Zinc: Biological Functions and Coordination Motifs

BERT L. VALLEE*,† AND DAVID S. AULD†,‡

Center for Biochemical and Biophysical Sciences and Medicine and Department of Pathology, Harvard Medical School and Brigham and Women's Hospital, 250 Longwood Avenue, Boston, Massachusetts 02115

Received April 6, 1993

Early attitudes regarding the presence and role of metals in the structure and function of enzymes markedly influenced the evolution of an understanding of their mechanisms. On the basis of the presence of an iron-porphyrin in peroxidase and that enzyme's unusual resistance to hydrolysis by trypsin, Willstätter[1] asserted that enzymes essentially are metalloporphyrin-like,[2-4] a view which proved erroneous. Sumner demonstrated that urease is a protein.[5] However, the ensuing inference that enzymes do not contain metals essential to their function also proved flawed. In point of fact, urease itself was ultimately found to be a nickel enzyme.[6] Thus, the somewhat atypical characteristics of the particular enzymes studied by Willstätter and Sumner fueled a noisy scientific controversy, and the prolonged and bitter debate did not encourage third parties to help settle it through a broadening of the database which would have allowed valid generalizations then. Enyzmes are now generally accepted to be proteins though even that is now too narrow a definition considering the emerging body of knowledge on RNA and catalysis; while the majority of enzymes contain neither porphyrins nor metals, a significant number do.

Much time elapsed before many and highly purified enzymes became available and the roles of metals and/or metalloporphyrins in their catalytic mechanism could be studied. This delay was exceptionally long in the study of zinc enzymes. The metalloproteins which contain the intensely chromophoric metalloporphyrins and transition metals attracted attention to themselves—in sharp contrast with those containing colorless zinc atoms. The biological chemistry of zinc, its complexes with proteins, and the development of instrumental analytical methods critical to its identification were long delayed relative to that for, e.g., iron and copper coordination complexes. Similarly, the technical means to establish its presence in and selective binding to enzymes were developed only very slowly.

The remarkable progress in the development of procedures for the isolation and characterization of proteins, and the billionfold lowering of the limit of detection of zinc by atomic absorption spectroscopy during the last thirty years, has led to the identification of zinc enzymes in large numbers. Among the metalloenzymes they now, in fact, constitute the largest category.[7,8] Biology has evolved very varied motifs of zinc–protein coordination complexes which stabilize their structures and conformations to serve in both enzyme function and the expression of genetic messages. Motifs for each of these have been defined and are distinctive.

X-ray diffraction structure determination of both zinc complex ions and enzymes now provides absolute structural standards of reference for the identity and nature of zinc ligands in proteins. Seventeen zinc enzymes belonging to IUB classes I–IV provide unambiguous details and virtually conclusive data which identify both their zinc ligands and coordination geometries for the three types of zinc motifs that have now been recognized: *catalytic*, *structural*, and *cocatalytic*.[9-11] These structure determinations serve as standards of reference for other enzymes with similar functional characteristics in families of both identical and closely related enzymes from many different species. Both their structural identities—or the converse—and conformations of their active and structural enzymatic site zinc ligands can be recognized by comparing their amino acid sequences with those of the enzymes whose three-dimensional structures were determined.

The magnitudes of stability constants of bidentate zinc complex ions led to the attribution of sulfur donors

Bert L. Vallee received his B.S. degree from the University of Bern, Switzerland, in 1938 and his M.D. from New York University in 1943. He founded the Biophysics Research Laboratory at Harvard Medical School and Peter Bent Brigham Hospital in 1954. His current title is Edgar M. Bronfman Distinguished Senior Professor. More than 200 students have been trained in his laboratory, which has been designated the Center for Biochemical and Biophysical Sciences and Medicine since 1980. Prof. Vallee is a member of the American Academy of Arts and Sciences, the National Academy of Sciences and the American Chemical Society, among others, and has received numerous awards and honors including the Gibbs medal and the Rose award. He has studied the chemistry and biology of zinc throughout his career and has also devoted much of his research to the biochemical basis of alcohol abuse and the role of angiogenin in neovascularization and organogenesis.

David S. Auld received his B.S. and M.S. degrees in Chemistry from Lehigh University in 1960 and 1962, respectively, and his Ph.D. in Organic Chemistry from Cornell under Prof. T. C. Bruice in 1966. He then joined the Biophysics Research Laboratory at Harvard Medical School as a Research Fellow with B. L. Vallee. His present title is Associate Professor of Pathology (Biochemistry). He has been a member of the Board of Tutors in Biochemical Sciences at Harvard College since 1977, and has been on the Board of Editors of Biochemistry since 1980. His research interests include the kinetics and mechanism of zinc metalloenzymes, particularly carboxypeptidase A and the matrix metalloproteinases. He has developed specialized methods for kinetic analysis, particularly low temperature, stopped-flow analysis based on radiationless energy transfer and fluorescence. He also specializes in the synthesis of fluorescent substrates for metalloenzymes.

† Center for Biochemical and Biophysical Sciences and Medicine.
‡ Department of Pathology.

(1) He believed all enzymes are low molecular weight active compounds absorbed to inert carriers.
(2) Willstätter, R. *Aus meinem Leben*; Verlag Chemie: Weinheimbergstrade, 1949.
(3) Cori, C. F. *Trends Biochem. Sci.* **1981**, *6*, 194.
(4) Costa, A. B. *Chem. Br.* **1989**, *25*, 788.
(5) Sumner, J. B. *J. Biol. Chem.* **1926**, *69*, 435.
(6) Dixon, N. E.; Gazzola, C.; Blakeley, R. L.; Zerner, B. *J. Am. Chem. Soc.* **1975**, *97*, 4131.
(7) Vallee, B. L.; Galdes, A. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1984**, *56*, 283.
(8) Vallee, B. L.; Auld, D. S. In *Matrix Metalloproteinases and Inhibitors*; Birkedal-Hansen, H., Werb, Z., Welgus, H., Van Wart, H., Eds.; Gustav Fischer: Stuttgart, 1992; p 5.
(9) Vallee, B. L.; Auld, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 220.
(10) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990**, *29*, 5647.
(11) Vallee, B. L.; Auld, D. S. *Faraday Discuss.* **1992**, *93*, 47 and 114.

