Biochemistry

 Volume 29, Number 24 June 19, 1990

Perspectives in Biochemistry

Zinc Coordination, Function, and Structure of Zinc Enzymes and Other Proteins†

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 March 30, 1990; Revised Manuscript Received April 18, 1990

Zinc is an essential component of many enzymes involved in virtually all aspects of metabolism. It has been found to be an integral component of nearly 300 enzymes in different species of all phyla, indispensable to their functions, which encompass the synthesis and/or degradation of all major metabolites. Advances in the isolation and characterization of enzymes and analysis of metals were basic to the rapid growth of the field. The capacity to measure picograms of zinc in nanograms of proteins is now widespread, and such analyses of biological matter have ceased to be a problem (Vallee, 1988; Riordan & Vallee, 1988).

About a decade ago, the role of zinc in gene expression began to attract interest (Vallee, 1977a,b; Vallee & Falchuk, 1981; Hanas et al., 1983) and this field has recently gained wide attention. In addition to its roles in catalysis and gene expression, zinc stabilizes the structure of proteins and nucleic acids, preserves the integrity of subcellular organelles, participates in transport processes, and plays important roles in viral and immune phenomena (Vallee & Auld, 1990a). Its nutritional essentiality has focused attention on the pathology and clinical consequences of both its deficiency and toxicity (Vallee, 1986).

A perspective on the biological occurrence and role of zinc should not only seek out the generalizations that are now apparent but should also call attention to the participation of zinc in metabolic processes, the bases of which have yet to be explained. We shall therefore highlight the salient accomplishments of the past two decades, achieved primarily through the study of zinc enzymes, while calling attention to the urgent but unanswered questions regarding the functions of zinc in metallothionein and DNA-binding proteins.

We will first consider the 12 zinc enzymes whose crystal structures are known and are now the standards of reference for catalytically active and structural zinc sites. In all of these a catalytic zinc atom is coordinated to three amino acid residues of the protein and an activated water molecule, whereas structural zinc atoms are coordinated to four cysteine residues. In members of their respective families whose crystal structures and/or sequences are known, the situation is identical. We will further examine the predictive capacity of such crystal structures for other zinc enzymes whose functions may be related and whose sequence is known but whose three-dimensional structure is not. This perspective will also deal with the induction of enzymatic activity in the procollagenases, which is accomplished through alterations in zinc coordination chemistry. It will conclude with the zinc cluster structure of the metallothioneins and its potential relevance to the DNA-binding zinc proteins and the problem of zinc coordination in those proteins in general.

(A) Crystal Structures of Zinc Enzymes as Standards of Reference

The realization that the affinity of zinc for nitrogen and oxygen is nearly equal to that for sulfur ligands has brought about a major change of viewpoint regarding zinc coordination chemistry and its manifestation in biological systems. Present perceptions indicate that there is as great a need to inspect the chemical properties of zinc complex ions containing combinations of nitrogen, oxygen, and sulfur ligands as well as the properties of those that only contain sulfur ligands.

X-ray crystallographic analysis of 12 zinc enzymes has identified the zinc ligands and defined the modes of coordination at both their active and structural sites. The primary and tertiary structures of these zinc enzymes provide standards of reference for the zinc sites in the sequences of protein families (Vallee & Auld, 1989, 1990b). These 12 zinc enzymes represent examples from classes I, II, III, and IV of the six classes of enzymes (Table I). X-ray crystallographic analyses of zinc enzymes in classes V and VI are not yet on record. The ligands are now known with certainty, and they define the features of the catalytic and structural zinc binding sites of zinc enzymes.

† This work was supported by grants from the Endowment for Research in Human Biology, Inc. (Boston, MA).

* To whom correspondence should be addressed.

0006-2960/90/0429-5647$02.50/0

	97			100			103								111			Ref.
Horse E	P	Q	C	G	K	C	R	V	C	K	H	P	E	G	N	F	C	L	K	1
Human α	P	Q	C	G	K	C	R	I	C	K	N	P	E	S	N	Y	C	L	K	2
Human β	P	Q	C	G	K	C	R	V	C	K	N	P	E	S	N	Y	C	L	K	3
Human γ	P	Q	C	G	K	C	R	I	C	K	N	P	E	S	N	Y	C	L	K	4
Human π	P	L	C	R	K	C	K	F	C	L	S	P	L	T	N	L	C	G	K	5
Human χ	P	Q	C	G	E	C	K	F	C	L	N	P	K	T	N	L	C	Q	K	6
Baboon*	P	Q	C	G	K	C	R	V	C	K	S	P	E	G	N	Y	C	V	K	7
Horse χ	P	Q	C	G	E	C	K	F	C	L	N	P	Q	T	N	L	C	Q	K	8
Rat χ	P	Q	C	G	E	C	K	F	C	L	N	P	K	T	N	L	C	Q	K	9
Rat	P	Q	C	G	K	C	R	I	C	K	H	P	E	S	N	L	C	C	Q	10
Mouse*	P	Q	C	G	E	C	R	I	C	K	H	P	E	S	N	F	C	S	R	11
Chicken*	P	Q	C	G	E	C	R	S	C	L	S	T	K	G	N	L	C	I	K	12
Quail*	P	Q	C	G	E	C	R	S	C	L	S	T	K	G	N	L	C	I	K	13
Maize 1*	G	E	C	K	E	C	A	H	C	K	S	A	E	S	N	M	C	D	L	14
Maize 2*	G	E	C	K	E	C	A	H	C	K	S	E	E	S	N	M	C	D	L	15
Barley*	G	E	C	K	E	C	P	H	C	K	S	A	E	S	N	M	C	D	L	16
Potato*	G	E	C	K	D	C	A	H	C	K	S	E	E	S	N	M	C	S	L	17
Pea*	G	E	C	G	E	C	P	H	C	K	S	E	E	S	N	M	C	D	L	18
Arab.*	G	E	C	G	D	C	R	H	C	Q	S	E	E	S	N	M	C	D	L	19
Asper.*	G	S	C	L	S	C	E	M	C	M	Q	A	D	E	P	L	C	P	H	20
Yeast 1*	G	S	C	M	A	C	E	Y	C	E	L	G	N	E	S	N	C	P	H	21
Yeast 2*	G	S	C	M	A	C	E	Y	C	E	L	G	N	E	S	N	C	P	H	22
Yeast 3*	G	S	C	M	T	C	E	F	C	E	S	G	H	E	S	N	C	P	D	23
Yeast 4*	S	S	C	G	N	C	E	Y	C	M	K	A	E	E	T	I	C	P	H	24

