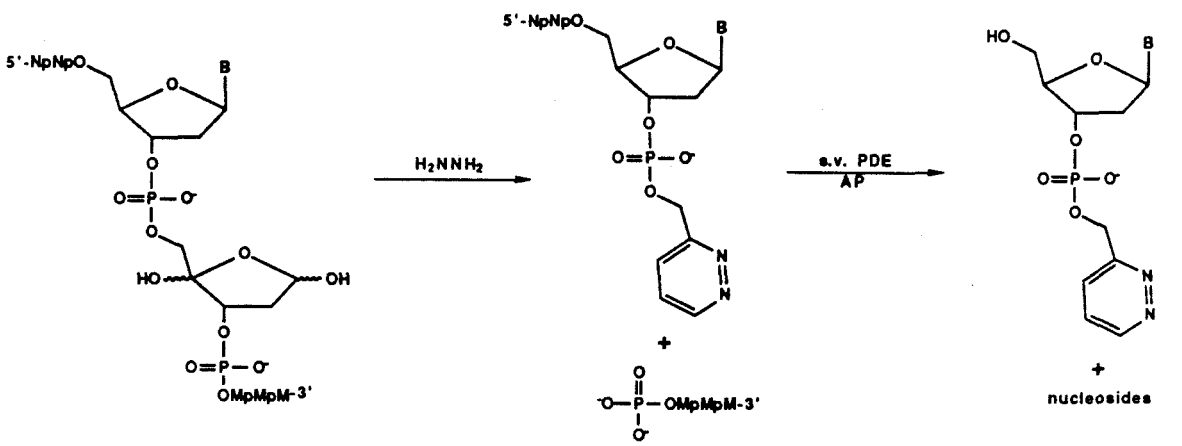


Figure 45 Scheme for the identification and quantitation of C-4' hydroxylated positions in DNA. Treatment with hydrazine followed by digestion with snake venom phosphodiesterase (s.v.PDE) and alkaline phosphatase (AP) lead to the 4 different 3'-(3-pyridazylmethyl)-nucleosides and the 4 different nucleosides (B = A, T, G, and C).

As a control experiment the reaction of Fe•Bleomycin was examined, which is known to produce some 4'-hydroxylation even under aerobic conditions.[45,46] Fe•bleomycin was allowed to react with sonicated calf thymus DNA (random DNA) in the presence of oxygen. The DNA was then treated with a 0.1 M solution of hydrazine for 5 min at 90 °C, followed by digestion with s.v. phospodiesterase and alkaline phosphotase. An aliquot of the digestion was then immediately injected onto a reverse-phase C_{18} HPLC column. The HPLC profile is shown in Figure 46A, and the presence of the G-pyridazine can be observed. The identity of the G-pyridazine was verified by coinjection with an authentic sample of the four different pyridazines (Figure 46B).[59]

Due to limitations in the amount of GGH(Hin 139-190) available and difficulties in finding suitable conditions for a large scale reaction of Ni(II)•GGH(Hin 139-190) with an oligonucleotide containing a hixL half-site, it was decided to attempt the reaction using a radiolabeled substrate. By placing a thymine which contains a tritium in the ribose ring at the site of cleavage on the DNA, it should be possible to produce a tritiated pyridazine product upon workup with hydrazine that could be observed by scintillation counting. In this methodology, it is only necessary to use a very small amount of tritiated DNA, thus making the reaction conditions analogous to the usual ^{32}P-end-labeled DNA experiments used for sequencing gels. For these experiments commercially available thymidine triphosphate, tritiated at the 1', 2', and methyl positions, was incorporated into the DNA duplex shown in Figure 47 using the sequenase enzyme. The DNA cleavage reaction of Ni(II)•GGH(Hin 139-190) using monoperoxyphthalic acid was performed on the tritiated substrate. The DNA was then treated with hydrazine followed by digestion with s.v. phosphodiesterase and alkaline phosphatase. The digestion products were injected onto the HPLC along with authentic samples of the 4 nucleosides and the 4 pyridazine-nucleosides (Figure 48). The peak corresponding to the thymine-pyridazine was collected along with leading and trailing fractions and assayed by scintillation counting.

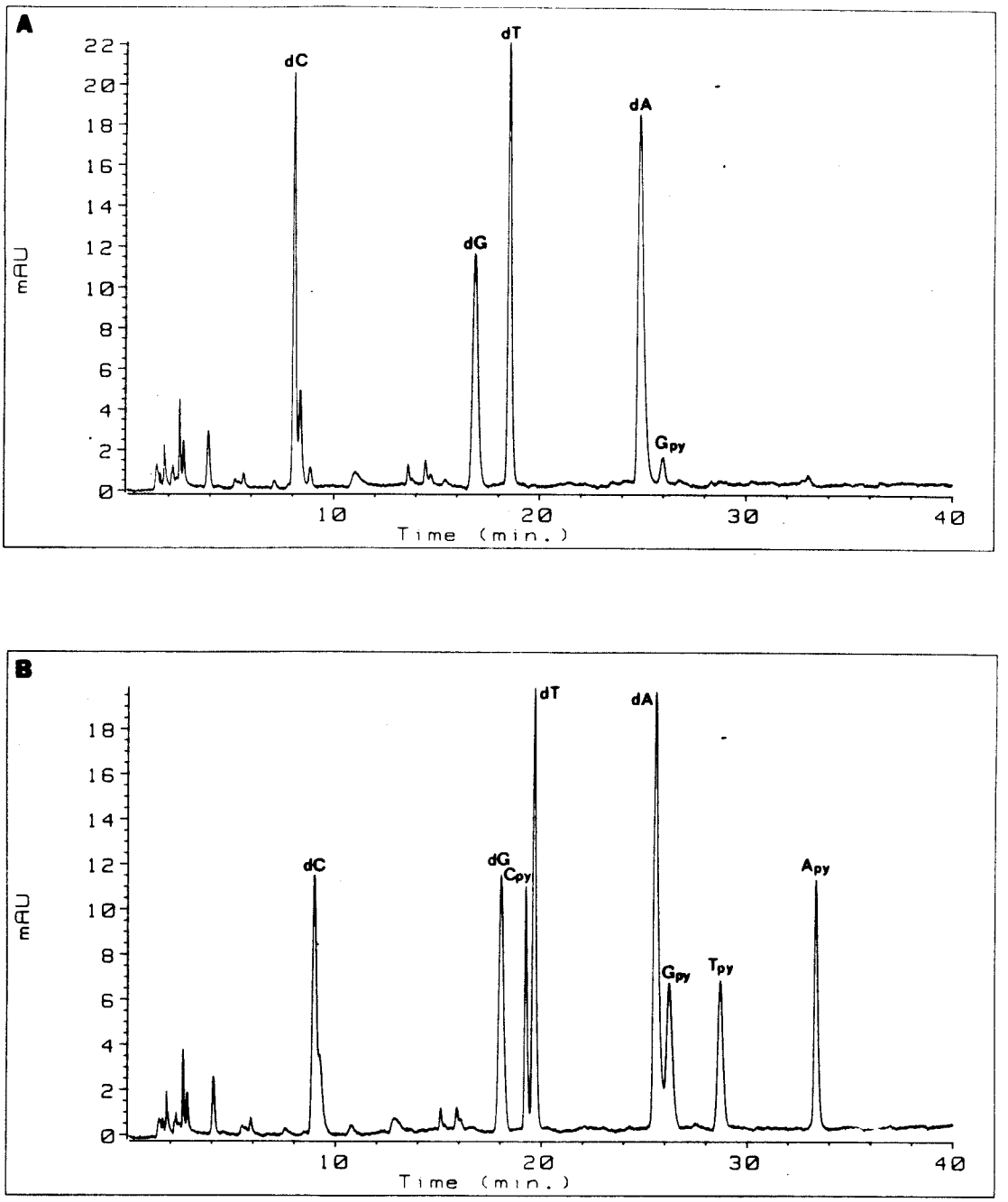


Figure 46 HPLC analysis of products formed by the reaction of Fe•Bleomycin in the presence of calf thymus DNA followed by hydrazine treatment and subsequent enzymatic digestion with snake venom phosphodiesterase and alkaline phosphatase. Panel A shows the mixture after enzymatic digestion and panel B shows the same digestion which had been coinjected with authentic samples of the 4 different 3'-pyridazine nucleosides.

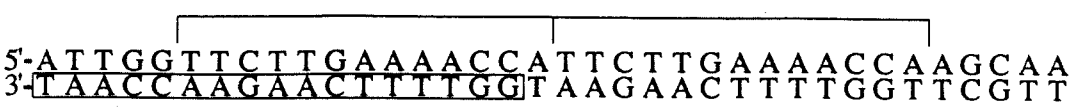


Figure 47 Oligonucleotide sequence used for tritium experiment. The boxed sequence was filled in with dATP, dGTP, dCTP, and (1',2', methyl-^{3}H)-TTP using the sequenase enzyme. The sequence contains the 13 bp left side of hixL (5'-TTCTTGAAAAACCA-3') repeated twice in a head to tail fashion.

Scintillation counting of both control reactions, containing no metal-protein complex, and cleavage reactions indicated the presence of tritium in and around the peak corresponding to the thymine-pyridazine. The amount of tritium was not significantly greater in the cleavage reactions than in the control reactions indicating that some background reaction took place in these experiments. One problem with the design of the experiment is the incorporation of 5 different tritiated thymine residues in the oligonucleotide. Since the enzymatic digestion cleaves the DNA into mononucleotides, any side reaction which causes a modification to the 3' side of a thymine could show up on the

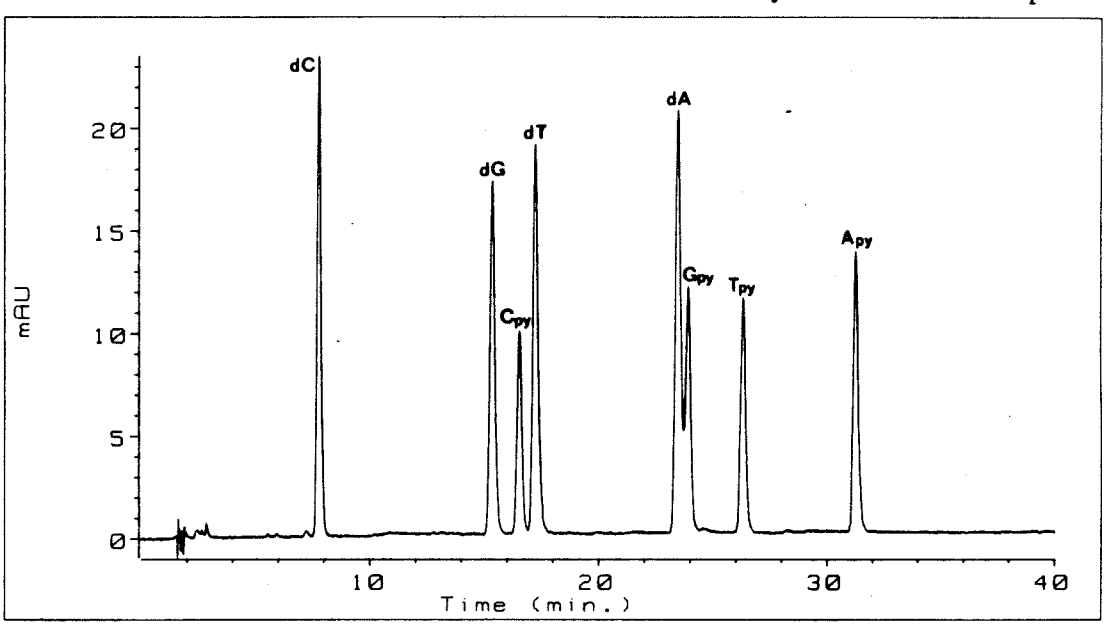


Figure 48 HPLC trace of products formed by the reaction of Ni(II)GGH(Hin 139-199) in the presence of the tritiated oligonucleotide (Figure 47) followed by hydrazine treatment and subsequent enzymatic digestion with snake venom phosphodiesterase and alkaline phosphatase and coinjected with 1 nmol of each of the 4 nucleosides and 3'-pyridazine nucleosides.