0001-4842/93/0126-0543$04.00/0

Table I. Zinc Ligands and Their Spacing for the Catalytic Zinc Sites[a]

enzyme	L_1	X	L_2	Y	L_3	Z	L_4
class I							
alcohol dehydrogenase	Cys_α	20	His_β	106	Cys_α(C)		H_2O
class III							
β-lactamase	His	1	His	121	His(C)		H_2O
adenosine deaminase	His	1	His	196	His(C)		H_2O
DD-carboxypeptidase	His_α	2	His	40	His_α(N)		H_2O
astacin	His_α	3	His_α	5	His(C)		H_2O
carboxypeptidase A	His_π	2	Glu_π	123	His_β(C)		H_2O
carboxypeptidase B	His	2	Glu_α	123	His_β(C)		H_2O
thermolysin	His_α	3	His_α	19	Glu_α(C)		H_2O
B. cereus neutral protease	His_α	3	His_α	19	Glu_α(C)		H_2O
P. aeruginosa elastase	His_α	3	His_α	19	Glu_α(C)		H_2O
phospholipase C	Glu_α	3	His_α	13	His_π(N)		H_2O
nuclease P1	Asp_α	3	His_α	12	His_β(N)		H_2O
alkaline phosphatase	Asp_β	3	His_β	80	His_β(C)		H_2O
class IV							
carbonic anhydrase I	His_β	1	His_β	22	His_β(C)		H_2O
carbonic anhydrase II	His_β	1	His_β	22	His_β(C)		H_2O

[a] X is the number of amino acids between L_1 and L_2; Y is the number of amino acids between L_3 and its nearest zinc ligand neighbor. L_3 is contributed by either the amino (N) or the carboxyl (C) portion of the protein with respect to L_1. The subscripts α, β, and π refer to the α-helix, β-sheet, or reverse turn structure which supplies the ligand. No structural information is given for the β-lactamase or adenosine deaminase.

as those which should be selective for zinc binding at enzyme active sites.[12] However, since then, examination of *multidentate* models has shown that both the imidazolyl nitrogen of histidine and the carboxyl and carbonyl oxygens of glutamic and aspartic acids are as effective in zinc complex ions and enzymes as the sulfhydryl group of cysteine.

Catalytic Zinc Sites

Ligands to Zinc. Three His, Glu, Asp, or Cys residues provide zinc ligands for all known enzyme *catalytic* zinc sites (Table I). H_2O is the fourth and universal ligand, but histidine is by far the most frequent amino acid among the catalytic site residues. Three histidines are typical of the lyases human carbonic anhydrases I and II;[13,14] the hydrolase *Bacillus cereus* β-lactamase;[15] the DD-carboxypeptidase of *Streptomyces albus* G;[16] adenosine deaminase;[17] and astacin.[18,19] Two histidines are characteristic of the hydrolases bovine carboxypeptidases A and B;[20,21] thermolysin, the neutral protease of *Bacillus thermoproteolyticus*;[22] *B. cereus* neutral protease;[23] *Pseudomonas aeruginosa* elastase;[24] and the catalytic-like zinc sites of the phosphate-metabolizing multi-zinc enzymes *B. cereus* phospholipase C,[25] *Penicillium citrinum* nuclear P1,[26] and *Escherichia coli* alkaline phosphatase.[27] The catalytic zinc site of alcohol dehydrogenase is the only one so far where there is only one histidine;[28] it is also unique among these enzyme active sites, where two cysteines participate. Glutamate is the oxygen donor in *six* and aspartate in *two* of these enzymes. Overall, zinc prefers the imidazolyl nitrogen by far in its catalytic sites; this finding is at variance with earlier suppositions.

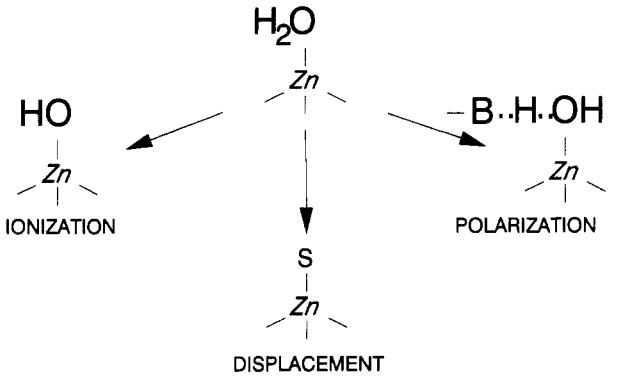

Figure 1. Schematic of the activation of H_2O in catalytic zinc sites. S, substrate; B, base.

Activation of H_2O. The fourth ligand (L_4) of active-site catalytic zinc atoms, H_2O, is not solely a critical feature but likely the very objective of this motif's design through its activation by ionization, polarization, or displacement (Figure 1). Ionization of the activated water or its polarization brought about by a base form of an active-site amino acid provides hydroxide ions at neutral pH, and its displacement results in Lewis acid catalysis on the part of the catalytic zinc atom. The structure of the active site implies that both the identity

(12) Gurd, F. R. N.; Wilcox, P. E. *Adv. Protein Chem.* **1956**, *11*, 311.
(13) Kannan, K. K.; Notstrand, B.; Fridborg, K.; Lövgren, S.; Orlsson, A.; Petef, M. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 51.
(14) Liljas, A.; Kannan, K. K.; Bergstén, P. C.; Waara, I.; Fridborg, K.; Strandberg, B.; Carlbom, V.; Järup, L.; Lövgren, S.; Petef, M. *Nature* **1972**, *235*, 131.
(15) Sutton, B. J.; Artymiuk, P. J.; Cordero-Borboa, A. E.; Little, C.; Phillips, D. C.; Waley, S. G. *Biochem. J.* **1987**, *248*, 181.
(16) Dideberg, O.; Charlier, P.; Dive, G.; Joris, B.; Frère, J. M.; Ghuysen, J. M. *Nature* **1982**, *299*, 469.
(17) Wilson, D. K.; Rudolph, F. B.; Quiocho, F. A. *Science* **1991**, *252*, 1278.
(18) Bode, W.; Gomis-Rüth, F. X.; Huber, R.; Zwilling, R.; Stöcker, W. *Nature* **1992**, *358*, 164.
(19) In three of these enzymes a fourth protein ligand with a relatively longer bonding distance has been noted. All of them are encountered in three histidyl-ligand catalytic zinc sites. In the Cd-substituted β-lactamase Cys-168 has been thought to be a weak ligand at 4.5 Å.[15] In astacin Tyr-149 is believed to be a ligand with a bond length of 2.6 Å,[18] and in the inhibitor-complexed form of adenosine deaminase Asp-295 ligates the zinc; the bond length is 2.5 Å.[17]
(20) Quiocho, F. A.; Lipscomb, W. N. *Adv. Protein Chem.* **1971**, *25*, 1.
(21) Schmid, M. F.; Herriott, J. R. *J. Mol. Biol.* **1976**, *103*, 175.
(22) Matthews, B. W.; Jansonius, J. N.; Colman, P. M.; Schoenborn, B. P.; Dupourque, D. *Nature, New Biol.* **1972**, *238*, 37.
(23) Pauptit, R. A.; Karlsson, R.; Picot, D.; Jenkins, J. A.; Niklaus-Reimer, A.-S.; Jansonius, J. N. *J. Mol. Biol.* **1988**, *168*, 525.
(24) Thayer, M. M.; Flaherty, K. M.; McKay, D. B. *J. Biol. Chem.* **1991**, *266*, 2864.
(25) Hough, E.; Hansen, L. K.; Birknes, B.; Jynge, K.; Hansen, S.; Horvik, A.; Little, C.; Dodson, E.; Derewenda, Z. *Nature* **1989**, *338*, 357.
(26) Volbeda, A.; Lahm, A.; Sakiyama, F.; Suck, D. *EMBO J.* **1991**, *10*, 1607.
(27) Kim, E. E.; Wyckoff, H. W. *J. Mol. Biol.* **1991**, *218*, 449.
(28) Brändén, C. I.; Jörnvall, M.; Eklund, M.; Furugren, B. In *Enzymes*; Boyer, P. D., Ed.; Academic: New York, 1975; Vol. 11, 3rd ed., p 103.

