8.6: Nature's Use of Metal/Nucleic-acid Interactions

In the context of what we understand about the fundamental interactions and reactions of metal ions and complexes with nucleic acids, and also in comparison to how chemists have been exploiting these interactions in probing nucleic acids, we can also consider how Nature has taken advantage of metal ions in the construction of metalloproteins, nucleic-acid assemblies, and smaller natural products containing metal ions that interact with DNA and RNA.

A Structural Role

One of the chief functions attributed to metal ions in biological systems is their ability to provide a structural center to direct the folding of a protein. Just as shape-selective recognition has been helpful in targeting metal complexes to specific sites on DNA, it appears that one element of the recognition of sites by DNA regulatory proteins may also involve the recognition of complementary shapes. Furthermore, metal ions appear to be used in these proteins to define the shape or folding pattern of the peptide domain that interacts specifically with the nucleic acid.

The DNA-binding metalloproteins that have received the greatest attention recently have been the "zinc-finger" regulatory proteins. It was discovered in 1983 that zinc ions played a role in the functioning of the nucleic acid-binding transcription factor IlIA (TFIIIA) from *Xenopus laevis*, which binds specifically both DNA, the internal control region of the 5S rRNA gene, and RNA, the 5S RNA itself.[81] The protein (actually the 7S storage particle) was found to contain two to three equivalents of zinc ion. Dialysis removed both the associated zinc ions and the nucleic-acid-binding ability of the protein. Importantly, treatment with zinc ion, or in later studies with higher concentrations of Co^{2+}, restored the specific binding ability. Hence, zinc ion was shown to be functionally important in these eukaryotic regulatory proteins.

The notion of a "zinc-finger structural domain" was first provided by Klug and coworkers, after examination of the amino-acid sequence in TFIIIA.[82] It was found that TFIIIA contained nine imperfect repeats of a sequence of approximately 30 amino acids, and furthermore that each repeat contained two cysteine residues, two histidine residues, and three hydrophobic residues, in conserved positions. In addition, subsequent metal analyses were revealing higher zinc contents (7 to 11 equivalents) associated with the protein, and protein-digestion experiments indicated that several repeated structural domains existed in the protein. The two cysteine thiolates and two histidine imidazoles in each repeated domain could certainly serve to coordinate a zinc ion. Thus it was proposed that each peptide repeat formed an independent nucleic-acid-binding domain, stabilized in its folded structure through coordination of a zinc ion. The peptide unit was termed a "zinc finger," which is illustrated schematically in Figure 8.17. TFIIIA was therefore proposed to contain nine zinc fingers, which would cooperatively bind in the internal control region of the 5S RNA gene.

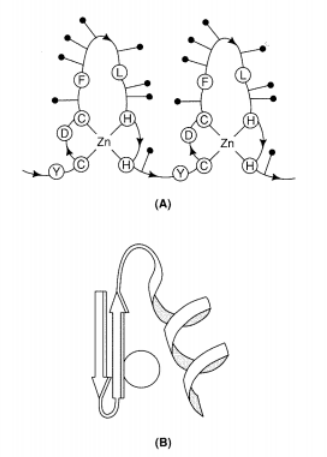

Figure 8.17 - (A) A schematic of a zinc-finger peptide domain.[82] (B) The proposed schematic structure of a zinc-finger domain based on comparisons to other structurally characterized metalloproteins.[85]

An enormous number of gene sequences from a variety of eukaryotic regulatory proteins was then found to encode strikingly similar amino-acid sequences,[83] and many were dubbed zinc-finger proteins. The bioinorganic chemist, however, should be aware that chemical analyses supporting such assignments are first required. Nonetheless, several legitimate examples of eukaryotic nucleic-acid-binding zinc-finger proteins containing multiple zinc-binding peptide domains have emerged since the first study of TFIIIA, including the proteins Xfin from *Xenopus*, the Kruppel protein from *Drosophila*, the Sp1 transcription factor, and human

testes-determining factor. It has therefore become clear that *the zinc-finger domain represents a ubiquitous structural motif for eukaryotic DNA-binding proteins.*[84]