HPLC in close proximity to the desired product. This could cause severe background problems as are observed in these experiments. A simple way to redesign the experiment to alleviate this background problem would be to synthesize a thymine specifically labeled with a tritium in the ribose ring only and incorporate it into an oligonucleotide at the single cleavage site by synthetic means. Thus only the single residue would contain a tritium and this should decrease any possible background problems.

Conclusions

Attachment of the metal-binding tripeptide, Gly-Gly-His, to the amino terminus of the DNA binding domain of Hin recombinase (residues 139-190) creates a new 55 residue protein containing only naturally occurring α-amino acids, with two structural domains each with distinct functions, sequence specific recognition and cleavage of double helical DNA. This protein, in the presence of Ni(II) and an oxygen atom donor (eg. iodosylbenzene), produces rapid, efficient modification of the DNA which leads to strand scission. DNA modification likely occurs via abstraction of the C-4' hydrogen atom on the deoxyribose backbone by a high-valent nickel-oxo species. Although mechanistic details remain to be clarified, this result has implications for design of protein catalysts. The tripeptide GGH is a metal-specific structural domain consisting of naturally occurring amino acids that could be incorporated at the amino terminus of a variety of recombinant proteins (such as other DNA binding proteins, receptors, or antibodies) with the function of substrate-directed oxidation in the presence of Ni(II) and an oxygen atom source.

Experimental

General. UV-Vis sprctra were recorded on a Beckman model 25 spectrophotometer, a Cary model 219 spectrophotometer, or a Perkin-Elmer lambda 4C spectrophotometer. Laser densitometry of gel autoradiograms was performed on an LKB Ultrascan XL densitometer. Gels were also exposed to storage phosphor screens and quantitated using a Molecular Dynamics model 400S phosphrimager.

Peptide synthesis and purification was carried out using both the manual and automated methods described in part I. tBoc-d-His(DNP) was purchased from Fluka, tBoc-d-Ala and CBZ-Gly from Bachem Bioscience, and tBoc-Aib from Peninsula Labs.

Doubly distilled or Milli-Q water was used for all aqueous reactions and dillutions. Calf thymus DNA was purchased from Sigma and was sonicated, deproteinized, and dialyzed. Enzymes were purchased from Boehringer Mannheim or New England Biolabs.

Peptide Synthesis. The original GGH(Hin139-190) was synthesized by manual protocols on phenylacetamidomethyl (PAM) resin solid support substituted with tBoc-Asn with Boc-protected amino acids with an average yield per step of 99.4% as obtained by quantitative ninhydrin analysis.[10,34] Subsequent syntheses were performed by automated methodology to obtain the protected resin-bound Hin(139-190). The resin was removed from the machine and further amino acids were coupled to ≈100 mg samples of the resin. GG(d)H- and AibAibH-(Hin 139-190) came from one synthesis and MGGH-, AcGGH-, AAH-, and (d)A(d)AH-(Hin 139-190) came from another. Although in these syntheses a resin sample was removed by the machine for ninhydrin analysis after each double coupling step, it was found that a white precipitate formed in a number of the samples which did not allow quantitation of the sample. Visual comparison of the sample with those of known concentration indicated that the couplings were proceeding in excellent yield in all cases. From each synthesis a sample of the resin was used to produce the GGH(Hin 139-190) protein for comparison with the original material from the manual synthesis. In all cases these proteins behaved exactly like the one from the original synthesis.

Oligonucleotide Synthesis and Purification. All oligonucleotides were synthesized on either a Beckman System 1 Plus DNA synthesizer or an Applied Biosystems 380B DNA synthesizer employing solid-phase phosphite-triester chemistry.[60-62] Oligonucleotides were cleaved from the solid support and deprotected using concentrated ammonium hydroxide at 55 °C for 15-20 hours. Following removal of the ammonium hydroxide the

oligonucleotides were purified by preparative polyacrylamide gel electrophoresis (40 cm long X 4 mm thick). The bands were visualized by UV illumination and the desired band excized from the gel. The gel slice was crushed and placed in a minimun amount of 0.2M NaCl and eluted overnight at 37 °C. The solution was filtered and the liquid placed in dialysis tubing (2000 MW cutoff) and dialyzed against water for 3 days (water changes every 12 hrs).

5' End-labeling of Oligonucleotides. About 30 pmol of single stranded oligonucleotide was 5' end-labeled with γ-^{32}P ATP and T4 polynucleotide kinase. After ethanol precipitation and equal molar amount of the appropriate opposite strand was added in a total final volume of 20 μL. The mixture was heated to 90 °C for 5 minutes and allowed to cool slowly to room temperature. The duplex DNA was then isolated on a 15% nondenaturing polyacrylamide prep gel (15 cm long X 2 mm thick). The bands were visualized by autoradiography and the desired band was excized from the gel. The gel slice was crushed and placed in a minimum volume of 0.2M NaCl and eluted overnight at 37 °C. The solution was filtered and extracted with dry n-butanol until the volume was ≈200 μL. The DNA was ethanol precipitated and resuspended in water at ≈30,000 cpm per μL.

Plasmid Construction. Two plasmids were constructed to investigate the end products of DNA cleavage produced by the metal-GGH(Hin 139-190) proteins. Each plasmid was prepared by hybridizing a pair of synthetic oligonucleotides to give a short duplex possessing cohesive EcoRI and HindIII termini. Table 3 lists the oligonucleotides used to construct the plasmids. Each duplex was ligated into the EcoRI/HindIII fragment of plasmid pBR322. The resulting ligation mixture was used to transform E. coli strain HB101 according to the standard $CaCl_2$ procedure.[63] Recombinant clones were selected for ampicillin resistance and tetracycline sinsitivity due to the disruption of the tetracycline gene by the between EcoRI and HindIII.[64] Plasmid DNA (pDPM12, pDPM34, and pMFB36[1]) were isolated from E. coli and purified by CsCl centrifugation.[65,66]

Table 3. Oligonucleotides Used for Plasmid Preparation	
pDPM12	5′AATTCTATTCGTTCTTGAAAACCAAGGTTTTTGATAAGCAATCA GATAAGCAAGAACTTTTGGTTCCAAAAACTATTCGTTAGTTCGA5′
PDPM34	5′AATTCTATTCGTTCTTCCTTATCTGATGTAAAGGAGAAGCAATCA GATAAGCAAGAAGGAATAGACTACATTTCCTCTTCGTTAGTTCGA5

Labeling of Restriction Fragments from plasmids. The procedure described is for the XbaI/EcoRI fragment from pMFB36,[1] but is the same for other plasmids but with the appropriate restriction enzymes. Plasmis pMFB36[1] was linearized by cleavage with restriction endonuclease XbaI. Labeling at the 3' end was accomplished with [α-^{32}P]dATP and the Klenow fragment of DNA polymerase I. The 5' end was labeled with ^{32}P by dephosphorylation with calf alkaline phosphatase (CAP) followed by treatment with [γ-^{32}P]ATP and T$ polynucleotide kinase. Cleavage with restriction endonuclease EcoRI yielded 3'- and 5'- end-labeled fragments 557 bp, which were isolated on a 5% nondenaturing polyacrylamide prep gel (20 cm long X 2 mm thick). The bands were visualized by autoradiography and the desired band excized from the gel. The gel slice was crushed and placed in a minimum volume of 0.2M NaCl and eluted overnight at 37 °C. The solution was filtered and n-butanol extracted until the volume was ≈200 μL. The DNA was ethanol precipitated twice and resuspended in water to a final concentration of ≈30,000 cpm per μL. A specific sequencing reactions were then preformed on the purified restriction fragments.[67]

End-labeling of Linerarized Plasmid pDPM12. pDPM12 was linearized with AvaI, and labeled on the 3' terminus of the counterclockwise strand with [α-^{32}P]TTP using the Klenow fragment from DNA polymerase I. The radiolabeled plasmid was then extracted with 50:50 phenol:chloroform to remove the enzyme. The unincorporated nucleotides were removed using a G-50 Sephadex spun-column[62] and the linearized plasmid diluted to 25,000 cpm per μL. Small portion (≈125,000 cpm) of the labeled plasmid were combined with 1 μg of CT DNA and digested with restriction enzymes SalI, EcoRI, XmnI, PstI, and

NdeI. The reactions were ethanol precipitated, counted and diluted such that 15 μL of the final solution contained 500 cpm of each band produced plus intact DNA.

MPE•Fe(II) Reactions. MPE•Fe(II) reactions were done in a total volume of 20 μL. Final concentrations were 10 μM MPE•Fe(II), 20 mM phosphate, pH 7.5, 20 mM NaCl, 100 μM (in base pairs) CT DNA, ≈15,000 cpm of ^{32}P-end-labeled DNA, 5 mM dithiothreitol (DTT), and appropriate concentrations of proteins if a footprinting experiment was being done. If protein was present, it was allowed to equilibrate with the DNA for 10 min. prior to addition of MPE•Fe(II). After a 1 min. incubations with MPE•Fe(II) the reaction was initiated by the addition of DTT and allowed to proceed for 12 min at 25 °C. The reactions were terminated by ethanol precipitation, dried, and resuspended in 5 μL of 100 mM Tris-borate-EDTA and 80% formamide solution. The ^{32}P-end-labeled products were analyzed by denaturing polyacrylamide gel electrophoresis followed by autoradiography.