of the three protein ligands and their spacing underlie mechanistic pathways to activate water, determining details of the ensuing catalytic reactions accomplished in conjunction with other active-site residues. Thus, carbonic anhydrase converts carbon dioxide into bicarbonate. The three His ligands readily allow formation of zinc hydroxide, which can then add OH^- to CO_2 to form HCO_3^-. Hydrolytic enzymes such as the metalloproteases, thermolysin, and carboxypeptidase A need to both add hydroxide to the carbonyl group of the peptide amide bond and protonate the departing amine group. These enzymes employ two His and one Glu as ligands, thus reducing the charge on the zinc and retarding the ionization of the zinc-bound H_2O. However, there also is an adjacent ionized Glu residue which can H-bond to the metal-bound water, thus activating it by polarization. The resultant complex can be thought of as a protonated carboxyl group hydrogen-bonded to the zinc hydroxide. This adduct can thus be viewed as providing both hydroxide and a proton during the cleavage of the amide bond. In the case of alcohol dehydrogenase, the zinc function needed is that of a Lewis acid since the adjacent coenzyme, NADH, provides the hydride to the substrate aldehyde carbonyl group. The presence of two negative cysteinyl ligands should retard the formation of zinc hydroxide and allow the ready displacement of the zinc-bound water. The large polarizable nature of the sulfur ligands should still allow the zinc to act as a Lewis acid. Thus the amino acid ligands, their spacing in the protein sequence (see below), and the vicinal properties of the active center created by protein folding are critical for the mechanism of action of each particular enzyme.

Short and Long Spacers. The regularity of amino acid spacing between the ligands of catalytic zinc atoms is striking.[9,10,29] In 14 out of the 15 enzymes listed, "short" spacers, consisting of from 1 to 3 amino acids, separate the first two ligands, L_1 and L_2 (Table I). Apparently, when properly oriented, L_1 and L_2 can form a bidentate zinc complex. It is equally characteristic that a "long" spacer, 5–196 residues long, separates L_3 from either L_1 or L_2, generally in the C-terminal end of the protein. This long spacer arm could contribute to the induction of the active catalytic site, substrate-binding groups, and hydrogen bonds to form the active center. Remarkably, short and long spacers characteristic of the catalytic sites of mono-zinc enzymes recur in coactive zinc–metal bridges (see below and Table II).

The long spacers intervening between the residues that bind the *catalytic* zinc atoms imply a much more flexible coordination geometry of the active-site complex than that which is characteristic of *structural* zinc sites; there the interligand distances are relatively short, as might be consistent with the possibility that zinc stabilizes both overall protein structure and/or local conformation. Seemingly, the multiplicity of possible coordination numbers combined with the adaptability of zinc geometries can alternately impart local, structural, and functional rigidity or flexibility. On the one hand, this could poise the zinc atom for catalysis, creating an entatic state susceptible to change through substrate and/or product interactions during catalysis.[9,10] On the other, such adaptability could generate multiple conformational states which would both suitably align and organize those amino acid side chains and hydrogen bonds that participate in the catalytic process and substrate binding as well as accommodate an outer-sphere ligand whose coordination might activate water. Flexible zinc coordination could thus prove instrumental to the induction of conformationally "elastic" substrate-binding pockets which would have catalytic potential. Thus, differences in spacer lengths might be determinants of substrate specificity, the functions of water, and the details of catalytic mechanisms.[9,29]

Secondary Structure. The zinc binding sites frequently have an α-helical or β-structural region of the protein that supplies the zinc ligands of the catalytic zinc sites (Table I). The length of the short spacer conditions the support structure. Thus, in carbonic anhydrase the short spacer consists of but one amino acid. As a consequence, the ligands are provided by the same side of the β-sheet. In contrast, when the spacer consists of three amino acids, as in thermolysin, an α-helix support structure becomes feasible and juxtaposes the ligands in an orientation suitable to establish a tetrahedral-like coordination sphere.

These structural features of zinc metalloenzymes further call attention to the importance of protein folding and conformation, which are the basis of and maintain the structure and function of proteins. In zinc enzymes, the primary and secondary structure enveloping the short and long spacer arms might contain information on and give directions to the creation of zinc complexes with suitable coordination geometries and numbers. Zinc, and perhaps other metals, may in fact turn out to monitor and probe the folding process.

Such considerations will no doubt bear on the design and synthesis of enzyme model systems. Overall, the catalytic potentials of zinc enzymes seem closely related to the nature of the short and long amino acid spacers and the environment that they create for metal ligands. Their incorporation into synthetic designs should reflect the potentials of catalysis and specificity inherent in zinc enzymes.

Structural Zinc Sites

The *structural* zinc atom of equine alcohol dehydrogenase is bonded to four cysteines,[28] as is the case for zinc in the regulatory subunit of aspartate carbamoyltransferase.[30] In the linear sequence these cysteines are close to one another, separated by 2, 2, and 7 or 4, 22, and 2 amino acids, respectively.

For both ADH and ATCase the zinc is tetrahedrally coordinated to the four cysteines, preventing access of water or substrate to its coordination sphere. The role of the zinc in both instances cited apparently is to maintain the structure of the protein in its immediate vicinity. The effects of zinc in a structural site might be comparable to those of disulfide bonds and/or calcium. The predominance of sulfur ligands in these sites is in agreement with earlier views from both inorganic and geochemistry for predilection of zinc for sulfur ligands.

(29) Vallee, B. L.; Auld, D. S. *FEBS Lett.* **1989**, *257*, 138.
(30) Honzatko, R. B.; Crawford, J. L.; Monaco, H. L.; Ladner, J. E.; Edwards, B. F. P.; Evans, D. R.; Warren, S. G.; Wiley, D. C.; Ladner, R. C.; Lipscomb, W. N. *J. Mol. Biol.* **1982**, *160*, 219.

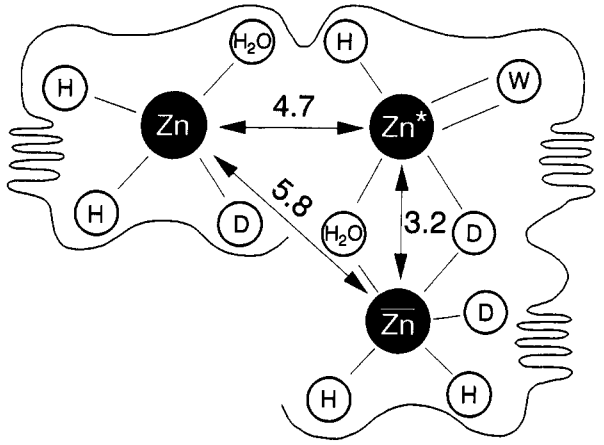

Figure 2. Schematic of the cocatalytic zinc site of nuclease P1.

