Cold Spring Harbor Symposia on Quantitative Biology

Metal Ions in Proteins: Structural and Functional Roles

J.M. Berg

Cold Spring Harb Symp Quant Biol 1987 52: 579-585
Access the most recent version at doi:10.1101/SQB.1987.052.01.066

Metal Ions in Proteins: Structural and Functional Roles

J.M. BERG
Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218

Metal ions play important roles in many biological systems. In particular, metal ions are required for the activity of a large number of enzymes and proteins. Metal ions are well suited for these functions because of the following properties: (1) Metal ions are almost always positively charged and, hence, electrophilic. They can act as Lewis acids for binding and activating substrates. (2) Many metals can exist stably in a number of different oxidation states differing by one or by several units. This allows these metals to participate in various types of oxidation-reduction processes. (3) Metal ions generally bind four or more ligands. By binding several protein side chains, metals can act as multidentate cross-linking agents.

In some metalloproteins, the coordination sphere around the metal ion is relatively static in that the coordination number and gross coordination geometry do not change as the protein executes its function. In others, the coordination environment is quite dynamic with changes in geometry, including changes in the ligands bound, playing an essential role in protein function. Furthermore, as noted above, some proteins contain metal ions that undergo changes in oxidation level, whereas others have ions that remain in a single oxidation state. Consideration of the static versus dynamic natures of the coordination spheres and oxidation states of various metal ions in proteins can be used to develop the classification scheme shown in Figure 1.

Metal Ions In Proteins

Oxidation State \ Coordination Sphere	Static	Dynamic
Static	**Binding Proteins** Aspartate Transcarbamylase Regulatory Subunit (Zn) Calmodulin (Ca)	**Hydrolases** Carboxypeptidase A (Zn) Urease (Ni)
Dynamic	**Electron Carriers** Cytochrome c (Fe) Plastocyanin (Cu)	**Oxidoreductases and Oxygen Carriers** Cytochrome P-450 (Fe) Hemocyanin (Cu)

Figure 1. Classification scheme for the functions of metal ions in proteins according to the static versus dynamic natures of their coordination spheres and oxidation levels.

The first class includes proteins for which both the coordination sphere and the oxidation state change during the course of function. The most obvious members of this class are oxidoreductases, enzymes that execute oxidation-reduction reactions of substrates. A specific example is cytochrome P-450. In the cytochromes P-450, an iron atom is bound to the four nitrogens of a porphyrin ring and to a thiolate group from a protein cysteine residue with the sixth position in an octahedral structure available for substrate binding (Poulos et al. 1985). During the catalytic cycle the iron undergoes the following changes: $Fe^{+++} \rightarrow Fe^{++} \rightarrow (Fe^{++}[O_2] \leftrightarrow Fe^{+++}[O_2^-]) \rightarrow (Fe=O)^{+++} \rightarrow Fe^{+++}$ (Coon and White 1980). Thus, the dynamic nature of the coordination sphere is required for dioxygen binding, and changes in oxidation state are required for dioxygen binding and activation. Another less obvious group of proteins in this class are the oxygen carriers. Although changes in the coordination sphere are clearly required for dioxygen binding, the necessity for changes in oxidation state is not as apparent. However, for the three well-characterized types of oxygen carriers, dioxygen is not bound as O_2^0 but as a more reduced form (O_2^-, O_2^{--}, or HO_2^-), indicating that an internal oxidation-reduction process has occurred. In hemocyanin, for example, the deoxy form has two Cu(I) ions, whereas the oxy form contains two Cu(II) ions bridged by a peroxide ligand (Eickman et al. 1979).

The second class consists of proteins in which the coordination sphere is dynamic, but the oxidation state of the metal ion remains constant. The prototypical examples of such proteins are the zinc(II)-containing peptidases such as carboxypeptidase A (Vallee et al. 1983). In these proteins, the zinc ions act as Lewis acids to bind and polarize substrate carbonyl groups and/or to bind and activate water/hydroxide. Since zinc has only one stable oxidation level, no oxidation-reduction process can be involved. Another example is provided by urease, which has an active site that appears to contain two Ni(II) ions (Blakeley and Zerner 1984). Although oxidation-reduction processes are possible with nickel, a mechanism has been proposed in which one Ni(II) ion acts to bind and activate urea and the other activates a bound hydroxide ion for attack (Blakeley and Zerner 1984).

The third class includes those proteins in which the coordination sphere remains relatively static while the metal ion is involved in an oxidation-reduction process. These proteins are generally electron carriers involved in transferring single electrons. Examples include cyto-

chrome *c* and plastocyanin. The coordination spheres of the metal ion in these proteins are changed only very slightly as the metal ions changes oxidation state. Indeed, this small structural change is certainly important in facilitating fast electron transfer by minimizing the reorganizational energy component to the activation energy.