Cu(II)•GGH(Hin 139-190) Cleavage Reactions. Cleavage reactions were done in a total volume of 20 μL. Final concentrations were 20 mM phosphate, pH 7.5, 20 mM NaCl, 100 μM (in base pairs) CT DNA, ≈15,000 cpm of ^{32}P-end-labeled DNA, 1 mM H_2O_2, 1 mM sodium ascorbate and appropriate concentrations of Cu(II)•GGH(Hin 139-190). The protein was allowed to equilibrate with the DNA for 10 min. prior to addition of H_2O_2 and sodium ascorbate. The reaction was initiated by the addition of H_2O_2 and sodium ascorbate and allowed to proceed for 90 min at 25 °C. The reactions were terminated by ethanol precipitation, dried, and resuspended in 5 μL of 100 mM Tris-borate-EDTA and 80% formamide solution. The ^{32}P-end-labeled products were analyzed by denaturing polyacrylamide gel electrophoresis followed by autoradiography.

Ni(II)•GGH(Hin 139-190) Cleavage Reactions. Cleavage reactions were done in a total volume of 20 μL. Final concentrations were 20 mM phosphate, pH 7.5, 20 mM NaCl, 100 μM (in base pairs) CT DNA, ≈15,000 cpm of ^{32}P-end-labeled DNA, 5 μM monoperoxyphthalic acid (or appropriate concentrations of other oxidants) and 5 μM

Ni(II)•GGH(Hin 139-190). The protein was allowed to equilibrate with the DNA for 10 min prior to addition of oxidant. The reaction was initiated by the addition of oxidant and allowed to proceed for 15 min at 25 °C. The reactions were terminated by ethanol precipitation, dried, resuspended in 50 μL of 0.1 N n-Butylamine, heated to 90 °C for 30 min, frozen and lyophilized, and resuspended in 5 μL of 100 mM Tris-borate-EDTA and 80% formamide solution. The ^{32}P-end-labeled products were analyzed by denaturing polyacrylamide gel electrophoresis followed by autoradiography.

Analysis of DNA Termini by Gel Electrophoresis. Cleavage reacions were run as described above for MPE•Fe(II), Cu(II)•GGH(Hin 139-190), and Ni(II)•GGH(Hin 139-190). The DNA cleavage fragments were then analyzed directly by electrophoresing the 20% gel until the desired product was as close to the bottom as possible.

The presence of 5' phosphate groups was tested using calf alkaline phosphotase to remove the 5' phosphate groups. After the DNA reaction was completed and the DNA ethanol precipitated (if base workup was necessary the solution was neutralized with NaOAc and ethanol precipitated), the DNA was dissolved in 45 μL of 5 mM Tris, pH 7.5, 10 μM EDTA and heated to 90 °C for 5 min. Calf alkaline phosphatase was added (5 μL of 1 unit per μL) and incubated at 37 °C for 30 min. The reaction was stopped by ethanol precipitation and analyzed on a 20% polyacrylamide gel.

The presence of 3' phosphate groups was tested using T4 polynucliotide kinase to remove the 3' phosphate groups.[17] After the DNA reaction was completed and the DNA ethanol precipitated (if base workup was necessary the solution was neutralized with NaOAc and ethanol precipitated), the DNA was dissolved in 25 μL of water and heat denatured for 5 min at 90 °C and cooled on ice. 25 μL of a buffer containing 20 mM Tris•HCl, pH 6.6, 20 mM magnesium chloride, and 10 mM β-mercaptoehtanol (made fresh) was added followed by 1 μL T4 polynucleotide kinase (10 units per μL). The reaction was incubated for 1 hour at 37 °C and ethanol precipitated for gel electrophoresis.

Double Strand Cleavage Reactions. Cleavage reactions were done in a total volume of 45 μL. Final concentrations were 20 mM phosphate, pH 7.5, 20 mM NaCl, 100 μM (in base pairs) CT DNA, ≈25,000 cpm of ^{32}P-end-labeled DNA, 2.5 μM monoperoxyphthalic acid (or 5 mM DTT) and 2.5 μM Ni(II)•GGH(Hin 139-190) (or 2.5 μM Fe•EDTA-Hin (139-190)). The protein was allowed to equilibrate with the DNA for 10 min prior to addition of oxidant (or reductant). The reaction was initiated by the addition of oxidant (or reductant) and allowed to proceed for 15 min (or 60 min) at 25 °C. The EDTA-protein reactions were terminated by addition of 5μL of a 10x ficol solution and stored frozen. To the Ni(II)-protein reactions was added 2 μL 2 M NaCl and 1 μL of a 5% piperidine in water solution and heated to 55 °C for 15-18 hours. 5 μL of a 10x ficol solution was added and the cleavage products analyzed on a 1% agarose gel (25 cm long X 4 mm thick) and visualized by autoradiography.

Deuterium Isotope Effects. DNA cleavage reactions of oligonucleotides containing all proteo, C-1', C-2', C-4', or C-5' thymine at the site of strongest cleavage for the Ni(II)•GGH(Hin 139-190) were done as described above. The exact oligonucleotide sequence is shown in Figure 40 and contains two copies of the hixL half site oriented in a head to tail fashion. Cleavage reactions were performed using 5 μM Ni(II)•GGH(Hin 139-190) and 1 μM monoperoxyphthalic acid. To insure as much day to day reproducibility in the measurements a stock solution was made up which contained the NaCl, phosphate buffer, and calf thymus DNA. This solution was then combined with the labeled oligonucleotide, protein, and oxidant to make a total final volume of 20 μL. Cleavage products were analyzed on 20% polyacrylamide denaturing gels and quantitated by use of the phosphorimager. Each gel contained 10 lanes of all proteo duplex as controls and 10 lanes of the specifically deuterated duplex. The deuterated thymine was always positioned at the site of cleavage closest to the label (lowest on the gel). The ratio of the lower to the upper band was calculated for each lane and the 10 control and 10 deuterated lanes were

averaged separately and standard deviations calculated. The ratio of the ratios (control to deutero) then gives k_H/k_D.

Pyridazine Formation from C-4' Hydroxylation. Fe•Bleomycin reacion. The DNA cleavage reaction was performed for 2 hours at 25 °C with the following final concentrations; 1 mM (in base pairs) calf thymus DNA, 25 mM phosphate, pH 7.2, 0.5 mM Fe•Bleomycin, in a total volume of 100 μL. To 20 μL of this solution was added 1 μL of a 2 M hydrazine solution followed by heating at 90 °C for 5 minutes. This solution was then diluted to 100 μL with final concentrations of 50 mM Tris, pH 7.5, 0.00015 units per μL snake venom phosphodiesterase and 0.05 units per μL calf alkaline phosphotase and incubated at 37 °C for 3 hrs. 20 μL of the digest was then injected onto the HPLC for analysis. HPLC was performed on a Hewlett-Packard HP1090M instrument equipped with a diode array detector and autosampler using a Regis Rexchrom C_{18} reverse phase column (25 cm X 4.6 mm). The separation was obtained using a linear gradient of 0 to 8% acetonitrile in 0.05 M ammonium formate at a flow rate of 1.5 mL per minute. The identity of the peaks was verified by coinjecting 20 μL of the digest with 5 μL of a solution which contained 200 μM each of the 4 different pyridazine derivatives.[59]

Ni(II)•GGH(Hin 139-190) reaction. The oligonucleotide containing a hixL half site which has [1', 2', methyl- ^{3}H]Thymine incorporated was prepared by using the sequenase enzyme (U.S. Biochemical Corp.), dATP, dGTP, dCTP, and [1', 2', methyl- ^{3}H]TTP (Amersham) to fill in the overhang of the duplex oligonucleotide shown in Figure 53. The full length duplex was purified on a 15% nondenaturing polyacrylamide prep gel (20 cm long X 2 mm thick). The bands were visualized by staining with ethidium bromide and the desired band was extracted from the gel. The gel slice was crushed and placed in a minimum volume of 0.2 M NaCl and eluted overnight at 37 °C. The solution was filtered and extracted extensively with n-butanol to remove the ethidium bromide. The DNA was ethanol precipitated and resuspended to ≈250,000 dpm per μL.

The Ni(II)•GGH(Hin 139-190) cleavage reaction was run exactly as for the sequencing reactions but with ≈500,000 dpm of ^{3}H-oligonucleotide and no n-butylamine treatment. The entire 20 μL reaction was then treated with hydrazine and digested with s.v. PDE and CAP as described for the Fe•Bleomycin reaction above. 25 μL of the digest was placed in an injection vial which had 1 nmol of each of the four nucleosides and each of the four pyridazine nucleosides. This entire volume was injected onto the HPLC and the peak corresponding to the Thymine-pyridazine product was collected and analyzed by scintillation counting.

References

1. Bruist, M.F.; Horvath, S.J.; Hood, L.E.; Steitz, T.A.; Simon, M.I., *Science*, **235**, 777 (1987).
2. Sluka, J.P.; Horvath, S.J.; Bruist, M.F.; Simon, M.I.; Dervan, P.B., *Science*, **238**, 1129 (1987).
3. Sluka, J.P.; Horvath, S.J.; Glasgow, A.C.; Simon, M.I.; Dervan, P.B., *Biochemistry*, **29**, 6551 (1990).
4. Sluka, J.P.; Griffin, J.H.; Mack, D.P.; Dervan, P.B., *J. Am. Chem. Soc.*, in press.
5. Lau, S.-J.; Kurck, T.P.A.; Sarkar, B., *J. Biol. Chem.*, **249**, 5878 (1974).
6. Kruck, T.P.A.; Lau, S.-J.; Sarkar, B., *Can. J. Chem.*, **54**, 1300 (1976).
7. Camerman, N.; Camerman, A.; Sarkar, B., *Can. J. Chem.*, **54**, 1309 (1976).
8. Chiou, S.-H., *J. Biochem.*, **57**, 957 (1983).
9. Kent, S.B.H., *Ann. Rev. Boichem.*, **57**, 957 (1988).
10. Sarin, V.K.; Kent, S.B.H.; Tam, J.P.; Merrifield, R.B., *Anal. Biochem.*, **117**, 147 (1981).
11. Sequence analysis by Edmann degradation was performed by David Teplow and Haul Wong at CIT.
12. Dervan, P.B., *Science*, **232**, 464 (1986).
13. Kross, J.; Henner, W.D., Hecht, S.M.; Haseltine, W.A., *Biochemistry*, **21**, 4310 (1982).
14. Tapper, D.P.; Clayton, D.A., *Nucleic Acids Res.*, **9**, 6787 (1981).
15. Hertzberg, R.P.; Dervan, P.B., *Biochemistry*, **23**, 3934 (1984).
16. Chaconas, G.; Van de Sande, J.H., *Methods Enz.*, **65**, 75 (1980).