Cocatalytic Zinc Sites

We have named a third type of zinc enzyme binding site *cocatalytic* or *coactive*.[11,31,32] Its characteristics emerge from inspection of the X-ray structure of the multi-zinc enzymes which have been determined thus far. *Cocatalytic* zinc binding sites occur in enzymes that contain two or more zinc atoms in close proximity to one another. Together they operate as a catalytic unit. X-ray studies reveal that in such zinc sites a single amino acid residue, either Asp or Glu, simultaneously binds to two zinc atoms or a zinc and a magnesium atom to form a bridge between two of the participating metal atoms. Alteration in the bond lengths or dissociation of either or both of these atoms from their respective metals during substrate interaction with the enzyme or formation of intermediates could profoundly affect catalysis due to the change in charge imposed upon the metals. In this manner these residues become pivotal components of the resultant functional template which consists of a network of closely spaced metal atoms and their ligands.

In the past literature, the zinc—or other metal—atoms that turn out to form such bridges have been described functionally as "modulating" or "regulatory".[8,33] Thus far, in addition to the active-site zinc atoms, these have been identified as zinc and magnesium atoms in alkaline phosphatase,[27] two zinc atoms in phospholipase C[25] and *P. citrinum* nuclease P1,[26] and zinc and/or magnesium in bovine lens leucine aminopeptidase.[34,35] In all of these enzymes one zinc atom is catalytic (**Zn**), while the remaining pair of metal atoms, **Zn*** and $\overline{\mathbf{Zn}}$ or Mg, is linked by a bridging amino acid, usually Asp, so that all three metal atoms are in close physical proximity to one another, as exemplified for nuclease P1 in Figure 2.

There are yet other features that set the coordination of their cocatalytic (**Zn*** and $\overline{\mathbf{Zn}}$/Mg) apart from that of **catalytic Zn** sites. While His and Cys predominate as ligands of catalytic and structural zinc atoms, Asp predominates in cocatalytic zinc sites where the frequency is Asp > His ≫ Glu (Table II). In the multi-zinc enzymes here considered, Ser, Thr, Lys, and Trp turn out to be ligands to **Zn*** and $\overline{\mathbf{Zn}}$/Mg sites. The ionizable groups of such ligands have high pK_a values. Hence, at neutrality their affinity for binding zinc is relatively lower than that of the His imidazole nitrogen, Glu and Asp carboxylate oxygen, and Cys sulfur donor groups of catalytic and structural sites. Moreover, amide carbonyl ligands should also coordinate zinc weakly. Thus, when substrate enters the coordination sphere, ligand displacement and/or increased coordination would be plausible with a cocatalytic mechanism.

Individual zinc—and other metal—atoms and their ligands may play alternate roles in different steps of catalysis. Thus, for the phosphate ester-hydrolyzing enzymes interaction of the substrate with their cocatalytic zinc site may involve initial contact with all three metals, as is observed in their final phosphate product complex.[25-27] One of the zinc atoms may then polarize a nucleophile, e.g., either water or Ser-102 in the case of alkaline phosphatase, to attack an amide or a phosphoester bond. A second zinc atom could act transiently as the receptor for the alcohol or amine leaving group. In subsequent steps the role of these zinc atoms may be reversed. The first zinc atom may polarize a phospho–oxygen or carbonyl bond of an acyl intermediate for the attack of water promoted by the second zinc atom. The bridging Asp and H_2O ligands could have critical roles in this process (Figure 2). Thus, their dissociation from either or both metal atoms during catalysis could change the charge on the metal, promoting the metal's action as a Lewis acid or allowing interaction with an electronegative atom of the substrate. Alternatively, the bridging ligand may participate transiently in the reaction as a nucleophile or general acid/base catalyst. In this manner the metal atoms and their associated ligands would play specific roles in each step of the reaction that works in concert to bring about catalysis.

It is apparent that the formation of cocatalytic sites in zinc metalloenzymes likely reflects sophistication of catalytic mechanisms brought about by innovations in chelate chemistry, structural design of coordinating sites, and modification/modulation of mechanistic arrangements. It would seem to represent an unusual biological utilization of zinc chemistry directed toward different ends of biological needs. The discernment of the structure/function relationships in *cocatalytic* sites likely heralds yet other motifs to be discerned.

Metalloexopeptidases (Unknown X-ray Structure)

Mono-zinc aminopeptidases (EC 3.4.11.2) have been identified in and isolated from a wide range of tissues and bacteria, but their X-ray or NMR structures have not yet been reported. Aminopeptidase M and human liver, porcine kidney, and two bacterial aminopeptidases from *Bacillus licheniformis* and *Bacillus subtilis* each contain one zinc atom.[10] In contrast, the aminopeptidase from bovine lens[36] features two zinc atoms (see

(31) Vallee, B. L.; Auld, D. S. *Biochemistry* **1993**, *32*, 6493.
(32) Vallee, B. L.; Auld, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2715.
(33) Vallee, B. L. In *Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhäuser: Boston, 1983; p 1.
(34) Burley, S. K.; David, P. R.; Sweet, R. M.; Taylor, A.; Lipscomb, W. N. *J. Mol. Biol.* **1992**, *224*, 113.
(35) Zinc and magnesium have also been implicated in the 3′–5′ exonuclease activity of *E. coli* DNA polymerase I: Beese, L. S.; Steitz, T. A. *EMBO J.* **1991**, *10*, 25–33. In that case zinc binds to Asp-355, Glu-357, and Asp-501. Asp-355 bridges magnesium to this site. The Cu/Zn superoxide dismutases from bovine, human, spinach, and yeast all contain a bridging metal ligand, His-61.[31] Structurally, it resembles Asp and Glu in binding two metal atoms, but since Cu clearly is the catalytic metal in superoxide dismutase, this enzyme is not discussed here.
(36) Thompson, G. A.; Carpenter, F. H. *J. Biol. Chem.* **1976**, *251*, 53.