The final class of protein is that in which both the coordination sphere and the oxidation state are fixed. In these proteins, the metal ions act to cross-link several amino acid side chains and hence to stabilize a particular protein structure. Thus, the metal ions play a purely structural role, but since three-dimensional structure and function are intimately connected, they have dramatic effects on protein function as well. Examples of such proteins include calcium-binding proteins such as calmodulin, and zinc proteins such as the regulatory subunit of aspartate transcarbamylase. In the calcium-binding proteins, the Ca^{++} ions are bound in an α helix-loop–α helix structure termed the EF-hand with the ligands all from a short stretch (about 12 residues) of sequence (Seamon and Kretsinger 1983). These proteins respond to changes in calcium concentration in a manner that is fundamental to the function of Ca^{++} as a second messenger. In aspartate transcarbamylase, the Zn^{++} is bound by four cysteinate residues from a single protein domain that is involved in interactions between the regulatory and catalytic subunits (Honzatko et al. 1982); removal of the Zn^{++} causes dissociation of the holoenzyme (Nelbach et al. 1972).

Recognition of the importance of "structural zinc" has been greatly increased recently by the discovery that zinc appears to play such a role in several classes of proteins involved in nucleic acid binding and gene regulation. This development has been made possible largely by the observation that these proteins contain characteristic patterns of cysteine and histidine residues within relatively short segments of amino acid sequence, suggesting the presence of small structural domains organized around bound metal ions. The remainder of this paper will discuss the structural and functional roles of such metal-binding domains.

RESULTS AND DISCUSSION

Transcription Factor IIIA and the "Zinc Finger" Proteins

The occurrence of metal-binding domains in nucleic-acid-binding proteins was first hypothesized based on analysis of the sequence of the protein transcription factor IIIA from *Xenopus* (TFIIIA). This protein binds both the approximately 50-bp internal control region of the 5S RNA genes to activate transcription and the 5S RNA molecule itself in a 1:1 complex that acts as a storage particle (Pelham and Brown 1980). The protein sequence was determined from a cDNA clone (Ginsberg et al. 1984), and two groups reported that the sequence contained nine tandem imperfect repeats of an approximately 30-amino-acid sequence with two conserved cysteine residues and two conserved histidine residues with some hydrophobic residues conserved as well (Brown et al. 1985; Miller et al. 1985). The consensus sequence of these repeats is (Phe,Tyr)-X - Cys - $(X)_4$ - Cys - $(X)_3$ - Phe - $(X)_5$ - Leu - $(X)_2$ - His - $(X)_3$ -His-$(X)_5$. Furthermore, it was shown that the storage particle contains 7–11 zinc ions if it was isolated under conditions that avoided the use of chelating agents (Miller et al. 1985), extending an earlier report that the particle and protein contained 2–3 zinc ions (Hanas et al. 1983). Based on these observations, it was proposed that each repeat bound a zinc ion tetrahedrally through the cysteine and histidine residues to form an independent structural domain that was termed a "zinc finger" (Miller et al. 1985).

A variety of evidence supports the existence of these domains. First, the protein was known to have a highly asymmetrical shape, based on sedimentation studies (Bieker and Roeder 1984) and on the fact that a single protein molecule protected about 50 bp of DNA from digestion (Bieker and Roeder 1984; Smith et al. 1984). The hypothesis that three-quarters of the protein are built of nine relatively independent structural units fits nicely with these observations. Second, limited proteolysis of the storage particle yielded fragments with sizes consistent with the existence of domains of approximately 3–3.5 kD (Miller et al. 1985). Third, the structure of the TFIIIA gene has been determined, and it was found that introns occur between repeats 1–2,2–3,3–4,4–5,5–6, and 6–7, strongly suggesting that these units have structural significance (Tso et al. 1986). Fourth, a 30-amino-acid peptide corresponding to one repeat has been prepared, and it has been shown to fold into a stable structure in the presence of Zn^{++} or similar divalent metal ions, but not in their absence (Frankel et al. 1987). Finally, spectroscopic studies of the metal-binding site are consistent with the hypothesis. X-ray absorption studies of the storage particle suggest that each zinc has two nitrogen ligands at 2.0 Å and two sulfur ligands at 2.3 Å (Diakun et al. 1986). Furthermore, the spectroscopic properties of the Co^{++} complex of the single-repeat peptide suggest a tetrahedral site and are entirely consistent with the proposed structure (Frankel et al. 1987).