17. Cameron, V.; Uhlenbeck, O.C., *Biochemistry*, **16**, 5120 (1977).
18. Margerum, D.W.; Chellappa, K.L.; Bossu, F.P.; Burce, G.L., *J. Am. Chem. Soc.*, **97**, 6894 (1975).
19. Levitzki, A.; Anbar, M.; Berger, A., *Biochemistry*, **6**, 3757 (1967).
20. Kurtz, J.L.; Burce, G.L.; Margerum, D.W., *Inorg. Chem.*, **17**, 2454 (1978).
21. Margerum, D.W.; Dukes, G.R., *Metal Ions Biol. Syst.*, **1**, 158 (1974).
22. Margerum, D.W.; Wong, L.F.; Bossu, F.P.; Chellappa, K.L., Czarnecki, J.J.; Dirksey, S.T.; Neubecker, T.A., *Adv. Chem. Ser.*, **162**, 281 (Bioinorg. Chem. 2nd Symp.) (1977).
23. Margerum, D.W., *Pure and Appl. Chem.*, **55**, 23 (1983).
24. Bryce, C.F.; Roeske, R.W.; Gurd, F.R.N., *J. Biol. Chem.*, **241**, 1072 (1966).
25. Bossu, F.P.; Margerum, D.W., *Inorg. Chem.*, **16**, 1210 (1977).
26. Bannister, C.E.; Raycheba, J.M.T.; Margerum, D.W., *Inorg. Chem.*, **21**, 1106 (1982).
27. Sakuri, T.; Nakahara, A.; *Inorg. Chim. Acta.*, **34**, L243 (1979).
28. Freeman, H.C.; Taylor, M.R., *Acta Crystallogr.*, **18**, 939 (1965).
29. Freeman, H.C.; Guss; J.M., Sinclair, R.L., *Chem. Comm.*, **485** (1968).
30. Bossu, F.P.; Paniago, E.B.; Margerum, D.W.; Kirksey, S.T.; Kurtz, J.L., *Inorg. Chem*, **17**, 1034 (1978).
31. Sugiyama, H.; Xu, C.; Murugesan, N.; Hecht, S.M., *Biochemistry*, **27**, 58 (1988).
32. Schultz, P.G.; Dervan, P.B., *Proc. Nat. Acad. Sci., USA*, **80**, 6834 (1983).
33. Moser, H.; Dervan, P.B., *Science*, **238**, 645 (1987).
34. Mack, D.P.; Iverson, B.I.; Dervan, P.B., *J. Am. Chem. Soc.*, **110**, 7572 (1988).
35. Bachmair, A.; Finley, D.; Varshavsky, A., *Science*, **234**, 179 (1986).
36. Huse, W.D.; Sastry, L.; Iverson, S.A.; Kang, A.S.; Alting-Mees, M.; Burton, D.R.; Benkovic, S.J.; Lerner, R.A., *Science*, **246**, 1275 (1989).
37. Kirksey, S.T.; Neubecker, T.A.; Margerum, D.W., *J. Am. Chem. Soc.*, **101**, 1631 (1979).
38. Hamburg, A.W.; Nemeth, M.T.; Margerum, D.W., *Inorg. Chem.*, **22**, 3535 (1983).
39. Diaddario, L.L.; Robinson, W.R.; Margerum, D.W., *Inorg. Chem.*, **22**, 1021 (1983).
40. Kennedy, W.R.; Margerum, D.W., *Inorg. Chem*, **24**, 2490 (1985).
41. Dizdaroglu, M.; Schulte-Frohlinde, D.; von Sonntag, C.Z.; Natruforsch, C., *Biosci.*, **32C**, 1021 (1977).
42. Sigman, D.S., *Acc. Chem. Res.*, **19**, 180 (1986).
43. Kappen, L.S.; Goldberg, I.H., *Nucleic Acids Res.*, **13**, 1637 (1985).

44. Chin, D.-H.; Kappen, L.S.; Goldberg, I.H., *Proc. Nat. Acad. Sci., USA*, **84**, 7070 (1987).
45. Hecht, S.M., *Acc. Chem. Res.*, **19**, 83 (1986).
46. Stubbe, J.; Kozarich, J.W., *Chem. Rev.*, **87**, 1107 (1987).
47. Kinneary, J.F.; Wagler, T.R.; Burrows, C.J., *Tetrahedron Lett*, **29**, 877 (1988).
48. Wagler, T.R.; Burrows, C.J., *Tetrahedron Lett.*, **29**, 5091 (1988).
49. Yoon, H.; Burrows, C.J., *J. Am. Chem. Soc.*, **110**, 4087 (1988).
50. Kinneary, J.F.; Albert, J.S.; Burows, C.J., *J. Am. Chem. Soc.*, **110**, 6124 (1988).
51. C-1', C-2', C-4', and C-5' specifically deuterated DMT-thymidine-phophoramidites were graciously provided by Dr. M.M. Greenberg.
52. The kinetic isotope effect (KIE) = $\frac{(k_{1H}/k_{2H})}{(k_{1D}/k_{2H})}$
53. The measured KIE is a lower limit due to the nature of the experimental technique since the cleavage band quantitated may have contributions from all of the hydrogen atom positions. The measured KIE is given by the following equation:
KIE(C-1') = $\frac{k_{1'H} + k_{2'H} + k_{3'H} + k_{4'H} + k_{5'H}}{k_{1'D} + k_{2'H} + k_{3'H} + k_{4'H} + k_{5'H}}$
Thus, reaction at any other center acts to diminish the isotope effect observed.
54. Kozarich, J.W.; Worth, L.; Frank, B.L.; Christner, D.F.; Vanderwall, D.E.; Stubbe, J.; *Science*, **245**, 1396 (1989).
55. Baldwin, J.E.; Andist, A.H., *Chem. Comm.*, 1512 (1971).
56. Pasto, D.J.; Meyer, G.R.; Lepeska, B., *J. Am. Chem. Soc.*, **96**, 1858 (1974).
57. Wigfield, D.C.; Phelps, D.J.; Pottie, R.F.; Sander, R., *J. Am. Chem. Soc.*, **97**, 897 (1975).
58. Sugiyama, H.; Kawabata, H.; Fujiwara, T.; Dannoue, Y.; Saito, I., *J. Am. Chem. Soc.*, **112**, 5252 (1990).
59. The 3'-pyridazine nucleosides were a gracious gift from H. Sugiyama.
60. Matteucci, M.D.; Caruthers, M.H., *J. Am. Chem. Soc.*, **103**, 3185 (1981).
61. Beaucage, S.L.; Caruthers, M.H., *Tetrahedron Lett.*, **22**, 1859 (1984).
62. Atkinson, T.; Smith, M., In *Oligonucleotide Synthesis: A Practical Approach* (Gait, M.J., ed.), pp. 35-81, IRL Press, Oxford and Washington D.C. (1984).
63. Maniatis, T.; Fritsch, E.F.; Sambrook, J., In *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1982).
64. Sutcliffe, J.G., *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 77 (1979).
65. Tanaka, T.; Weisblum, B., *J. Bacteriol.*, **121**, 354 (1975).
66. Mendel, D., Ph.D. thesis, California Institute of Technology, Pasadena, CA (1989)
67. Iverson, B.L.; Dervan, P.B., *Nucleic Acids Res.*, **15**, 7823 (1987).

Appendix A

Conditions Gels for the Cu(II) and Ni(II)•GGH(Hin 139-190) DNA Cleavage reactions.

Figure 1 Autoradiogram of high-resolution denaturing gel of Cu(II)•GGH(Hin 139-190) and Ni(II)•GGH(Hin 139-190) cleavage reaction time courses on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Nickel-mediated reactions (and Cu-monoperoxyphthalic acid reactions) were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 5'-end-labeled DNA was used. Lane 1 is a DNA control lane. Lanes 2 is an A specific sequencing reaction. Lanes 3 thru 5 contain $Ni(OAc)_2$ (5 μM) and GGH(Hin 139-190) (5 μM) followed by monoperoxyphthalic acid, reactions were run for 15, 30, and 60 minutes, respectively. Lanes 6 thru 10 contain $Ni(OAc)_2$ (5 μM) and GGH(Hin 139-190) (5 μM) followed by hydrogen peroxide (5 mM), reactions were run for 30, 60, 90, 120, and 180 minutes, respectively. Lanes 11 thru 15 contain $CuCl_2$ (2.5 μM) and GGH(Hin 139-190) (5 μM) followed by monoperoxyphthalic acid (1 mM), reactions were run for 30, 60, 90, 120, and 180 minutes, respectively. Lanes 16 thru 20 contain $CuCl_2$ (2.5 μ M), GGH(Hin 139-190) (5 μM) followed by sodium ascorbate and hydrogen peroxide (1 mM each), reactions were run for 30, 60, 90, 120, and 180 minutes, respectively.