What is the structure of a zinc finger, and how is this structure important for binding a specific nucleic-acid site? Based on a search of crystallographic databases for metalloproteins and an examination of the consensus sequence emerging for zinc fingers (that is, which residues were truly conserved and common to the different putative zinc fingers), Berg proposed a three-dimensional structure for a zinc finger, shown schematically in Figure 8.17.[85] The proposed structure included the tetrahedral coordination of zinc by the two cysteine and histidine residues at the base of the finger and an $alpha$-helical region running almost the length of the domain. EXAFS studies also supported the tetrahedral zinc site. Since this proposal, two detailed two-dimensional NMR studies have been reported that are consistent with the tetrahedral zinc coordination and the a-helical segment.[86] More recently, a crystal structure of a three-finger binding domain associated with an oligonucleotide was determined.[87] The zinc fingers lie in the major groove of DNA, the α-helical region being within the groove. Not surprisingly, given basic coordination chemistry, the zinc does not interact directly with the nucleic acid. Instead, the zinc ion must serve a structural role, defining the folding and three-dimensional structure of the protein scaffolding about it. This structure, defined by the metal at its center, like other coordination complexes, is able to recognize its complementary structure on the nucleicacid polymer.

It should also be noted that this zinc-finger structural motif is not the only metal-containing or even zinc-containing structural motif important in nucleic-acid-binding proteins.[88] A clearly different domain is evident in the protein GAL4, a transcription factor required for galactose utilization in *S. cerevisiae*.[88a] A recent crystal structure of the protein bound to an oligonucleotide shows the protein to bind to DNA as a dimer; each monomer contains a binuclear zinc cluster with two zinc ions tetrahedrally coordinated by six cysteines (two cysteines are bridging), not dissimilar from proposed structures in metallothionein. Still another structural motif is found in the glucocorticoid receptor DNA-binding domain. Crystallography[89] has revealed that this domain also binds DNA as a dimer; here each monomer contains two zinc-nucleated substructures of distinct conformation. The zinc ions are each tetrahedrally coordinated to four cysteine residues. Likely this too represents another structural motif for proteins that bind nucleic acids, and one again in which the metal serves a structural role.

Lastly, one might consider why zinc ion has been used by Nature in these nucleic-acid binding proteins. Certainly, the natural abundance of zinc is an important criterion. But also important is the absence of any redox activity associated with the metal ion, activity that could promote DNA damage [as with Fe(II) or Cu(II), for example]. In addition, other softer, heavier metal ions might bind preferentially to the DNA bases, promoting sequence-specific covalent interactions. Zinc ion, therefore, is clearly well-chosen for the structural center of these various nucleic-acid-binding proteins.

A Regulatory Role

Metalloregulatory proteins, like the transcription factors described above, affect the expression of genetic information through structural interactions that depend upon the metal ions, but unlike the zinc-finger proteins, metalloregulatory proteins act as triggers, repressing or activating transcription given the presence or absence of metal ion. In some respects, even more than zinc fingers, these systems resemble the Ca^{2+}-activated proteins described in Chapter 3.

Consider the biological system that must respond to changing intracellular metal concentrations. At high concentrations many metal ions become toxic to the cell; hence, a full system of proteins must be synthesized that will chelate and detoxify the bound-metal-ion pool. In order to actively engage these proteins, the genes that encode them must be rapidly transcribed. But at the same time, the DNA itself must be protected from the high concentrations of metal ion. Hence the need for these metalloregulatory proteins, which bind DNA in the absence of metal ion, usually repressing transcription, but in the presence of metal ion bind the metal ion tightly and specifically, and as a consequence amplify transcription.

Perhaps the best-characterized metalloregulatory system thus far is the MerR system, regulating mercury resistance in bacteria.[90] An inducible set of genes arranged in a single operon is under the control of the metal-sensing MerR protein, and it is this system that mediates mercury resistance. Mercury resistance depends upon the expression of these genes to import toxic Hg(lI), reduce Hg(II) to the volatile Hg(0) by NADPH, and often additionally to cleave organomercurials to their corresponding hydrocarbon and Hg(II) species. The MerR protein regulates mercury resistance both negatively and positively. As illustrated in Figure 8.18, MerR in the absence of Hg(II) binds tightly and sequence-specifically to the promoter. In so doing, MerR inhibits binding to the promoter by RNA polymerase. When Hg^{2+} is added at low concentrations, the metal ion binds specifically and with high affinity to the DNA-bound MerR, and causes a DNA conformational change detectable by using other metal reagents as conformational probes. This conformational change now facilitates the binding of RNA polymerase and hence activates expression of the gene family.