Subsequently, several other proteins have been discovered to contain quite similar sequences. The consensus sequences are essentially identical, except that two residues are present between the cysteine residues rather than four as observed in TFIIIA. Four of these were discovered in a manner entirely independent of the knowledge of the zinc finger sequence. These include the product of the *Drosophila* segmentation gene *Krüppel* (Rosenberg et al. 1986), the β and δ open reading frames from the *Drosophila serendipity* locus (Vincent et al. 1985), and the yeast positive control protein ADR1 (Hartshorne et al. 1986). More recently, several open reading frames have been isolated using nucleic acid probes derived from the *Krüppel* gene. These include one from *Drosophila* (Kr h) (Schuh et al. 1986) and two from mouse (MKR1 and

MKR2) (Chowdhury et al. 1987). In addition, hybridization studies using *Krüppel*-derived probes and studies using antisera against a 12-amino-acid peptide derived from *Krüppel* indicate that a wide variety of eukaryotes have nuclear proteins that contain the zinc finger motif (Schuh et al. 1986).

Other Potential Metal-binding Domains

The widespread occurrence of TFIIIA-like metal-binding sequences raised the question of whether other classes of metal-binding domains also occur in other nucleic-acid-binding proteins. Such domains would be similar to the TFIIIA-like structures in that they would have four metal-binding residues within a short stretch of amino acid sequence, but they would not be homologous to the TFIIIA-like sequences in a formal sense. To address this question, a systematic search procedure was devised (Berg 1986). A search template of the form Cys-$(X)_{2-4}$-Cys-$(X)_{2-15}$-a-$(X)_{2-4}$-b or a-$(X)_{2-4}$-b-$(X)_{2-15}$-Cys-$(X)_{2-4}$-Cys, where a and b can be either Cys or His, was developed. The form of this template was loosely based on the TFIIIA-like sequences, and the spacing of two to four residues between the outer sets of metal-binding ligands was chosen based on an analysis of structurally characterized metalloproteins. It should be noted that such a procedure cannot (and is not intended to) find all potential binding sites since many such sites involve amino acids that are widely separated in the amino acid sequence. The involvement of all metal-binding residues within a short stretch of sequence provides a definition of a metal-binding domain. Furthermore, the goal of the search was to identify *potential* metal-binding domains. Once identified, the actual role of metal ions in these systems (if any) must be investigated experimentally.

The template was used to search a computer-based protein sequence library. Five classes of proteins involved in nucleic-acid-binding or gene regulation were identified (Berg 1986). Of these, I shall discuss two in detail. The first group consists of the low-molecular-weight nucleic-acid-binding proteins from retroviruses. The proteins are encoded within the *gag* gene, expressed as part of a polyprotein, and freed by proteolysis. The final products have less than 100 amino acids and contain either one or two sequences of the form Cys-$(X)_2$-Cys-$(X)_4$-His-$(X)_4$-Cys. Although no direct evidence is yet available concerning the involvement of metal ions in these proteins, several observations support this hypothesis. First, some proteins contain one such sequence, whereas others contain two, with the spacing between the two varying from 5 to 22 residues. These facts suggest that each sequence forms an independent structural unit. Second, Gly residues occur frequently in positions 5 and, especially, 8. These are analogous to the Gly residues in the metal-binding sequences of rubredoxins (Cys-$[X]_2$-Cys-Gly) and azurins (Cys-$[X]_3$-Gly-His), where they allow type II β turns and, concomitantly, additional hydrogen bonds to form. Finally, these sequences are quite similar to the sequence Cys-$(X)_3$-His-$(X)_5$-Cys-$(X)_2$-Cys from the bacteriophage T4 helix-destabilizing (gene 32) protein, which was also identified as a potential metal-binding domain by the search procedure (Berg 1986). This protein has recently been shown to contain one zinc ion per molecule, and spectroscopic studies of the Co^{++}-substituted protein are consistent with this binding site (Giedroc et al. 1986).

The function of these proteins is not completely understood. In vitro studies of the isolated proteins indicated that they are nonspecific nucleic-acid-binding proteins with a preference for single-stranded substrates (Davis et al. 1976; Schulein et al. 1978; Smith and Bailey 1979). In vivo photo-cross-linking studies, however, indicated the presence of specific binding sites on the viral RNA (Darlix and Spahr 1982; Méric et al. 1984). More recent studies of mutants have strongly suggested that the proteins play a role in RNA packaging (Méric and Spahr 1986). A 2-amino-acid insertion within one of the two potential metal-binding domains of the Rous sarcoma virus protein p12 caused a marked decrease in the amount of dimeric viral RNA present in the virion, whereas more dramatic mutations led to little RNA incorporation whatsoever. These results suggest that the proteins may recognize a site (sequence or structure) on the viral RNA that is required for packaging. One possible explanation for the differences between the in vitro and in vivo studies is that the proteins have been isolated in the presence of EDTA and other potential chelating agents. Interestingly, treatment of free TFIIIA with low concentrations of EDTA causes loss of zinc with concomitant abolition of site-specific DNA-binding activity, although the protein still binds nucleic acids (particularly single stranded) nonspecifically (Hanas et al. 1983). TFIIIA is normally isolated as a complex with 5S RNA, which is substantially less sensitive to chelating agents.