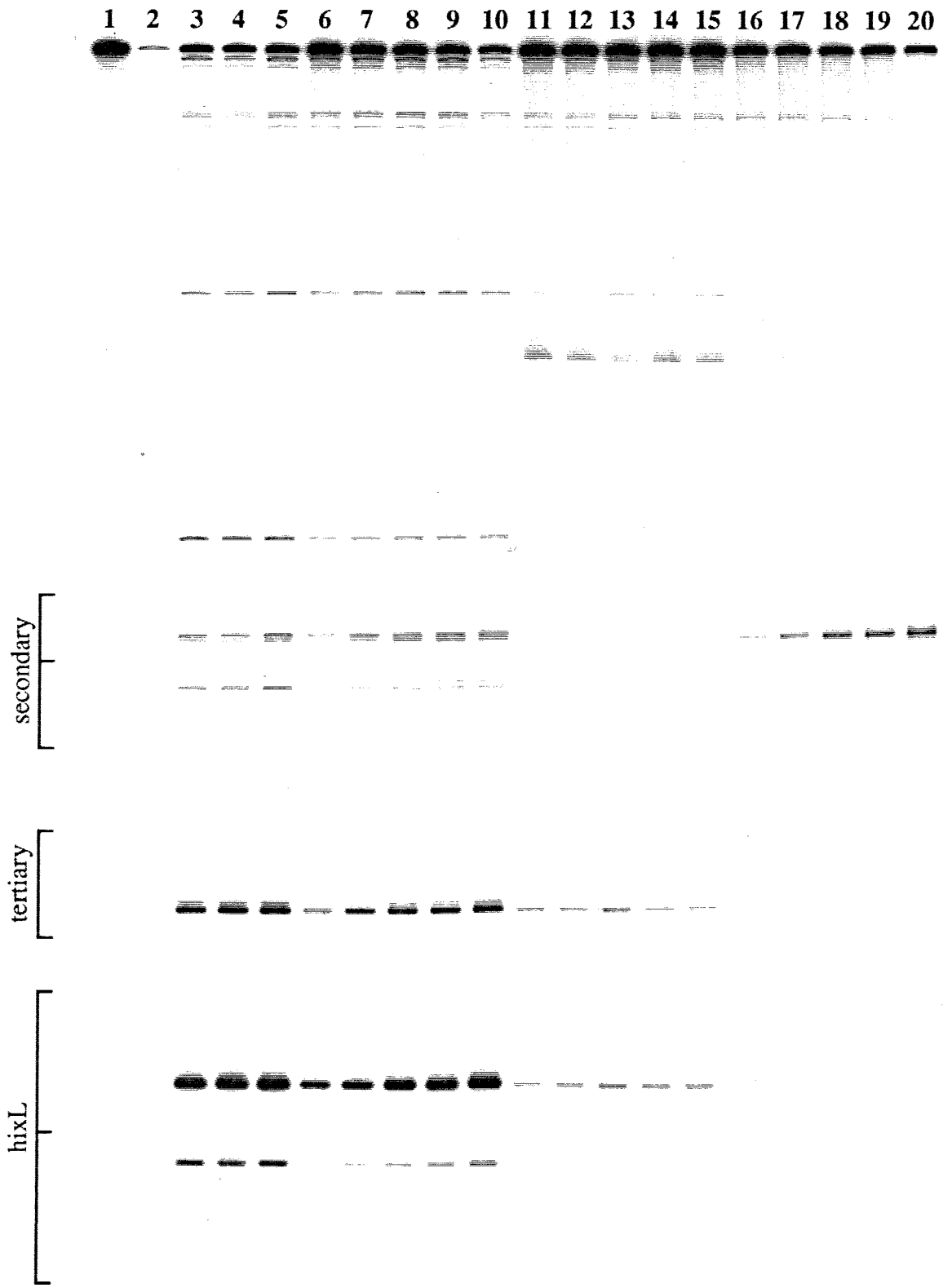
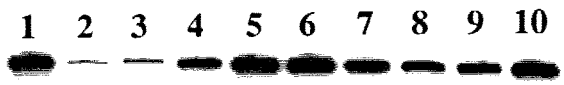

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
secondary
tertiary
hixL

Figure 2 Autoradiogram of high-resolution denaturing gel of Cu(II)•GGH(Hin 139-190) cleavage reaction varying sodium ascorbate and hydrogen peroxide concentrations on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 2.5 μM $CuCl_2$, 5 μM GGH(Hin 139-190) 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions were run for 90 minutes at 25 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 3'-end-labeled DNA was used. Lane 1 is a DNA control lane. Lane 2 is an A specific sequencing reaction. Lanes 3 thru 6 contain 1 mM sodium ascorbate and 5, 1, 0.1, and 0.01 mM hydrogen peroxide, respectively. Lanes 7 thru 10 contain 1 mM hydrogen peroxide and 5, 1, 0.1, and 0.01 mM sodium ascorbate.


1 2 3 4 5 6 7 8 9 10

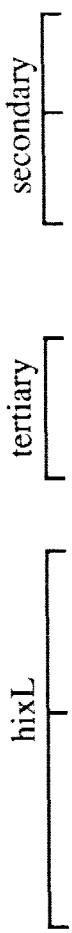

secondary
tertiary
hixL

Figure 3 Autoradiogram of high-resolution denaturing gel of Cu(II)•GGH(Hin 139-190)-sodium ascorbate-hydrogen peroxide DNA cleavage reaction pH study on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 2.5 μM $CuCl_2$, 5 μM GGH(Hin 139-190), 20 mM NaCl, 20 mM phosphate, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions were run for 90 minutes at 25 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 3'-end-labeled DNA was used. Lane 1 is a DNA control lane. Lane 2 is an A specific sequencing reaction. Lanes 3 thru 7 are at pH 6.0, 6.5, 7.0, 7.5, and 8.0, respectively.

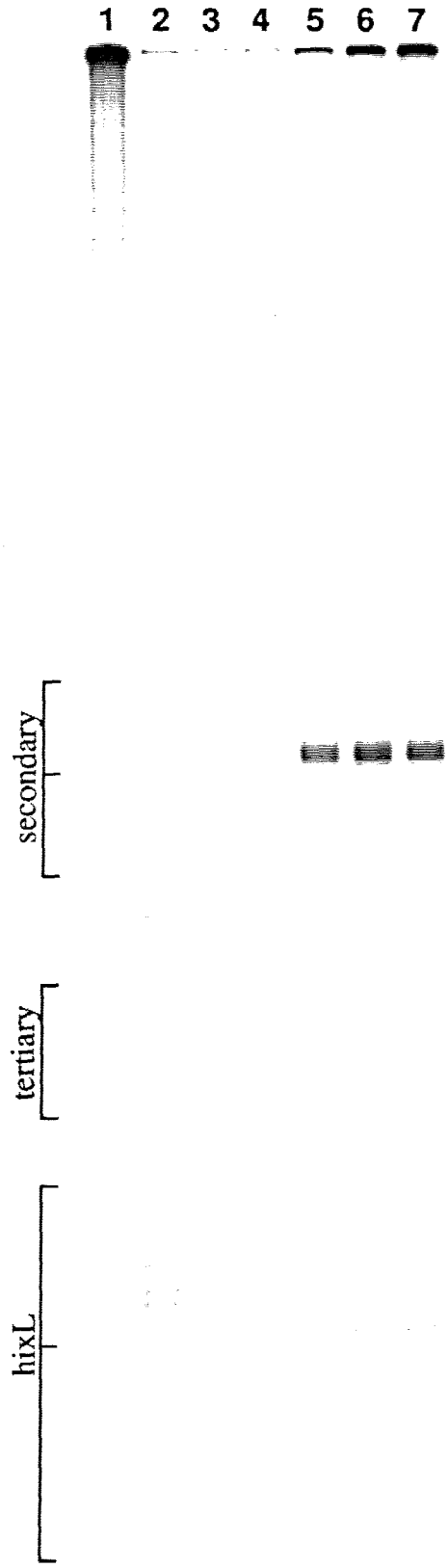

1 2 3 4 5 6 7
secondary
tertiary
hixL

Figure 4 Autoradiogram of high-resolution denaturing gel of Cu(II)•GGH(Hin 139-190) monoperoxyphthalic acid cleavage reaction peracid concentration study on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 2.5 μM $CuCl_2$,5 μM GGH(Hin 139-190), 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reacions were run for 30 min at 25 °C. Copper-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 3'-end-labeled DNA was used. Lane 1 is a DNA control lane. Lane 2 is an A specific sequencing reaction. Lanes 3 thru 8 contain 5, 1, 0.5, 0.05, and 0.005 mM monoperoxyphthalic acid, respectively.

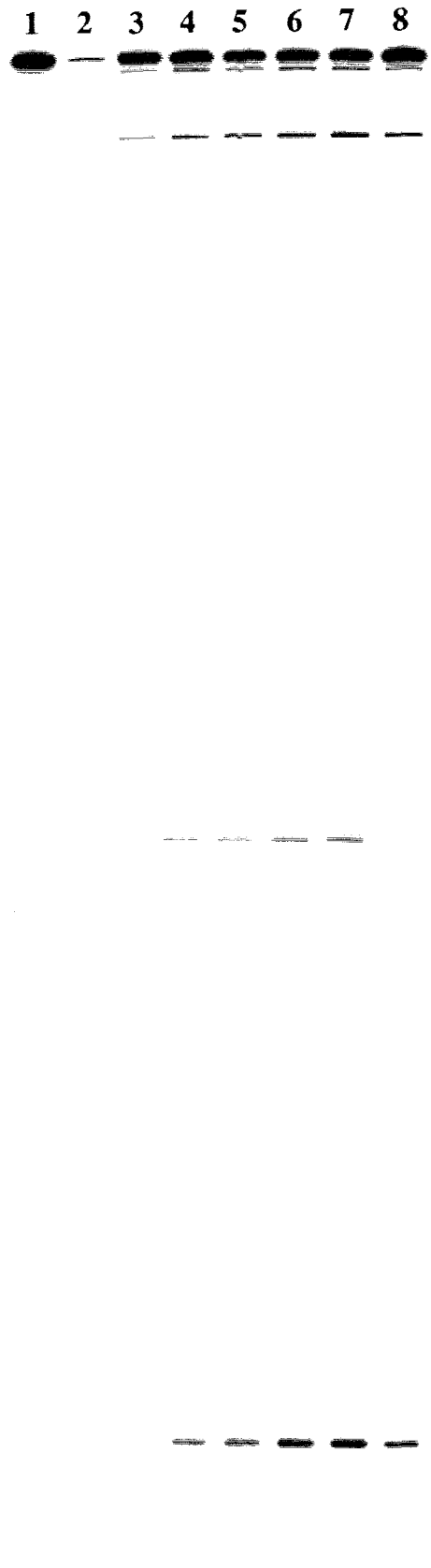

1 2 3 4 5 6 7 8
secondary
tertiary
hixL

Figure 5 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190)-hydrogen peroxide cleavage reaction H_2O_2 concentration study on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 5 μM $Ni(OAc)_2$, 5 μM GGH(Hin 139-190), 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions were run for 90 minutes at 25 °C. Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 3'-end-labeled DNA was used. Lane 1 is a DNA control lane with 100 mM hydrgogen peroxide and no base workup. Lane 2 is a DNA control lane with 100 mM hydrogen peroxide followed by base workup. Lane 3 is an A specific sequencing reaction. Lanes 5 thru 10 contain 100, 10, 5, 1, 0.1, and 0.01 mM hydrogen peroxide, respectively.