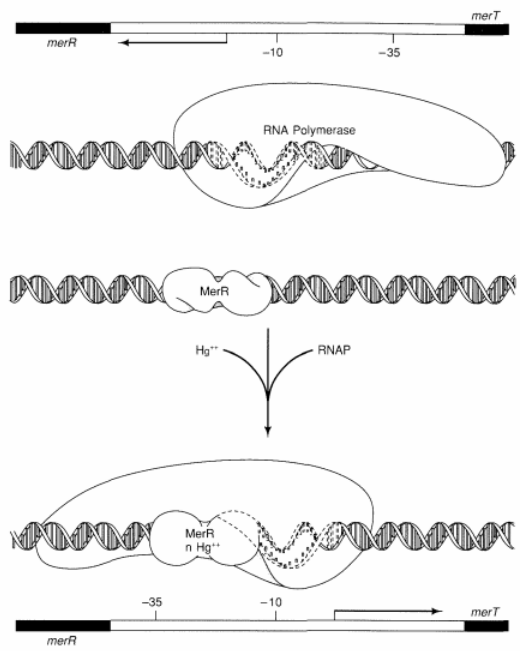

Figure 8.18 - A model for MerR metalloregulation.[90a] In the absence of MerR, RNA polymerase binds and transcribes the MerR promoter. In the presence of MerR, the preferential binding of MerR to the promoter is observed that inhibits transcription by the polymerase. The addition of Hg^{2+} then leads to a conformational change that promotes binding of the polymerase, substantially increasing transcription. Reproduced with permission from Reference 90a.

What are the structural requirements in the metal-binding site? Certainly one requirement is Hg(II) specificity, so that other metal ions will not also trigger transcriptional activation. Another is high metal-binding affinity to protect the DNA from direct coordination of the Hg(lI). The MerR protein is dimeric, and contains four cysteine residues per monomer. Site-directed mutagenesis studies[91] have indicated that three of these four cysteine residues are needed for Hg(II) binding, and EXAFS studies[92] have been consistent with tricoordinate ligation, clearly a well-designed system for Hg(II) specificity. Perhaps even more interesting, the site-directed mutagenesis studies[91] on heterodimers (a mixture of mutant and wild-type monomers) have indicated that the coordinated Hg(II) bridges the dimer, ligating two cysteines of one monomer and one cysteine of the other. This scheme may provide the basis also for the kinetic lability needed in a rapidly responsive cellular system.

Model systems are also being constructed to explore metal modulation of DNA binding. One system involves the assembly of two dipeptides linked by a central acyclic metal-binding polyether ligand, with $Fe(EDTA)^{2-}$ tethered to one end to mark site-specific binding.[93] In the presence of alkaline earth cations, which induce a conformational change that generates a central macrocycle, the linked peptides become oriented to promote sequence-specific binding in the minor groove. In the absence of the alkaline earth ion, no site-specific binding, or cleavage of DNA, is evident. One might consider this system as a simple, first-order synthetic model for the metalloregulatory proteins.

MerR is clearly only one natural metalloregulatory system. Other metal ions bind regulatory factors to mediate the regulation of gene expression in a metalspecific manner. Two examples include the Fe(II)-binding Fur protein from enteric bacteria[94] and the copper-binding protein ACE1/CUP2 from *S. cerevisiae*.[95] Both copper and iron are essential trace elements for which high concentrations are toxic; for nucleic acids this toxicity is certainly the result of redox-mediated strand damage. Other metal-specific regulatory systems are surely present as well. Both the MerR and the synthetic system may exemplify how these various systems function, how Nature might construct a ligand system to facilitate toxicmetal-specific binding in the presence of DNA that then alters or triggers how other moieties bind and access the nucleic acid.

A Pharmaceutical Role

With the exception of cisplatin (see Chapter 9), most pharmaceuticals currently being used as DNA-binding agents were first isolated as natural products from bacteria, fungi, plants, or other organisms. For the most part they represent complex organic moieties, including peptide and/or saccharide functionalities, and often a unique functionality, such as the ene-diyne in calichimycin. These natural products bind DNA quite avidly, through intercalation, groove binding, or a mixture thereof. Often the efficacy of these antitumor antibiotics stems from subsequent alkylation or DNA strand-cleavage reactions that damage the DNA.