The second class of proteins consists of one class of products of the adenovirus E1A genes (Berk 1986). These proteins contain a sequence of the form Cys-$(X)_2$-Cys-$(X)_{13}$-Cys-$(X)_2$-Cys (Berg 1986). Two overlapping mRNAs are transcribed from the E1A genes. These differ by a small internal sequence (93–138 nucleotides) that is removed by splicing from one of the messages but not the other. The messages remain in the same reading frame, however, so that the two protein products have identical amino- and carboxy-terminal ends and differ only by the presence of an internal sequence of 31–46 amino acids. Importantly, the potential metal-binding domain sequences occur in this region, which is unique to the larger protein product.

The E1A gene products have several activities (Berk 1986). They increase the efficiency of transcription of other early viral genes and some endogenous cellular genes, they play a role in cell transformation, and they repress the effect of certain enhancer elements. Studies of mutant genes have revealed that only the larger protein is active in transcriptional activation, whereas both proteins have repression and transformation ac-

tivities (Lillie et al. 1986). A variety of studies have indicated that the larger E1A protein is not itself a DNA-binding protein but instead interacts with cellular transcription factors to increase the efficiency of transcription of certain genes (Berk 1986). In particular, a TATA-box-binding transcription factor has been implicated in studies of E1A protein *trans*-activation of the adenovirus E1B promoter (Wu et al. 1987).

These results suggest that the potential metal-binding domain in the larger E1A protein may be involved in a protein-protein contact. As noted in the introduction, such a role for a metal-binding domain has a precedent in the regulatory subunit of aspartate transcarbamylase (Honzatko et al. 1982). This protein is a dimer of a polypeptide that has two structural domains called the allosteric effector domain and the zinc domain. The dimer is held together by interactions between β sheets of the allosteric effector domains, and the zinc domain mediates interactions between the regulatory dimer and the catalytic subunits. The zinc domain consists of residues 101–153 and contains the metal-binding sequence Cys-109–$(X)_4$–Cys-114–$(X)_{25}$–Cys-138–$(X)_2$–Cys-141. The interactions with the catalytic subunits involve residues close to the zinc-binding residues, namely, residues Asp-111, Lys-139, Tyr-140, and Glu-142. Thus, the bound zinc stabilizes a structure that is essential for this protein-protein interaction.

A Predicted Structure for the Zinc Finger Domain

A question arises from the observations of the TFIIIA-like and other potential metal-binding sequences: Can these sequences be used to predict more detailed three-dimensional structures for the metal-binding domains? Recent analysis of the TFIIIA-like zinc finger sequences suggests that the answer may be "yes" for certain cases (Berg 1987).

As noted above, the TFIIIA-like repeats have the consensus sequence ([Phe,Tyr]-X-Cys-$[X]_{2or4}$-Cys-$[X]_3$-Phe)-($[X]_4$)-(X-Leu-$[X]_2$-His-$[X]_3$-His)-($[X]_5$). The sequence has been divided into four parts that will be discussed below. The first part of the sequence contains the two cysteines and two conserved aromatic residues. Two structurally characterized proteins have sequences of the isolated sequences of the form Cys-$(X)_{2or4}$-Cys that are involved in binding single metal ions: rubredoxin (Fe) (Watenpaugh et al. 1979) and aspartate transcarbamylase regulatory subunit (Zn) (Honzatko et al. 1982). Each of these proteins contains two such sites. Analysis of the structures of these metal-binding sequences revealed that the structures are quite conserved from one protein to another. The regions with sequences shown below are of particular interest.