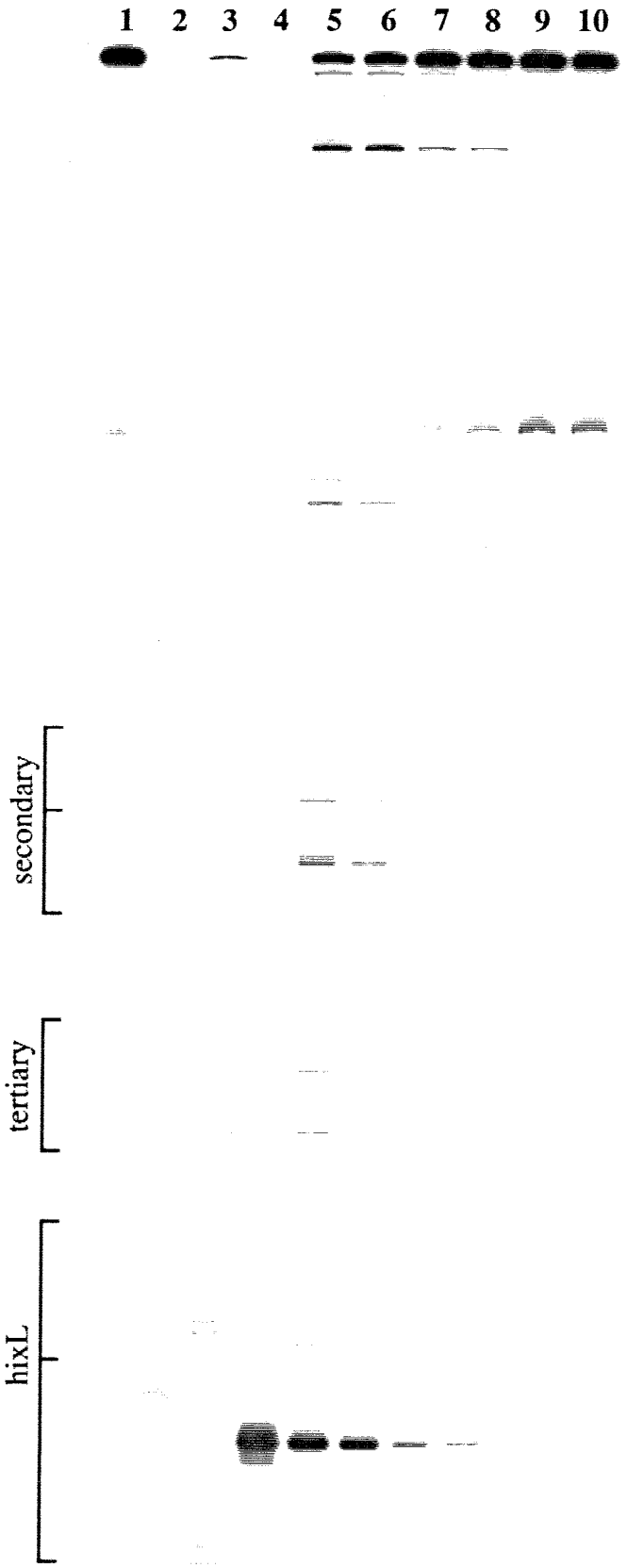

1 2 3 4 5 6 7 8 9 10
secondary
tertiary
hixL

Figure 6 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190) monoperoxyphthalic acid cleavage reaction peracid concentration study on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 5 μM $Ni(OAc)_2$,5 μM GGH(Hin 139-190), 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions were run for 15 min at 25 °C. Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 3'-end-labeled DNA was used. Lane 1 is a DNA control lane. Lane 2 is an A specific sequencing reaction. Lanes 3 thru 9 contain 1000, 500, 100, 50, 10, 5, and 1 μM monoperoxyphthalic acid, respectively.

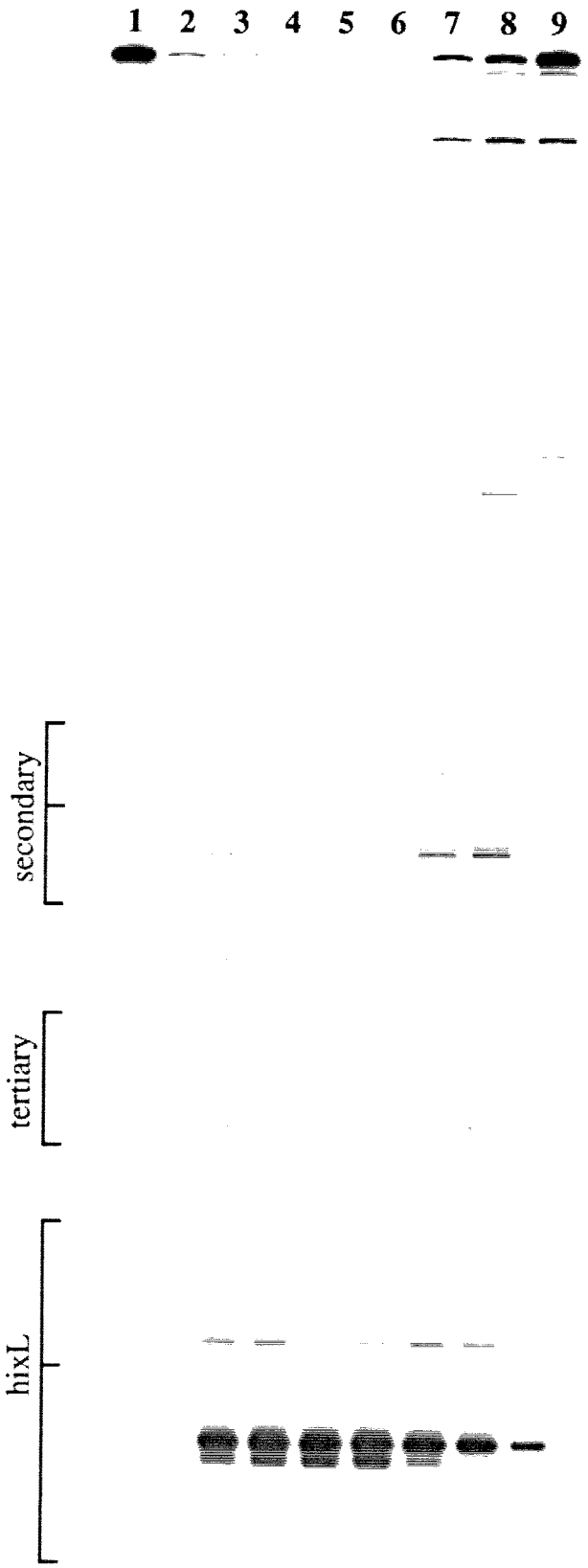

1 2 3 4 5 6 7 8 9
secondary
tertiary
hixL

Figure 7 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190)-sodium ascorbate-hydrogen peroxide DNA cleavage reaction pH study on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 5 μM $Ni(OAc)_2$, 5 μM GGH(Hin 139-190), 20 mM NaCl, 20 mM phosphate, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions were run for 15 minutes at 25 °C. Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 3'-end-labeled DNA was used. Lane 1 is a DNA control lane. Lane 2 is an A specific sequencing reaction. Lanes 3 thru 7 are at pH 6.0, 6.5, 7.0, 7.5, and 8.0, respectively.

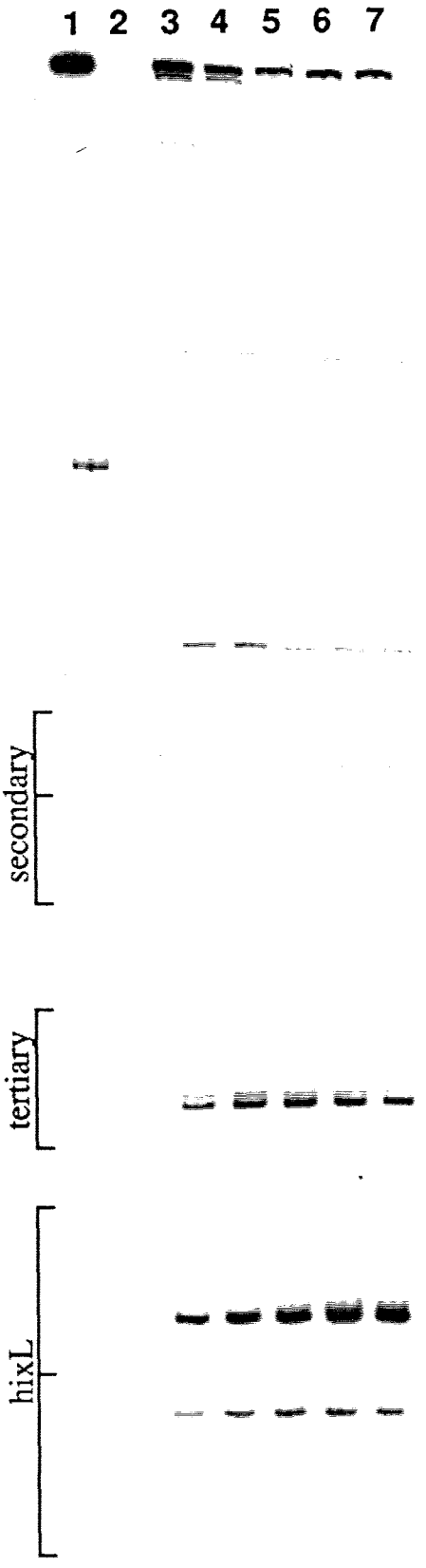
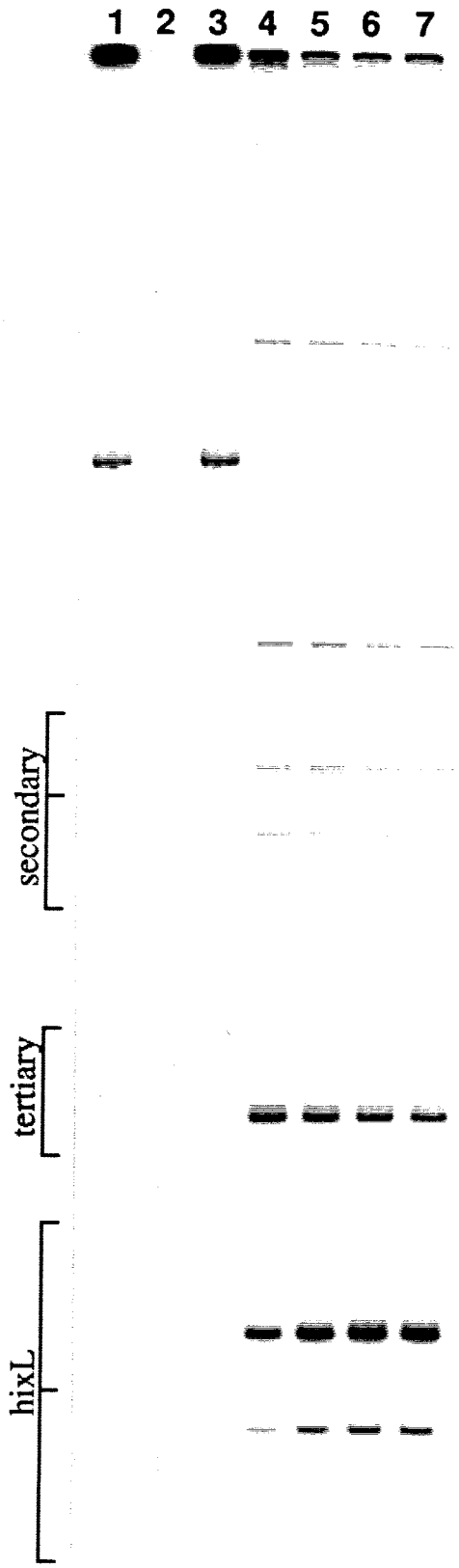