Table II. Zinc Ligands and Their Spacing in Amino Acid Bridging Zinc–Metal Sites[a]

enzyme	metal	L_1	X	L_2	Y	L_3	Z	L_4
E. coli alkaline phosphatase	**Zn***	Asp_β	0	His_β	317	***Asp***$_\beta$(N)	50	Ser(N)
	Mg	***Asp***$_\beta$	103	Thr_β	166	Glu_β(C)		H_2O[b]
B. cereus phospholipase C	**Zn***	Trp[c]	12	His_α	107	***Asp***$_\alpha$(C)		H_2O
	Zn	His_α	3	***Asp***$_\alpha$	48	His(N)	13	Asp_α(N)[d]
P. citrinum nuclease P1	**Zn***	Trp[c]	4	His_α	113	***Asp***$_\alpha$(C)		H_2O
	Zn	His	3	***Asp***$_\alpha$	55	His_α(N)	14	Asp_α(N)[d]
bovine lens aminopeptidase	Zn	Asp[e]	1	***Glu***$_\alpha$	76	Asp_β(N)		H_2O[f]
	Zn*	Lys_β	22	Asp	60	***Glu***$_\alpha$(C)		

[a] The amino acid spacer between ligands L_3 and L_4 is Z. See Table I caption for other definitions. Amino acid residues which bridge the two metal sites are shown in italic boldface. [b] Three water molecules complete this hexacoordinate site. [c] Ligands are the α-amino and carbonyl group of the N-terminal Trp. [d] A bridging water molecule or hydroxide ion from **Zn*** completes this pentacoordinate site. [e] The Asp is coordinated to **catalytic Zn** through both the carboxyl group and amide carbonyl. [f] A water molecule is observed at 3.2 Å.

above). These two categories of aminopeptidase either contain one or two zinc atoms or contain one zinc plus a magnesium atom in lieu of the second zinc. Both feature the one zinc atom essential for catalytic activity.

In the mono-zinc human intestinal, rat kidney, and *E. coli* aminopeptidases, a domain of ~300 amino acids contains two histidines and one glutamic acid, similar to the zinc binding site of thermolysin. The three amino acids of the short spacer in thermolysin are identical to those of the mono-zinc aminopeptidases. The long spacer in these aminopeptidases contains 18 amino acids, corresponding to 19 in thermolysin.[9] This structural analysis of the zinc binding sites in those zinc enzymes predicted the presence of zinc and a hitherto unrecognized aminopeptidase activity in leukotriene A_4 hydrolase (EC 3.3.2.6). It contains an amino acid segment that is homologous to the zinc binding domain of intestinal aminopeptidase. On this basis, the leukotriene A_4 hydrolase was then shown to contain 1 mol of zinc/mol of protein, to exhibit aminopeptidase activity, and to be inhibited by bestatin and captopril, specific peptidase inhibitors.[37] Moreover, mutagenic replacements of leukotriene A_4 hydrolase zinc ligands completely abolish its activity,[11,37,38] the first proof of the identity of a hitherto unknown zinc enzyme binding site by mutagenesis.

Metalloendopeptidases (Unknown X-ray Structure)

The matrix metalloproteinases (MMPs) have long been proposed to be involved in both normal and pathological invasive processes of the extracellular matrix (ECM). Two MMPs, pump-1[39] and human stromelysin-1,[40] have been shown to contain zinc, 1.0 and 1.9 mol, respectively.

The zinc binding site "signature sequence" of astacin, HExxHxxGxxH, is characteristic of this family of metalloproteinases.[8,41-43] The X-ray structure of astacin[18] confirmed the prediction of a new type of a zinc protease binding site in which all three histidines are ligands to the catalytic zinc.[43] On the basis of this zinc binding site the extended astacin family of metalloproteinases likely includes (1) all matrix metalloproteinases, (2) the hemorrhagic toxins, and (3) certain bacterial zinc enzymes involved in inflammatory processes. The vast majority of these proteases degrade the ECM and share with astacin a preference for substrates containing Pro residues one or two amino acids removed from the cleavage site.[42]

The MMPs range in size from the 92-kDa collagenase to the 28-kDa pump-1. All, including HE6 and MME, have a domain of approximately 270 amino acids which is homologous to pump-1, the prototype of the MMP family.[43] Pump-1 has the features essential for both the latent form of the enzyme, a propeptide that contains a single Cys believed to be involved in zymogen activation,[10,44] and the catalytic domain containing the catalytic zinc binding site. The larger forms contain additional domains which may be important for their binding to the matrix components.

"Velcro" Mechanism of Matrix Prometalloproteinases Activation. The MMPs are synthesized as inactive precursors and then converted to the active form. The mechanism of activation of the prometalloproteases does not conform with either a "one-by-one" or a "zipper" mechanism,[45] but rather by the "Velcro" mechanism.[10]

In a highly conserved region, PRCGVPDV, the propeptides of the procollagenases and progelatinases contain one single cysteine residue.[10] This cysteine forms a mercaptide with the single zinc atom of the enzyme, making it inactive. Apparently, the dissociation and/or displacement of that cysteine from the zinc atom generates enzymatic activity.[44] While details of this process still require delineation, the removal of the SH group from the zinc is clearly the principal chemical event to induce activity. Thus, in this instance the displacement of cysteine converts a tetradentate (and presumably structural-like site) into an enzymatically functional zinc atom, providing another example of the adaptability of zinc chemistry to biological needs.

Zinc binding sites at the interface of protein–protein interactions having short and long spacer characteristics and four amino acid ligands have also been recently observed in the serine protease tonin[46] and in the binding of human growth hormone to human prolactin

(37) Haeggström, J. Z.; Wetterholm, A.; Shapiro, R.; Vallee, B. L.; Samuelsson, B. *Biochem. Biophys. Res. Commun.* **1990**, *173*, 431.
(38) Medina, J. F.; Wetterholm, A.; Rådmark, O.; Shapiro, R.; Haeggström, J. Z.; Vallee, B. L.; Samuelsson, B. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7620.
(39) Crabbe, T.; Willenbrock, F.; Eaton, D.; Hynds, P.; Carne, A. F.; Murphy, G.; Docherty, A. J. P. *Biochemistry* **1992**, *31*, 8500.
(40) Salowe, S. P.; Marcy, A. I.; Cuca, G. C.; Smith, C. K.; Kopka, I. E.; Hagman, W. K.; Hermes, J. D. *Biochemistry* **1992**, *31*, 4537.
(41) Stöcker, W.; Wolz, R. L.; Zwilling, R.; Strydom, D. A.; Auld, D. S. *Biochemistry* **1988**, 427, 5026.
(42) Stöcker, W.; Ng, M.; Auld, D. S. *Biochemistry* **1990**, *29*, 10418.
(43) Auld, D. S. *Faraday Discuss.* **1992**, *93*, 117.
(44) Springman, E. B.; Angleton, E. L.; Birkedal-Hansen, H.; Van Wart, H. W. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 364.
(45) Linderström-Lang, K. *Stanford Univ. Publ., Univ. Ser., Med. Sci.* **1952**, *6*, 1.
(46) Fujinaga, M.; James, M. N. G. *J. Mol. Biol.* **1987**, *195*, 373.

receptors.[47] In both of these instances one protein molecule provides a short and long spacer while the second molecule supplies the fourth ligand, a situation reminiscent of the chemical basis of the Velcro mechanism.[10]

Metallothioneins

The metallothioneins are another class of zinc proteins that have become increasingly important. Their structure is known but their function is not, though they have been suggested to play an important role in a multiplicity of metabolic and toxicological processes.[48] Their molecular weight is low (~6700) but their metal content is high, usually 7 (or 6) mol/mol including zinc, cadmium, and/or copper. There are 62 (or 61) amino acids, among them 20 cysteines, although neither cystine, heterocyclic, nor aromatic amino acids are present. The amino acid sequences of metallothioneins are unique, the distribution of cysteinyl residues is characteristic, and the spectroscopic features are typical of metal thiolates (mercaptides) in general and clusters in particular.