Rubredoxin:
Tyr-4-Thr-Cys-Thr-Val-Cys-Gly-Tyr-Ile-**Tyr**

Aspartate transcarbamylase:
Leu-136-Lys-Cys-Lys-Tyr-Cys-Glu-Lys-Glu-**Phe**

In addition to the two cysteine residues, these sequences also have hydrophobic amino acids in the positions corresponding to the conserved hydrophobic residues from the TFIIIA-like sequences. Importantly, structures of these two regions are extremely similar. The α carbon atoms for these stretches of 10 residues may be superimposed with a rms deviation of 0.72 Å. The structures are antiparallel β sheets with two hydrogen bonds between the first and last residues and an additional hydrogen bond between the NH group of the first cysteine and the carbonyl group of the second residue past the second cysteine. A major factor contributing to the similarity of the Cys-$(X)_{2\text{ or }4}$-Cys structures is the presence of NH to cysteine sulfur hydrogen bonds (Adman et al. 1975). Bonds are formed between the cysteine sulfur of residue n and the NH group of residue $n + 2$. These bonds orient the peptide units in a way that places large constraints on the total conformation.

The histidine residues occur in regions with the form X-Leu-$(X)_2$-His-$(X)_3$-His. Three structurally characterized proteins have two metal-binding histidine residues separated by three residues: thermolysin (Zn) (Matthews et al. 1972), hemerythrin (Fe) (Stenkamp et al. 1982), and hemocyanin (Cu) (Gaykema et al. 1985). In each case, the region is α helical with the metal bound to the ϵ-nitrogens of the two histidines. The presence of an α helix in the zinc finger structure is supported by two additional observations. First, the conserved Leu residue is positioned such that it would lie on the same face of an α helix as the histidines. Second, secondary structure prediction methods averaged over 39 of the zinc finger sequences have been used to predict that the region X-Leu-$(X)_2$-His-$(X)_2$ is helical (Brown and Argos 1986).

The two substructures (the β sheet containing the cysteine residues and the α helix containing the histidines) can be combined around a tetrahedral metal ion and connected via a bend consisting of the four residues that lie between them. Only one arrangement (corresponding to one absolute configuration around the tetrahedral zinc ion) is possible due to the length of the bend region. This structure places the conserved leucine residue in position to neatly pack against the two conserved aromatic residues to form a small hydrophobic core. Two schematic views of the predicted structure are shown in Figure 2. Pairs of these structures would be connected by the remaining linker sequence. This sequence generally contains five residues with the sequence Thr-Gly-Glu-Lys-Pro occurring frequently. This region is predicted to be in a relatively extended conformation.

Fairall, Rhodes, and Klug proposed two models for the interaction between TFIIIA and the 5S RNA gene based on nuclease digestion and chemical protection experiments (Fairall et al. 1986). In model I TFIIIA wraps around the DNA following the major groove, whereas in model II the protein lies on one face of the DNA with alternate fingers lying in two different planes at an angle to one another. The two models differ in their structural requirements for the individual metal-

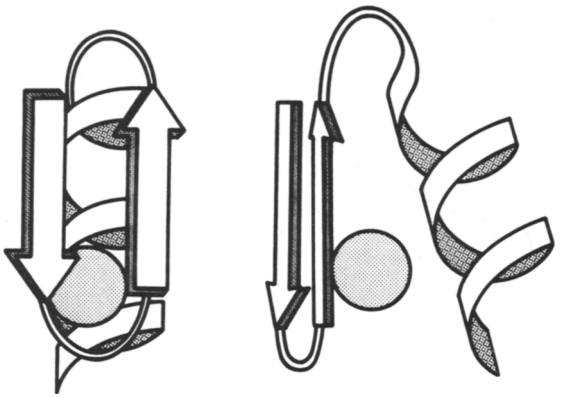

Figure 2. Two views of a proposed structure for the metal-binding domains from TFIIIA and related proteins (zinc finger domains). The structure has the form β-β-α with a zinc ion tetrahedrally coordinated by two cysteine and two histidine residues.

binding domains. In model II, the true repeat is two domains so that alternate domains contact the DNA through different faces, suggesting a symmetrical structure for the domain. In addition, this model suggests that the amino terminus and carboxyl terminus of the domain should exit from the same end of the finger. In contrast, each domain in model I interacts with the DNA in the same manner, and the amino and carboxyl termini should exit from opposite ends of the finger. The structure for the zinc finger domain developed clearly favors model I, with the protein lying largely in the major groove of the DNA (Berg 1987). A schematic view for the proposed protein-DNA complex is shown in Figure 3. It must be noted that the derived structure is based on the consensus sequence; variations from the consensus do occur and are probably crucial for the process of protein-nucleic acid recognition.