Among the various natural products used clinically as antitumor antibiotics are bleomycins, a family of glycopeptide-derived species isolated from cultures of *Streptomyces*.[25,96] The structure of bleomycin A_2 is shown schematically in Figure 8.19. The molecular mode of action of these species clearly involves binding to DNA and the promotion of single-stranded cleavage at GT and GC sequences. Importantly, as demonstrated by Horwitz, Peisach, and coworkers, this DNA cleavage requires the presence of Fe(II) and oxygen.[97] Thus, one might consider Fe-bleomycin as a naturally occurring inorganic pharmaceutical.

What is the role of the metal ion in these reactions? As one might imagine, based upon our earlier discussions of metal-promoted DNA cleavage, the iron center is essential for the oxidative cleavage of the strand through reaction with the sugar moiety. The reaction of Fe(II)-bleomycin can, however, clearly be distinguished from the $Fe(EDTA)^{2-}$ reactions discussed earlier in that here no diffusible intermediate appears to be involved. Instead of generating hydroxyl radicals, the Fe center must be positioned near the sugar-phosphate backbone and activated in some fashion to promote strand scission *directly*.

Despite extensive studies, in fact little is known about how Fe(II)-bleomycin is oriented on the DNA. Indeed, the coordination about the metal is the subject of some debate. The structure of Cu(II)-P-3A,[98] a metallobleomycin derivative, is also shown in Figure 8.19. On the basis of this structure and substantive spectroscopic studies on Fe-bleomycin itself, it is likely that, as with Cu(II), in the Fe(II)-bleomycin complex the metal coordinates the β-hydroxyimidazole nitrogen, the secondary amine of β-aminoalanine, and the N1 nitrogen of the pyrimidine. Whether in addition the primary amines of the amino alanine and of the histidine coordinate the metal is still not settled. Possibly bithiazole coordination or some coordination of the sugar moieties is involved. Nonetheless, given five different coordination sites to the bleomycin, the sixth axial site is available for direct coordination of dioxygen. How is this Fe—O_2 complex oriented on the DNA? It is likely that at least in part the complex binds against the minor groove of the helix. There is some evidence that suggests that the bithiazole moiety intercalates in the helix. It is now becoming clear, however, that the structure of the metal complex itself, its three-dimensional shape, rather than simply the tethered bithiazole or saccharide, is needed for the sequence selectivity associated with its mode of action.

Although the coordination and orientation of the metal complex are still not understood, extensive studies have been conducted concerning the remarkable chemistry of this species.[96,99] The overall mechanism of action is described in Figure 8.19. In the presence of oxygen, the Fe(Il) O_2 species is formed and is likely rapidly converted to a ferric superoxide species. The one-electron reduction of this species, using either an organic reductant or another equivalent of Fe(II)-bleomycin, leads formally to an Fe(III)-peroxide, which then undergoes O—O bond scission to form what has been termed "activated bleomycin." This species might be best described as Fe(V)=O (or $[Fe\ O]^{3+}$). This species is comparable in many respects to activated cytochrome P-450 or perhaps even more closely to the Fe center in chloroperoxidase (see Chapter 5). Like these systems, activated bleomycin can also epoxidize olefins and can generally function as an oxo transferase. In contrast to these systems, Fe-bleomycin clearly lacks a heme. How this species can easily shuttle electrons in and out, forming and reacting through a high-valent intermediate, without either the porphyrin sink or another metal linked in some fashion, is difficult to understand. In fact, understanding this process, even independently of our fascination with how the reaction is exploited on a DNA helix, forms the focus of a substantial effort of bioinorganic chemists today.

What has been elucidated in great detail is the reaction of activated bleomycin with DNA. It has been established that the activated species promotes hydrogen abstraction of the C4'-H atom, which is positioned in the minor groove of the helix (Figure 8.19). Addition of another equivalent of dioxgen to this C4'-radical leads to degradation of the sugar to form a 5'-phosphate, a 3'-phosphoglycolate, and free base propenal. Alternatively, oxidation of the C4'-radical followed by hydroxylation in the absence of oxygen yields, after treatment with base, a 5'-phosphate, an oxidized sugar phosphate, and free base.