Coordination Chemistry. In all classes of metallothioneins the metal complexes are organized in metal–thiolate clusters, a structural feature not hitherto described for naturally occurring inorganic zinc and/or cadmium complexes. This mode of coordination, proven by NMR[49] and X-ray diffraction analysis,[50] is in the form of tetrahedral tetrathiolate bonds, with some of the thiolate ligands sharing a bound divalent metal ion atom.[51] In mammalian metallothioneins eight Cys serve as doubly coordinated bridging and twelve as singly coordinated terminal thiolate ligands. The overall measurement of about 3 SH ligands per metal atom[51] is in accord with this.

Thermodynamic and Kinetic Stability of the Clusters. Cluster B preferentially loses cadmium and/or zinc, which tend to accumulate in cluster A and B, respectively, the latter being the site of lower affinity.[52] Indications are that the stability of the two metal–thiolate clusters differs significantly, resulting in random distribution of the two clusters, presumably accounting for the resultant cooperativity.[53] It is conceivable that these cluster features serve to regulate the metabolic roles of these metals, as indicated by fluctuations in the clusters and consequent intramolecular metal exchange.[54] Spontaneous mutual exchange among binding sites occurs within minutes and seconds in cluster A and B, respectively.[55] This metal-free form of the protein, thionein, can remove zinc from the gene activator protein TFIIIA.[56] This may bear on potential interprotein metal exchange and the physiological function of metallothionein, particularly in regard to the GAL4 DNA binding proteins, as will be discussed.

Zinc Fingers, Twists, and Clusters

In the 1970s evidence emerged that zinc is crucial to DNA and RNA synthesis and cell division,[11] followed by the discovery of Wu that *Xenopus* transcription factor IIIA (TFIIIA), which activates the transcription of the 5S RNA gene, contains from 2 to 3 mol of zinc/mol of protein,[57] thus focusing interest on the role of zinc in the transcription process. It was proposed that the two Cys and two His residues per 30 amino acid unit of TFIIIA form a tetrahedral coordination complex with each of nine zinc atoms to generate peptide domains—zinc fingers[58]—that interact with DNA. However, the first three zinc fingers bind to the internal control region of the 5S RNA as tightly as the full-length TFIIIA,[59] correlating better with the zinc content of Wu and co-workers than with suppositions based on the sequence alone. NMR data on single or double "fingers" produced by organic synthesis, biochemical engineering, and mutagenesis provided the first evidence for the structure of zinc fingers.

The TFIIIA sequence and the features of its zinc complexes have become models for zinc proteins that bind to DNA. Numerous sequences have been thought to reflect both structural and functional similarities, leading to revision of past ideas and predictions of new mechanisms for DNA–protein interactions. The strictest classification would confine the term zinc finger to those proteins in which (1) there is one or multiple repeats of about 30 amino acids each, (2) both two Cys and two His residues and their spacing are conserved, (3) two aromatic residues and leucine are conserved in the TFIIIA repeat units,[60] and (4) the three-dimensional structure of the sequence is presumed to resemble a finger. However, the designation "zinc finger" has served to describe virtually any relatively short protein sequence that contains four or more Cys and/or His residues which is believed to interact with a nucleic acid binding domain.[61]

(A) Synthetic "Zinc Finger" Peptides. Small peptide domains have been synthesized to serve as zinc finger models of those encountered in *Xenopus* TFIIIA and *Xfin* proteins,[62,63] the yeast transcriptional activator SW15,[64] the yeast alcohol dehydrogenase regulatory gene ADR1,[65] and mouse Krüppel-like gene mKr2.[66]

(47) Cunningham, B. C.; Bass, S.; Fuh, G.; Wells, J. A. *Science* **1990**, *250*, 1709.
(48) Vallee, B. L. *Experientia Suppl.* **1987**, *34*, 19.
(49) Messerle, B. A.; Schäffer, A.; Vašák, M.; Kägi, J. H. R.; Wüthrich, K. *J. Mol. Biol.* **1990**, *214*, 765.
(50) Robbins, A. H.; McRee, D. A.; Williamson, M.; Collett, S. A.; Xuong, . H.; Furey, W. F.; Wang, B. C.; Stout, C. D. *J. Mol. Biol.* **1991**, *221*, 1269.
(51) Kägi, J. H. R.; Vallee, B. L. *J. Biol. Chem.* **1961**, *236*, 2435.
(52) *Metallothionein II, Proceedings of the Second International Meeting, Zurich, Switzerland*; Kägi, J. H. R., Kojima, Y., Eds.; Birkhäuser: Basel, 1987.
(53) Avdeet, A.; Zelazowski, A. J.; Garvey, J. S. *Inorg. Chem.* **1985**, *24*, 1928.
(54) Vašák, M.; Hawkes, G. E.; Nicholson, J. K.; Sadler, P. J. *Biochemistry* **1985**, *24*, 740.
(55) Otvos, J.; Chen, S.-M.; Liu, X. In *Metal Ion Homeostasis: Molecular Biology and Chemistry*; Alan R. Liss: New York, 1989; p 197.
(56) Zeng, J.; Vallee, B. L.; Kägi, J. H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9984.
(57) Hanas, J. S.; Hazuda, D. J.; Bogenhagen, D. F.; Wu, F. Y.-H.; Wu, C.-W. *J. Biol. Chem.* **1983**, *258*, 14120.
(58) Klug, A.; Rhodes, D. *Trends Biochem. Sci.* **1987**, *12*, 464.
(59) Liao, X.; Clemens, K. R.; Tennant, L.; Wright, P. *J. Mol. Biol.* **1992**, *223*, 857.
(60) Frankel, A. D.; Bredt, D. S.; Pabo, C. O. *Science* **1988**, *240*, 70.
(61) Berg, J. M. *Curr. Opin. Struct. Biol.* **1993**, *3*, 11.
(62) Frankel, A. D.; Berg, J. M.; Pabo, C. O. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4841.
(63) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635.
(64) Neuhaus, D.; Nakaseko, Y.; Nagai, K.; Klug, A. *FEBS Lett.* **1990**, *262*, 179.
(65) Párraga, G.; Horvath, S. J.; Eisen, A.; Taylor, W. E.; Hood, L.; Young, E. T.; Klevit, R. E. *Science* **1988**, *241*, 1489.
(66) Carr, M. D.; Pastore, A.; Gausepohl, H.; Frank, R.; Roesch, P. *Eur. J. Biochem.* **1990**, *188*, 455.