Structural and Potential Regulatory Roles of Metal-binding Domains

A final question remains about the occurrence of metal-binding domains in these nucleic-acid-binding and gene-regulatory proteins: Why are metal ions involved at all? A structural role is evident but the particular characteristics of metal ion stabilization of a small domain need further elaboration. Consider the hypothetical folding pathway shown in Figure 4 using a TFIIIA-like peptide as an example. In the absence of metal ion the peptide exists as a random coil. There are insufficient stabilizing interactions in the folded state (hydrogen bonds, hydrophobic interactions, van der Waals interactions) to overcome the factors stabilizing the unfolded state (conformational entropy, interactions with solvent). Upon reaction with one equivalent of metal ion in aqueous solution, the metal-binding residues (in this case two cysteinates and two histidines) displace the six water molecules from the metal ion to form a peptide-metal ion complex, but it is assumed that no additional "folding" occurs. A major contribution driving this reaction is the entropy of release of the six water molecules from the metal ion. The enthalpy of the metal-binding process may be positive in that six metal-ligand bonds are broken and only four are formed (albeit with more tightly binding ligands). Finally, the peptide folds to assume its final conformation. This process will be driven by the stabilizing interactions noted above with the conformational entropy of the "unfolded" peptide-metal ion complex being significantly reduced by the cross-links involving the bound metal ion. Thus, metal ion stabilization of small domains (which are not expected to be folded in the absence of bound metal ion) is founded on two entropic contributions: the entropy of water molecules released from the aquo complex of the "free" metal ion and the reduced conformational entropy of the peptide with a metal ion bound to four side chains.

Do metal ions play any other roles in these proteins? One feature that many of the zinc finger proteins such as TFIIIA and *Krüppel* appear to have in common is

Figure 3. A model proposed for the interaction of a protein consisting of TFIIIA-like metal-binding domains and DNA based on the predicted structure for the individual domains. The α helix from each domain lies in the major groove of the DNA and makes sequence-specific contacts, whereas the β sheet lies further away from the DNA helical axis and makes contacts with the sugar-phosphate backbone. The linker region also lies in the major groove. This model is related to model I proposed by Fairall, Rhodes, and Klug (Fairall et al. 1986).

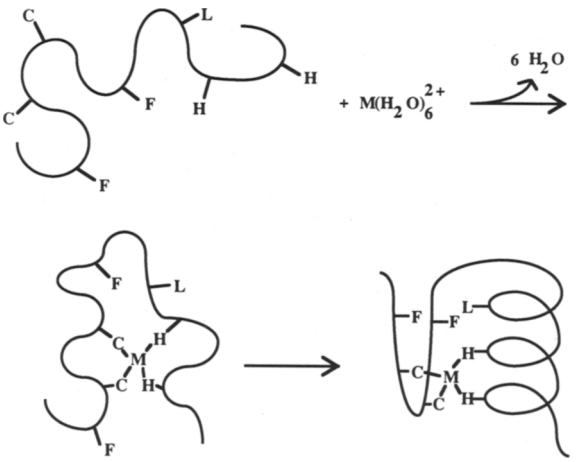

Figure 4. A hypothetical folding pathway for a TFIIIA-like peptide. The metal-free peptide exists as a random coil. This species binds a metal ion to form a relatively unfolded metal-peptide complex that then folds into its final conformation. This scheme is not intended to represent a real folding process but rather to illustrate some of the thermodynamic contributions to the stability of the final structure.

that they are involved in control of development. This suggests the possibility that zinc may be involved in developmental control in some manner. No direct information is yet available to support this hypothesis, but several additional observations allow the construction of one speculative scenario: (1) Free TFIIIA is quite sensitive to chelating agents (such as EDTA), losing its site-specific DNA-binding activity upon removal of zinc (Hanas et al. 1983); (2) TFIIIA bound to nucleic acids (at least 5S RNA) is less sensitive to such treatment (Hanas et al. 1983); and (3) metallothioneins, very cysteine-rich proteins involved in binding metal ions such as zinc and cadmium, are tightly regulated developmentally, with their expression being sensitive to the presence of steroid hormones and of certain metal ions. This tight regulation has suggested that these proteins have functions other than heavy-metal detoxification (Karin 1985). Early in development, metallothionein concentration is relatively low and free zinc concentration is expected to be relatively high. Thus, if free zinc finger proteins are produced, they bind zinc and are activated to bind their DNA target sites. Later, as metallothionein expression is elevated, the free zinc concentration should fall so that newly synthesized zinc finger proteins may not be activated, but existing zinc finger protein-nucleic acid complexes would be stable. Thus, the initial relatively high zinc concentration is reflected in the genome even after the free zinc concentration is reduced by increases in metallothionein concentration. Clearly, further chemical and biological studies are required to more completely elucidate the structural and functional roles of zinc in nucleic-acid-binding and gene-regulatory proteins.

ACKNOWLEDGMENTS

I thank the Camille and Henry Dreyfus Foundation and the National Institutes of Health for support of this work.