Other metals such as copper and cobalt can also activate bleomycins, although their mechanistic pathways for strand scission are clearly different from that of Fe(II)-bleomycin. Whether other natural products that bind DNA also chelate metal ions and exploit them for oxidative strand cleavage is not known, but several systems provide hints that they do. Furthermore, such a fact would not be surprising given our understanding of the utility of metal ions in promoting this chemistry. An even more detailed understanding of this chemistry might lead to the development of second-generation synthetic transition-metal pharmaceuticals that specifically and efficiently target and cleave DNA sites.

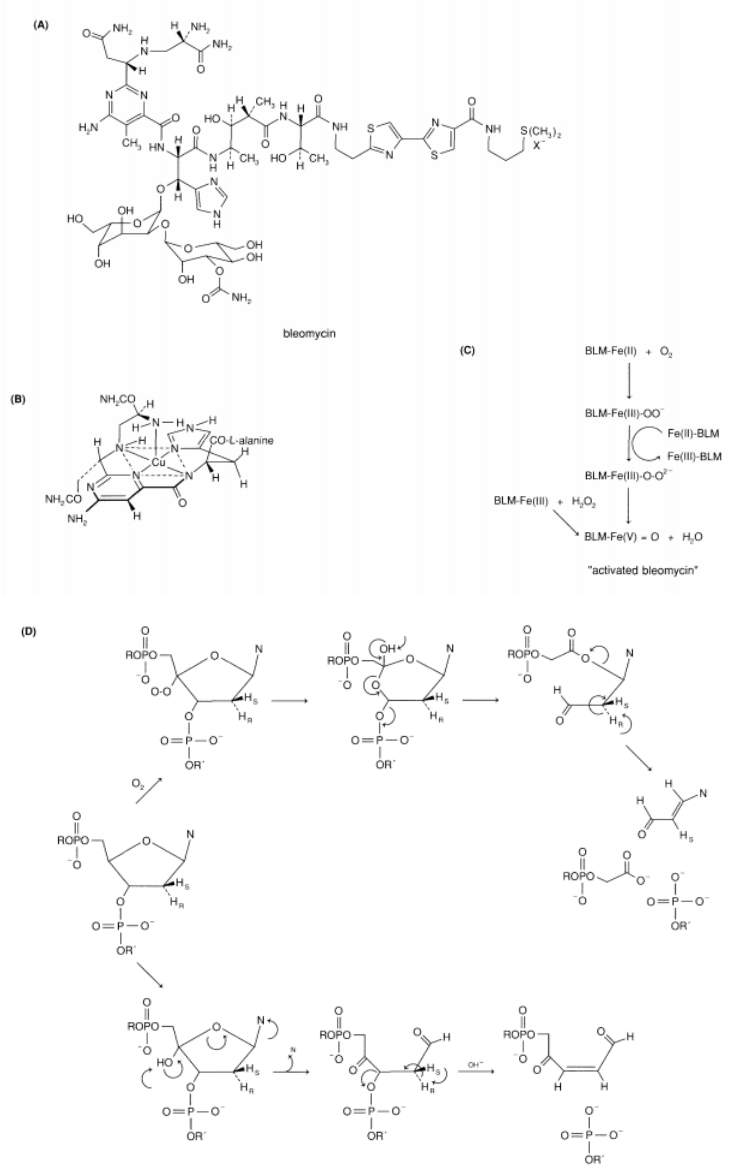

Figure 8.19 - (A) The structure of bleomycin A_2. (B) The crystallographically determined structure[98] of a copper derivative of bleomycin, P3A. (C) A scheme to generate "activated bleomycin." (D) The proposed mechanism of action of activated Fe-bleomycin.

A Catalytic Role

In addition to serving structural and modulating roles in proteins which bind nucleic acids, metal ions also appear to be essential to the functioning of various complex enzymes that act on nucleic acids. At this stage our understanding of the participation of the metal ion in the catalytic chemistry of these enzymes is somewhat sketchy, and we are relying more on our current understanding of the possible roles where metal ions may prove advantageous. These remain areas of biochemical focus where the inorganic chemist could make a major contribution.