Based on known structures of zinc binding sites in proteins and theoretical considerations, models for the structure of the peptide domains have been proposed.[67,68] These predict tetrahedral coordination to the two cysteines and two histidines and an α-helix through the two histidine ligands. They differ in suggesting either a β-sheet configuration for the cysteine region[67] or a two-stranded β-hairpin turn for the cysteine ligands.[68]

EXAFS and visible absorption spectrometry of cobalt complexes of single-finger domains support the cysteine/histidine binding of zinc in a tetrahedral arrangement for ADR1,[65] a TFIIIA domain,[62] and TFIIIA itself.[69] Addition of Zn^{2+} to the apo-ADR1,[70] TFIIIA,[60] and *Xfin*[71] peptides results in a significant increase in α-helical content and shifts in the ^1H NMR resonances for His, suggesting the zinc introduces novel structural features. Reducing the size of that loop of ADR1 from 12 to 11 amino acids by deleting Asn-138 prevents tetrahedral coordination of cobalt and results in a peptide which does not fold into a stable domain structure.[70] However, this peptide contains a third cysteine residue only one removed from the putative zinc binding cysteine. The effect of the deletion on the oxidation state of the cysteines in the deletion peptide was not reported.

In several model systems varying lengths and regions of the α-helix have been examined by NMR. In the 25 amino acid *Xfin* peptide domain a particularly long stretch of an 11 amino acid α-helix (E12–K24) incorporates and extends through both histidines.[63] A changing α- to 3_{10}-helix of similar length (N56–Q67) occurs in the second finger of the 35 amino acid peptide domain of SW15.[64] A much shorter amino acid α-helix (L146–K153) brackets only the first histidine ligand for the 30 amino acid yeast ADR1-2 peptide.[65]

In the synthetic 30 amino acid mKr2 zinc-free peptide, a shortened α- or 3_{10}-helix (S14–I20) terminates one amino acid short of the first histidine ligand.[66] In all four peptides examined thus far, an α-helix was observed, but its length varies from 5 to 11 amino acids and may include both, one, or no histidine ligands.

The putative cysteine ligand region has been described variously. The cysteines are part of two β-strands arranged in a hairpin structure of the *Xfin* peptide,[63] an irregular antiparallel β-sheet (Y42–F53) of the SW15 peptide,[64] some form of a turn for ADR1,[65] or no β-turn or -sheet for the mKr2 peptide.[66]

These model studies therefore indicate the zinc is critical for proper peptide folding and that some form of secondary structure exists for the peptide segments encompassing the Cys and His ligands similar to what has been observed in zinc binding sites of enzymes. Previous efforts to identify and define the zinc binding sites of zinc enzymes were based on the three-dimensional structures of a given member of a family determined by X-ray methods. The results were then compared with sequences of members of the families to which they belong but on which no structure analysis had been performed.[9,10] This scenario could not be followed for the DNA-binding zinc proteins until structural data were obtained.

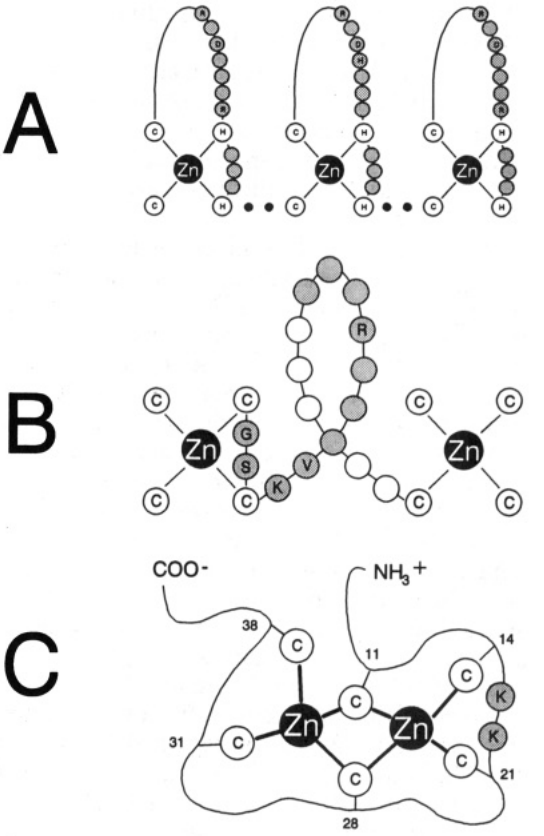

Figure 3. Schematic of (A) zinc fingers of Zif268, (B) zinc twist of glucocorticoid receptor, and (C) zinc cluster of GAL4. The amino acids involved in DNA recognition are labeled in shaded circles.

(B) Zinc Fingers. A definitive understanding of how these zinc finger domains serve in site-specific recognition of a particular DNA-binding protein has come from the X-ray diffraction analysis of the DNA-binding domain derived from the mouse immediate early protein Zif 268 (also known as Krox 24, NGFI-A, and Egr1). The three-zinc finger peptide was crystallized with a consensus DNA-binding site, and the structure was solved at 2.1 Å.[72]

Each zinc finger domain consists of an antiparallel β-sheet containing two Cys and an α-helix containing two His held together by coordination of the Cys and His residues to the central zinc ion and by a set of hydrophobic residues. The interatomic distance between Zn(1) and Zn(2) is 26.6 Å, and that between Zn(2) and Zn(3) is 27.4 Å.[11] Importantly, in each zinc finger one zinc atom is coordinated to two His and two Cys as its base.

Each of the three zinc fingers uses Arg, His, and Asp amino acid residues from the N-terminal portion of its α-helix to make contact with guanine bases in the major groove. These residues derive from the central peptide loop between the second Cys and first His of each "finger" and include the amino acid immediately preceding the α-helix and the second, third, and sixth ones of the α-helix (Figure 3A).

The results of these studies demonstrate how zinc makes a direct and unsuspected contribution to the

(67) Berg, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 99.
(68) Gibson, T. J.; Postma, J. P. M. Brown; R. S.; Argos, P. *Protein Eng.* **1988**, *2*, 209.
(69) Diakun, G. P.; Fairall, L.; Klug, A. *Nature* **1986**, *324*, 698.
(70) Párraga, G.; Horvath, S. J.; Young, E. T.; Klevit, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 137.
(71) Lee, M. S.; Gottesfield, J. M.; Wright, P. E. *FEBS Lett.* **1991**, *279*, 289.
(72) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809.

overall binding energy. The tetrahedral geometry around the zinc ion orients the finger for site-specific interaction. Zinc also promotes a specific interaction of the first ligating His to the DNA backbone; the N^{ϵ} of His coordinates the zinc while the N^{δ} hydrogen bonds to the phosphodiester oxygens. As multiple zinc finger proteins are studied, additional permutations may become apparent.

(C) Zinc Twists. The glucocorticoid (GR) and estrogen (ER) receptors belong to the nuclear receptor class of proteins. Upon binding to a hormone, the receptor translocates from the cytoplasm to the nucleus, where it binds to a specific DNA sequence (glucocorticoid or estrogen responsive elements, GRE or ERE) and thereby modulates transcription. The ^{1}H and ^{113}-Cd NMR solution structures of the 71 amino acid segment of the DNA-binding domain of GR[73-75] and the ^{1}H-NMR structure of a corresponding 84 amino acid fragment of ER[76] show two zinc atoms each coordinated to four cysteines in these particular receptors separated by ~13 and 12 Å, respectively.