1 2 3 4 5 6 7
secondary
tertiary
hixL

Figure 8 Autoradiogram of high-resolution denaturing gel of Ni(II)•GGH(Hin 139-190) monoperoxyphthalic acid cleavage reaction peracid concentration study on a ^{32}P end-labeled fragment (XbaI/EcoRI) from pMFB36. Reaction conditions were 5 μM GGH(Hin 139-190), 20 mM NaCl, 20 mM phosphate, pH 7.5, calf thymus DNA (100 μM in base pair), and ≈15,000 cpm end-labeled DNA in a total volume of 20 μL. Reactions were run for 15 min at 25 °C. Nickel-mediated reactions were treated with 0.1 M $BuNH_2$ for 30 min at 90 °C. Cleavage products were analyzed on an 8%, 1:20 crosslinked, 50% urea polyacrylamide gel, 0.4 mm thick. 3'-end-labeled DNA was used. Lane 1 is a DNA control lane. Lanes 2 is an A specific sequencing reaction. Lanes 3 thru 9 contain 0.0, 0.5, 2.5, 5, and 10 μM $Ni(OAc)_2$, respectively.


1 2 3 4 5 6 7
secondary
tertiary
hixL

Appendix B

Mass Spectrometry and Circular Dichroism Studies

Attempts were made to better determine the physical properties of the synthetic proteins produced in this work. Two important questions about the synthetic proteins were posed: 1) Are the purified materials the correct compounds?, and 2) What secondary structural elements do the proteins contain? To begin to answer these questions mass spectrometry and circular dichroism studies were undertaken.

Mass Spectrometry. It has been shown by amino acid analysis and sequencing that the synthetic 52 residue protein, Hin(139-190), and the 55 residue protein, GGH(Hin 139-190), contain the correct amino acids in the desired sequence. Experiments of this nature, however, are not easily accomplished when the unnatural amino acid EDTA is incorporated into the protein. Another possible way to determine if the material has the correct composition is to obtain the molecular weight by mass spectrometry. Thus, the purified proteins, Hin(139-190), GGH(Hin 139-190), Hin(139-190, Gly^{139}-GABA-EDTA), Hin(139-190, Lys^{171}-EDTA), Hin(139-184), Hin(139-184, Lys^{183}-EDTA), and Hin(139-184, Gly^{139}-GABA-EDTA, Lys^{183}-EDTA) were examined by time-of-flight mass spectrometry,[1] to determine whether they correspond to the desired products. The expected and observed masses for the purified proteins are shown in Table 1.

Table 1 Mass Spectrometry Data for Synthetic Proteins

Protein	Expected Mass[a]	Observed Mass[b]
Hin(139-190)	6036.1	6040.7 ± 3.8
GGH(Hin 139-190)	6287.3	6276.8 ± 11.2
Hin(139-190, Gly^{139}-GABA-EDTA)	6395.4	6399.5 ± 7.7
Hin(139-190, Lys^{171}-EDTA)	6325.3	6372.5 ± 22.5
Hin(139-184)	5265.1	5255.5 ± 2.1
Hin(139-184, Lys^{183}-EDTA)	5580.4	5574.8 ± 4.6
Hin(139-184, Gly^{139}-GABA-EDTA, Lys^{183}-EDTA)	5939.7	5928.1 ± 3.0

a. Average mass

b. Error in mass calculated from $(M+H)^+$ and $(M+2H)^{2+}$

As can be seen from Table 1 all of the observed masses match fairly well to the expected masses. Since the accuracy of the measurement is from 0.1-1% (6-60 mass units for a protein of molecular weight 6000) it is not expected that the exact molecular weight would be obtained. Instead, what these experiments can show are significant differences in molecular weight between two compounds. For instance, the observed mass difference between Hin(139-190) and Hin(139-190, Gly^{139}-GABA-EDTA) is 359 mass units. This difference of the correct size and is larger than the expected experimental error, thus indicating that indeed the GABA-EDTA unit is attached to the protein. The expected differences between Hin(139-190) and GGH(Hin 139-190) as well as between Hin(139-184), Hin(139-184, Lys^{183}-EDTA), and Hin(139-184, Gly^{139}-GABA-EDTA, Lys^{183}-EDTA) can also be seen to be of the expected values.[2] The expected difference between Hin(139-190) and Hin(139-190, Lys^{171}-EDTA) is larger than expected, but not out of the limits of experimental error.[3] This data has thus shown that the synthetic proteins do indeed correspond to the correct species, ant that EDTA has been attached to the proteins in the desired manner.

Another interesting measurement which can be attempted using mass spectrometry is whether or not a metal is bound to the protein. In the affinity cleaving experiments carried out in this work, Fe is added to the EDTA-proteins in order to form an active cleaving function. Due to the scale of the reactions ($\leq 5 \times 10^{-9}$ moles of protein) no direct measurement for the presence of the EDTA-Fe complex is possible. If the complex has a high enough binding constant it may be possible to observe it by mass spectrometry. Table 2 indeed shows that the expected difference (observed = 58 mass units, expected = 56 mass units) exists for the EDTA-Fe complex, confirming its existence.

The mass spectrometry results described in this section represent the only analytical method available fro determining whether unnatural amino acids, such as EDTA, have been incorporated into proteins as expected from the experimental design. Further evidence has

Table 2 Mass Spectrometry Data for the Fe•Protein Complex		
Protein	**Expected Mass[a]**	**Observed Mass[b]**
Hin(139-190, Gly^{139}-GABA-EDTA)	6395.4	6399.5 ± 7.7
Hin(139-190, Gly^{139}-GABA-EDTA•Fe)	6451.3	6457.5 ± 5.1

a. Average mass

b. Error in mass calculated from $(M+H)^+$ and $(M+2H)^{2+}$

been presented which imply that this technique may also be useful for examining metal coordination in proteins.

Circular Dichroism. The data presented throughout this work is consistent with Hin(139-190) containing a helix-turn-helix binding motif. This data by itself, however, is not proof of the presence of such a structural unit. If such a motif does exist, it is expected that the protein would contain a significant amount of α-helical structure. To begin directly testing for the presence of secondary structural elements in Hin(139-190) circular dichroism (CD) studies were undertaken.

Figure 1 shows the CD spectrum of Hin(139-190) (at 20 μM, in 20 mM NaCl and 20 mM Phosphate buffer pH=7.5).[4] A significant amount of secondary structure is present in the protein in solution, as can be seen by the large absorbance between 194 and 250 nm. An estimate of the percentage of each secondary structural type which makes up such a spectrum can also be determined. Figure 2 shows two separate fits to the data.[5] Both fits indicate that a significant amount of α-helical structure exists in the protein along with other structural elements. The amount of helical structure observed is lower than expected, but fitting routines such as the one used, are limited by the reference curves used and have considerable error.

Although a quantitative assessment of the structural makeup of the protein is not possible other issues may be addressed by this technique. One such issue of interest is that since the amount a helical structure observed in solution is lower than expected, does the structure change upon binding to DNA? It is possible that the structure observed in

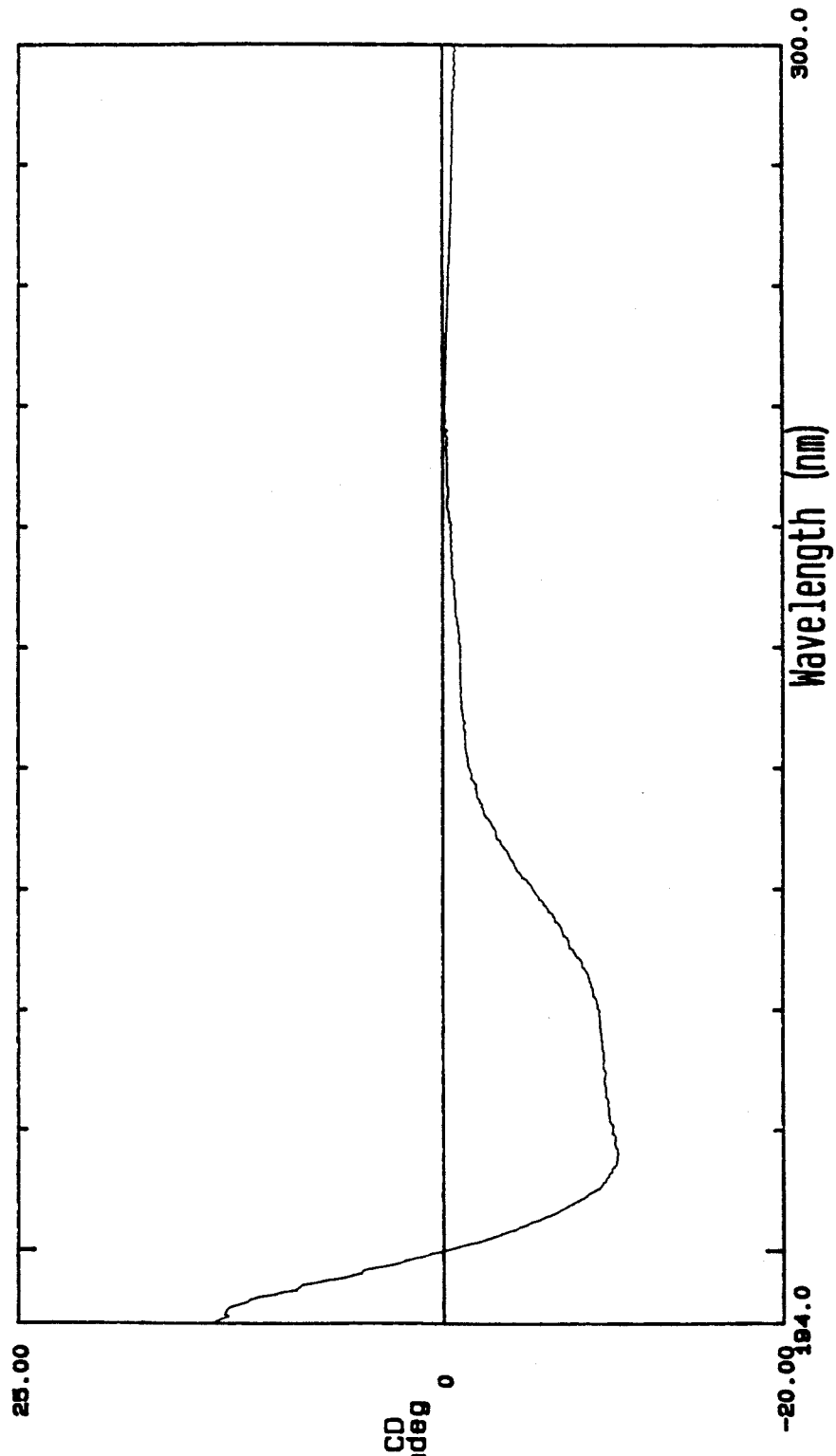


Figure 1 CD scectrum of Hin(139-190)