For example, zinc ion appears to be essential to the functioning of both RNA polymerases and DNA topoisomerases.[100-102] These multisubunit enzymes perform quite complex tasks. RNA polymerase must bind site-specifically to its DNA template, bind its nucleotide and primer substrates, and form a new phosphodiester bond in elongating the growing RNA. Two zinc ions appear to be involved. One may be involved in orienting the nucleotide substrate, and the other structurally in template recognition. It would not be surprising, however—indeed, it might be advantageous—if one or both metal ions also participated in the polymerization step. Our mechanistic understanding of how topoisomerases function is even more cursory. These complex enzymes bind supercoiled DNA, sequentially break one strand through hydrolytic chemistry, move the strand around the other (releasing one tertiary turn),

and religate the strand. Again, the zinc ion might participate in the hydrolytic chemistry, the ligation step, or both; alternatively, the metal might again serve a structural role in recognition of the site of reaction.

We do have some understanding of the role of metal ions in several endonucleases and exonucleases. As discussed in Section II.C, metal ions may effectively promote phosphodiester hydrolysis either by serving as a Lewis acid or by delivering a coordinated nucleophile. Staphylococcal nuclease[103] is an extracellular nuclease of *Staphylococcus aureus* that can hydrolyze both DNA and RNA in the presence of Ca^{2+}. The preference of the enzyme is for single-stranded DNA, in which it attacks the 5'-position of the phosphodiester linkage, cleaving the 5'-P-O bond to yield a 5'-hydroxyl and 3'-phosphate terminus. Ca^{2+} ions are added as cofactors and are strictly required for activity. The structure of staphylococcal nuclease, determined by x-ray crystallography and crystallized in the presence of Ca^{2+} and the enzyme competitive inhibitor pdTp, as well as subsequent NMR and EPR studies on mutant enzymes using Mn^{2+} as a substitute for the Ca^{2+} ion, have provided the basis for a detailed structural analysis of the mechanism of this enzyme. In this phosphodiester hydrolysis, the metal ion appears to function primarily as an electrophilic catalyst, polarizing the P-O bond, and stabilizing through its positive charge the evolving negative charge on the phosphorus in the transition state. The base is thought here not to be directly coordinated to the metal; instead, action of a general base is invoked.

Metal ions also participate in the functioning of other nucleases, although the structural details of their participation are not nearly as established as those for staphylococcal nuclease. DNAse I also requires Ca^{2+} for its catalytic activity.[104] S1 endonuclease, mung bean nuclease, and *Physarum polycephalem* nuclease require zinc ion either as cofactors or intrinsically for nuclease activity, and the restriction enzyme EcoRI may also require intrinsically bound zinc ion.[33] In terms of how the zinc ion might function in these enzymes, one can look both to staphylococcal nuclease and to bacterial alkaline phosphatase[105] for some illustrations. One would expect that this metal ion could serve both as an electrophilic catalyst and also in the delivery of a zinc-coordinated hydroxide, as it does in alkaline phosphatase, directly attacking the phosphate ester. More work needs to be done to establish the mechanisms by which zinc ion promotes phosphodiester hydrolysis in these enzymatic systems.

Probably most intriguing and mysterious at this stage is the metal participation in the very complex DNA-repair enzyme endonuclease III from *E. coli* (similar enzymes have also been isolated from eukaryotic sources). This enzyme is involved in the repair of DNA damaged by oxidizing agents and UV irradiation, and acts through an N-glycosylase activity to remove the damaged base, and through an apurinic/apyrimidinic endonuclease activity to cleave the phosphodiester bond adjacent to the damaged site. Although more complex in terms of recognition characteristics, this enzyme functions in hydrolyzing the DNA phosphodiester backbone. What is so intriguing about this enzyme is that it contains a 4Fe-4S cluster (see Chapter 7) that is essential for its activity![106] We think generally that Fe-S clusters best serve as electron-transfer agents. In the context of this repair enzyme, the cluster may be carrying out both an oxidation and a reduction, to effect hydrolysis, or alternatively perhaps a completely new function for this metal cluster will emerge. (Fe-S clusters may represent yet another structural motif for DNA-binding proteins and one which has the potential for regulation by iron concentration.) Currently the basic biochemical and spectroscopic characterization of the enzyme is being carried out. Understanding this very novel interaction of a metal center and nucleic acid will require some new ideas, and certainly represents one new challenge for the bioinorganic chemist.