REFERENCES

Adman, E., K.D. Watenpaugh, and L.H. Jensen. 1975. NH—S hydrogen bonds in *Peptococcus aerogenes* ferredoxin, *Clostridium pasteurianum* rubredoxin, and *Chromatium* high potential iron protein. *Proc. Natl. Acad. Sci.* **72:** 4854.

Berg, J.M. 1986. Potential metal binding domains in nucleic acid binding proteins. *Science* **232:** 485.

———. 1987. A proposed structure for the zinc binding domains from transcription factor IIIA and related proteins. *Proc. Natl. Acad. Sci.* (in press).

Berk, A.J. 1986. Adenovirus promoters and E1A transactivation. *Annu. Rev. Genet.* **20:** 45.

Bieker, J.J. and R.G. Roeder. 1984. Physical properties and DNA-binding stoichiometry of a 5S gene-specific transcription factor. *J. Biol. Chem.* **259:** 6158.

Blakeley, R.L. and B. Zerner. 1984. Jack bean urease: The first nickel enzyme. *J. Mol. Catal.* **23:** 263.

Brown, R.S. and P. Argos. 1986. Fingers and helices. *Nature* **324:** 215.

Brown, R.S., C. Sander, and P. Argos. 1985. The primary structure of transcription factor TFIIIA has 12 consecutive repeats. *FEBS Lett.* **186:** 271.

Chowdhury, K., U. Deutsch, and P. Gruss. 1987. A multigene family encoding several "finger" structures is present and differentially active in mammalian genomes. *Cell* **48:** 771.

Coon, M.J. and R.E. White. 1980. Cytochrome P-450, a versatile catalyst in monooxygenation reactions. In *Metal ion activation of dioxygen* (ed. T.G. Spiro), p. 73. Wiley-Interscience, New York.

Darlix, J.-L.and P.-F. Spahr. 1982. Binding sites of viral protein P19 onto Rous sarcoma virus RNA and possible controls of viral functions. *J. Mol. Biol.* **160:** 147.

Davis, J., M. Scherer, W.P. Tsai, and C. Long. 1976. Low-molecular-weight, Rauscher leukemia virus protein with preferential binding for single-stranded RNA and DNA. *J. Virol.* **18:** 709.

Diakun, G.P., L. Fairall, and A. Klug. 1986. EXAFS study of the zinc-binding sites in the protein transcription factor IIIA. *Nature* **324:** 698.

Eickman, N.C., R.S. Himmelwright, and E.I. Solomon. 1979. Geometric and electronic structure of oxyhemocyanin: Spectral and chemical correlations to met apo., half met, met, and dimer active sites. *Proc. Natl. Acad. Sci.* **76:** 2094.

Fairall, L., D. Rhodes, and A. Klug. 1986. Mapping of the sites of protection on a 5S RNA gene by the *Xenopus* transcription factor IIIA: A model for the interaction. *J. Mol. Biol.* **192:** 577.

Frankel, A.D., J.M. Berg, and C.O. Pabo. 1987. Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc. Natl. Acad. Sci.* **84:** 4841.

Gaykema, W.P.J., A. Volbeda, and W.G.J. Hol. 1985. Structure determination of *Panulirus interruptus* haemocyanin at 3.2 Å resolution. Successful phase extension by sixfold density averaging. *J. Mol. Biol.* **187:** 255.

Giedroc, D.P., K.M. Keating, K.R. Williams, W.H. Konigsberg, and J.E. Coleman. 1986. Gene *32* protein, the single-stranded DNA binding protein from bacteriophage T4, is a zinc metalloprotein. *Proc. Natl. Acad. Sci.* **83:** 8452.

Ginsberg, A.M., B.O. King, and R.G. Roeder. 1984. *Xenopus* 5S gene transcription factor, TFIIIA: Characterization of a cDNA clone and measurement of RNA levels throughout development. *Cell* **39:** 479.

Hanas, J.S., D.J. Hazuda, D.F. Bogenhagen, F.Y.-H. Wu, and C.-W. Wu. 1983. *Xenopus* transcription factor A requires zinc for binding to the 5S RNA gene. *J. Biol. Chem.* **258:** 14120.

Hartshorne, T.A., H. Blumberg, and E.T. Young. 1986. Sequence homology of the yeast regulatory protein ADR1 with *Xenopus* transcription factor IIIA. *Nature* **320:** 283.

Honzatko, R.B., J.L. Crawford, H.L. Monaco, J.E. Ladner, B.F.P. Edwards, D.R. Evans, S.G. Warren, D.C. Wiley, R.C. Ladner, and W.N. Lipscomb. 1982. Crystal and molecular structures of native and CTP-liganded aspartate carbamyltransferase from *Escherichia coli*. *J. Biol. Chem.* **160:** 219.