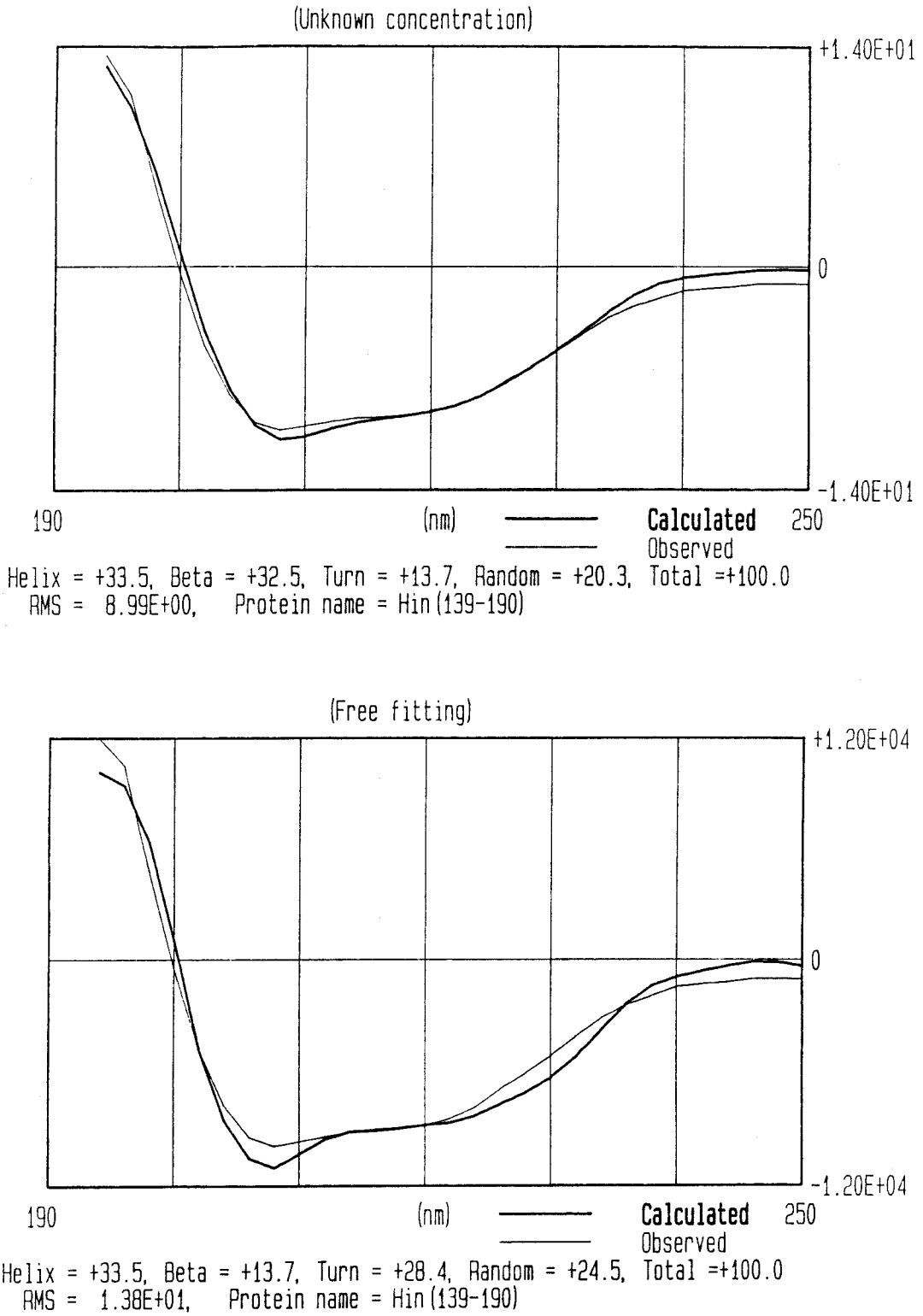


Figure 2 Computer fits to the data in Figure 1 indicating the percentage composition of secondary structural elements.

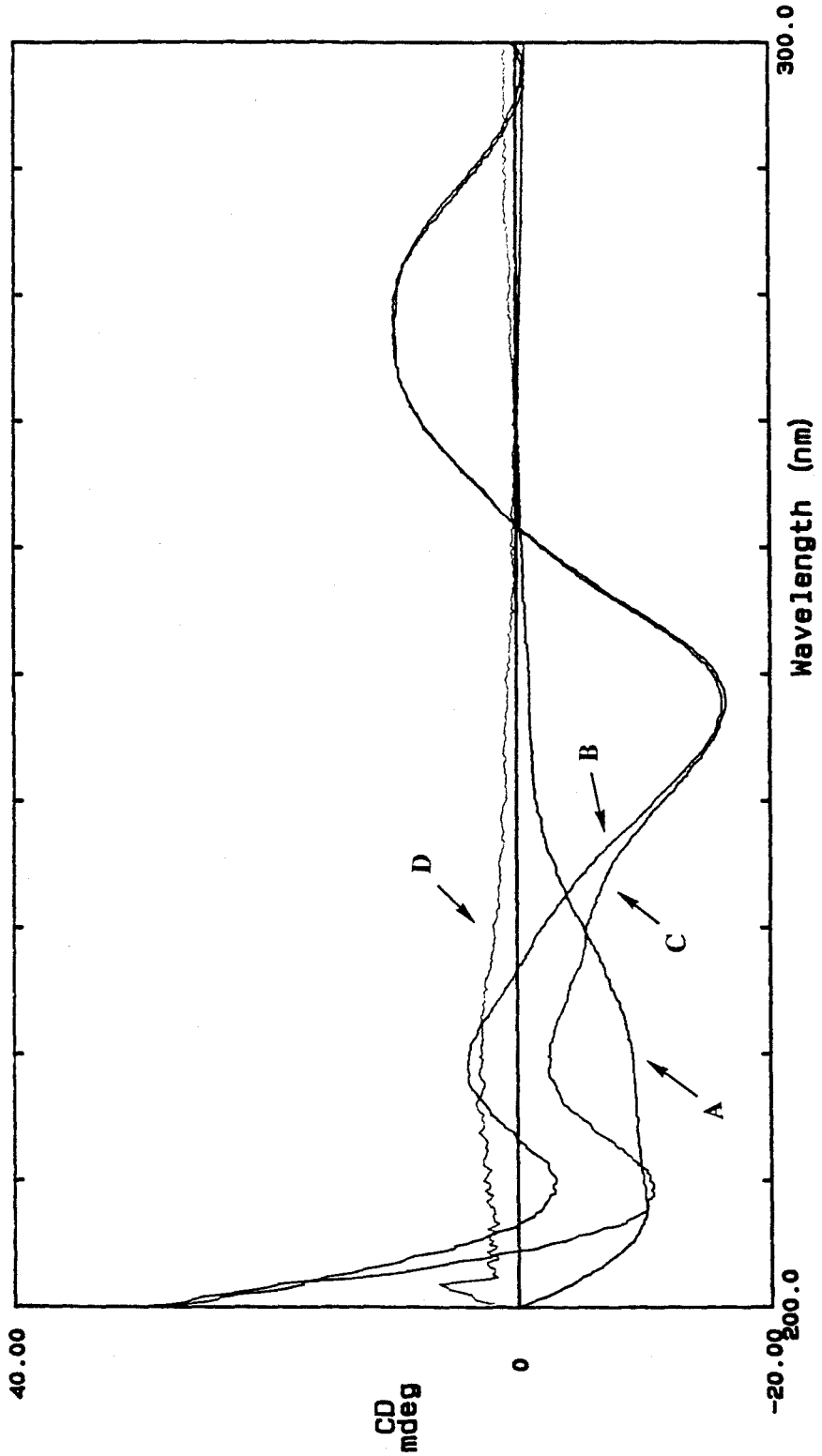


Figure 3 CD spectra of A) Hin(139-190) alone, B) DNA duplex alone, C) Hin(139-190)-DNA complex, D) Result of the subtraction of spectra A and B from C.

solution is not the helix-turn-helix motif and that upon binding to DNA the protein is induced to adopt the proposed structure. In order to test this hypothesis the CD spectrum of Hin(139-190) alone and in the presence of DNA, as well as the spectrum of the DNA alone were obtained (Figure 3).[6] As can be seen in Figure 3 when the spectra for DNA and protein alone are subtracted from that of the complex no additional signal, corresponding to changes in structure, are observed. This data then implies that the folded structure of the protein in solution is the same as that when the protein is bound to DNA. Other techniques such as NMR and crystallography will be needed to obtain a detailed picture of the folded structure of Hin(139-190).

References

1. Mass spectrum were obtained by Dr. David Teplow in the Biology Department at Caltech on a Bioion 20 time-of-flight plasma desorption mass spectrometer.
2. GGH = 251.2 mass units, Ser → Lys-EDTA = 315.3 mass units, GABA-EDTA = 359.3 mass units.
3. Ile → Lys-EDTA = 289.2 mass units.
4. CD measurements were performed on a Jasco J-600 spectrometer using a 0.1 mm path length cell.
5. Fits were obtained using a protein secondary structure estimation program supplied by Japan Spectroscopic Co., Ltd., which uses four reference curves; α-helix, β-sheet, β-turn, and random coil. The curves are determined from the CD spectra of proteins whose x-ray crystal structures are known.
6. The DNA sequence used was:

   5'-AAGCGACTAAAATTCTTCCTTATCTGATGTAAAGGAGAAAATCATGGCTACT
   TCGCTGATTTTAAGAAGGAATATACTACATTTCCTCTTTTAGTACCGATGAT-5'

   This sequence contains the secondary Hin binding site.