Remarkably, these zinc binding sites are not independent zinc finger-like substructures, but rather fold to form a single structural domain.[73,76] The DNA-binding recognition site derives from the α-helix that is part of the linking peptide region between the two zinc sites, accounting for our choice of calling this a zinc twist (Figure 3B). It clearly need not necessarily be an α-helix in all proteins, of course. A second α-helix which is anchored by the C-terminal coordinating cysteines of the second zinc site perpendicularly crosses the first helix near its midpoint, stabilizing it by hydrophobic interaction between the two helices. The globular peptide domain of the resultant zinc twists has a DNA recognition site between the two zinc atoms, each of which is coordinated to four Cys.

X-ray crystallographic analysis of the DNA-binding domain of GR complexed to either a GRE consensus sequence containing a symmetric but abnormal spacing of four nucleotides, GRE_{4s}, or a natural three-nucleotide space, GRE_{3s},[77] is in general agreement with the molecular models for dimer interaction with DNA as proposed on the basis of NMR studies of uncomplexed monomeric DNA-binding domains of GR and ER.[73,76] The two zinc atoms are coordinated tetrahedrally to the expected eight cysteines with the central α-helix of the zinc twist, forming a complex with the major groove of the DNA. These studies further stress the importance of Arg, Val, and Lys residues in this interaction (Figure 3B). Arg-466 makes two hydrogen bonds to a G(4) base in a manner similar to that seen for Zif 268. A methyl group of Val-462 makes a van der Waals contact with the 5-methyl group of the T(5) base while Lys-461 directly hydrogen bonds to N(7) of G(6) and indirectly provides hydrogen bonds through a water molecule to O(6) of G(7) and O(4) of T(6). The X-ray diffraction studies further reveal how the second zinc binding site influences the dimerization process. Residues from the five amino acid loop between the first two cysteines make critical dimer interface contacts with residues from the nine amino acid loop between the second and third cysteines.[77]

The function of zinc in the nuclear receptor family is clearly twofold: It stabilizes the helix involved in DNA recognition and aids in orientating the peptide fold of the second zinc site, which is critical to the dimerization process.

The difference of the motifs in the GR proteins from that of the zinc fingers calls for further discussion. The DNA recognition site of GR lies *between* 2 fully tetracoordinated zinc anchors complexed to 4 cysteine residues. Thus, the GR binding mode can only be expected to be found in proteins whch contain *at least two zinc atoms*. The DNA recognition site of zinc fingers, on the other hand, connects the two Cys with the two His residues which together bind to each of the zinc atoms such that it lies at the base of each finger. This could occur in a protein even when it contains just one zinc atom.

(D) Zinc Clusters. The GAL4 transcription factor from *Saccharomyces cerevisiae* has been shown to contain zinc.[11] ^{1}H and ^{113}Cd NMR studies of either two-zinc or two-cadmium GAL4(62*)[78] and GAL4(7–49)[79] have shown that the two metal atoms coordinate to the six cysteines in a binuclear zinc thiolate cluster (Figure 3C).[80]

The DNA-recognition sites of zinc clusters have been suggested[11] to be in the region of residues 12–19, based on two findings: this region has a propensity to form an α-helix,[79] and in the GAL4 family of fungal transcription factors it contains several Lys or Arg residues.[80] The side chain of such amino acids thus could interact with DNA bases and/or the phosphate backbone.

The crystal structure of the yeast transcriptional factor fragment GAL4(1–65) bound to a 17 bp DNA fragment that is a consensus of 11 known GAL4 binding sites has recently been reported.[81] The GAL4(1–65) binds to the DNA as a symmetrical dimer. Each subunit folds into three distinct modules: the zinc cluster (8–40) and extended linker (41–49) and an α-helical dimerization element (50–64). The zinc cluster contains two short α-helices, residues 10–27 and 27–39, with the first and fourth residues being zinc binding Cys residues. The zinc cluster recognizes a conserved CCG triplet at each end of the site through direct contacts with the major groove.

Specific base interactions occur at Lys-17 and -18 (Figure 3C). The carbonyl of the amide of Lys-17 accepts a hydrogen bond from the N(4) nitrogen of cytosine 8 while the carbonyl of Lys 18 accepts H bonds from N(4) nitrogens of both cytosines 7 and 8. The ϵ-amino group donates H-bonds to the O(6) of guanine 6 and to N(7) of guanine 7. The zinc stabilizes these interactions since it is coordinated to both Cys-14 and Cys-21, both part of this α-helical region and flanking the Lys residues critical to recognizing the DNA triplet.

(73) Härd, T.; Kellenbach, E.; Boelens, R.; Kaptein, R.; Dahlman, K.; Carlstedt-Duke, J.; Freedman, L. P.; Maler, B. A.; Hyde, E. I.; Gustafsson, J.-Å; Yamamoto, K. R. *Biochemistry* **1990**, *29*, 9015.

(74) Pan, T.; Freedman, L. P.; Coleman, J. E. *Biochemistry* **1990**, *29*, 9218.

(75) Kellenbach, E.; Maler, B. A.; Yamamoto, K. R.; Boelens, R.; Kaptein, R. *FEBS Lett.* **1991**, *291*, 367.

(76) Schwabe, J. W. R.; Neuhaus, D.; Rhodes, D. *Nature* **1990**, *348*, 458.

(77) Luisi, B. F.; Xu, W. X.; Otwinowski, Z.; Freedman, L. P.; Yamamoto, K. R.; Sigler, P. B. *Nature* **1991**, *352*, 497.

(78) Pan, T.; Coleman, J. E. *Biochemistry* **1991**, *30*, 4212.

(79) Gadhavi, P. L.; Davis, A. L.; Povey, J. F.; Keeler, J.; Laue, E. D. *FEBS Lett.* **1991**, *281*, 223.

(80) Vallee, B. L.; Coleman, J. E.; Auld, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 999.

(81) Marmorstein, R.; Corey, M.; Ptashne, M.; Harrison, S. H. *Nature* **1992**, *356*, 408.

The functional role of the zinc cluster seen for the GAL4 transcription factors may also be applicable to the Cu^{1+} binding yeast transcriptional factor ACE1[82] that regulates the expression of metallothionein genes. This factor has been proposed to coordinate metals in a metallothionein-like cluster arrangement. Strikingly, it has a sequence in the proposed metal binding domain A*C*ETC*R*G**HR** that is very similar to the DNA recognition site of GAL4, A*C*DI*C*RL**KK**, suggesting a similar DNA metal cluster interaction site for this factor.

(82) Fürst, P.; Hu, S.; Hackett, R.; Hamer, D. *Cell* **1988**, *55*, 705.

In both GR and GAL4, cysteine residues were known to be the sole zinc ligands (Figure 3B,C), but the hypotheses regarding their zinc binding sites advanced on the basis of the consensus sequences alone do not turn out to coincide with the results of the NMR structure determination.[80] The primary sequences of GAL4 and GR are necessary but not sufficient to permit definitive structural predictions. Considering the limited number of structure determinations of DNA-binding zinc proteins so far, broad extrapolations from primary to tertiary structure and their relationship to function would seem premature.