Karin, M. 1985. Metallothioneins: Proteins in search of function. *Cell* **41:** 9.

Lillie, J.W., M. Green, and M.R. Green. 1986. An adenovirus E1a protein region required for transformation and transcriptional repression. *Cell* **46:** 1043.

Matthews, B.W., J.N. Jansonius, P.M. Colman, B.P. Schoenborn, and D. Dupourque. 1972. Three-dimensional structure of thermolysin. *Nature* **238:** 37.

Méric, C. and P.-F. Spahr. 1986. Rous sarcoma virus nucleic acid-binding protein p12 is necessary for viral 70S RNA dimer formation and packaging. *J. Virol.* **60:** 450.

Méric, C., J.-L. Darlix, and P.-F. Spahr. 1984. It is Rous sarcoma virus protein P12 and not P19 that binds to Rous sarcoma virus RNA. *J. Mol. Biol.* **173:** 531.

Miller, J., A.D. McLachlan, and A. Klug. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4:** 1609.

Nelbach, M.E., V.P. Pigiet, Jr., J.C. Gerhart, and H.K.

Schachman. 1972. A role for zinc in the quaternary structure of aspartate transcarbamylase from *Escherichia coli*. *Biochemistry* **11:** 315.

Pelham, H.R.B. and D.D. Brown. 1980. A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. *Proc. Natl. Acad. Sci.* **77:** 4170.

Poulos, T.L., B.C. Finzel, I.C. Gunsalus, G.C. Wagner, and J. Kraut. 1985. The 2.6-Å crystal structure of *Pseudomonas putida* cytochrome P-450. *J. Biol. Chem.* **260:** 16122.

Rosenberg, U.B., S. Schröder, A. Preiss, A. Kienlin, S. Côté, I. Riede, and H. Jäckle. 1986. Structural homology of the product of the *Drosophila Krüppel* gene with *Xenopus* transcription factor IIIA. *Nature* **319:** 336.

Schuh, R., W. Aicher, U. Gaul, S. Côté, A. Preiss, D. Maier, E. Seifert, U. Nauber, C. Schröder, R. Kemler, and H. Jäckle. 1986. A conserved family of nuclear proteins containing structural elements of the finger protein encoded by *Krüppel*, a *Drosophila* segmentation gene. *Cell* **47:** 1025.

Schulein, M., W.N. Burnette, and J.T. August. 1978. Stoichiometry and specificity of binding of Rauscher oncovirus 10,000-dalton (p10) structural protein to nucleic acids. *J. Virol.* **26:** 54.

Seamon, K.B. and R.H. Kretsinger. 1983. Calcium-modulated proteins. In *Calcium in biology* (ed. T.G. Spiro), p. 1. Wiley-Interscience, New York.

Smith, B.J. and J.M. Bailey. 1979. The binding of an avian myeloblastosis virus basic 12,000 dalton protein to nucleic acids. *Nucleic Acids Res.* **7:** 2055.

Smith, D.R., I.J. Jackson, and D.D. Brown. 1984. Domains of the positive transcription factor specific for the *Xenopus* 5S RNA gene. *Cell* **37:** 645.

Stenkamp, R.E., L.C. Sieker, and L.H. Jensen. 1982. Restrained least-squares refinement of *Themiste dyscritum* methydroxohemerythrin at 2.0 angstroms resolution. *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **38:** 784.

Tso, J.Y., D.J. Van Den Berg, and L.J. Korn. 1986. Structure of the gene for *Xenopus* transcription factor TFIIIA. *Nucleic Acids Res.* **14:** 2187.

Vallee, B.L., A. Galdes, D.S. Auld, and J.F. Riordan. 1983. Carboxypeptidase A. In *Zinc enzymes* (ed. T.G. Spiro), p. 25. Wiley-Interscience, New York.

Vincent, A., H.V. Colot, and M. Rosbash. 1985. Sequence and structure of the *serendipity* locus of *Drosophila melanogaster*: A densely transcribed region including a blastoderm-specific gene. *J. Mol. Biol.* **186:** 149.

Watenpaugh, K.D., L.C. Sieker, and L.H. Jensen. 1979. The structure of rubredoxin at 1.2 angstroms resolution. *J. Mol. Biol.* **131:** 509.

Wu, L., D.S.E. Rosser, M.C. Schmidt, and A.J. Berk. 1987. A TATA box implicated in E1A transcriptional activation of a simple adenovirus 2 promoter. *Nature* **326:** 512.