FIGURE 1: Structural-site zinc ligands of alcohol dehydrogenase. Lightly stippled boxes denote the enzyme(s) X-ray standard of reference for each family. An asterisk denotes those for which zinc was not measured directly. Black vertical columns indicate the proposed metal-binding ligands based on the structure of the standard of reference. Tetrameric ADH enzymes are shown at the bottom. Arab. and Asper. refer to *Arabidopsis thaliana* and *Aspergillus nidulans*, respectively. Yeasts 1 and 2 are respectively *Saccharomyces cerevisiae* cytosolic isozymes 1 and 2, yeast 3 is a *S. cerevisiae* mitochondrial enzyme, and yeast 4 is a *Schizosaccharomyces pombe* enzyme. References: (1) Jörnvall, H. (1970) *Eur. J. Biochem. 16*, 41–49; (2) von Bahr-Lindström, H., Höög, J.-O., Hedén, L.-O., Kaiser, R., Fleetwood, L., Larsson, K., Lake, M., Holmquist, B., Holmgren, A., Hempel, J., Vallee, B. L., & Jörnvall, H. (1986) *Biochemistry 25*, 2465–2470; (3) Duester, G., Smith, M., Bilanchone, V., & Hatfield, G. W. (1986) *J. Biol. Chem. 261*, 2027–2033; (4) Höög, J. O., Hedén, L.-O., Larsson, K., Jörnvall, H., & von Bahr-Lindström, H. (1986) *Eur. J. Biochem. 159*, 215–218; (5) Höög, J.-O., von Bahr-Lindström, H., Hedén, L.-O., Holmquist, B., Larsson, K., Hempel, J., Vallee, B. L., & Jörnvall, H. (1987) *Biochemistry 26*, 1926–1932; (6) Kaiser, R., Holmquist, B., Hempel, J., Vallee, B. L., & Jörnvall, H. (1988) *Biochemistry 27*, 1132–1140; (7) Trezise, A. E. O., Godfrey, E. A., Holmes, R. S., & Beacham, I. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5454–5458; (8) H. Jörnvall, personal communication; (9) Julia, P., Pares, X., & Jörnvall, H. (1988) *Eur. J. Biochem. 172*, 73–83; (10) Crabb, D. W., & Edenberg, H. J. (1986) *Gene 48*, 287–291; (11) Edenberg, H. J., Zhang, K., Fong, K., Bosron, W. F., & Li, T. K. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2262–2266; (12) H. Jörnvall, personal communication; (13) *ibid.*; (14) Brandén, C. I., Eklund, H., Cambillau, C., & Pryor, A. J. (1984) *EMBO J. 3*, 1307–1310; (15) Dennis, E. S., Sachs, M. M., Gerlach, W. L., Finnegan, E. J., & Peacock, W. J. (1985) *Nucleic Acids Res. 13*, 727–743; (16) Good, A. G., Pelcher, L. E., & Crosby, W. L. (1988) *Nucleic Acids Res. 16*, 7182; (17) D. P. Matton and M. Brisson, in GenBank/Los Alamos; (18) Llewellyn, D. J., Finnegan, E. J., Ellis, J. G., Dennis, E. S., & Peacock, W. J. (1987) *J.* Mol. *Biol. 195*, 115–123; (19) Chang, C., & Meyerowitz, E. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1408–1412; (20) McKnight, G. L., Kato, H., Upshall, W., Parker, M. D., Saari, G., & O'Hara, P. J. (1985) *EMBO J. 4*, 2093–2099; (21) Jörnvall, H. (1977) *Eur. J. Biochem. 72*, 425–442; (22) Russell, D. W., Smith, M., Williamson, V. M., & Young, E. T. (1983) *J. Biol. Chem. 258*, 2674–2682; (23) Young, E. T., & Pilgrim, D. (1985) *Mol. Cell. Biol. 5*, 3024–3034; Young, E. T., & Pilgrim, D. (1986) *Mol. Cell. Biol. 6*, 2284; (24) Russell, P. R., & Hall, B. D. (1983) *J. Biol. Chem. 258*, 143–149.

Searches of the protein and gene sequence data banks and literature have ascertained the structural homology or identity and functional similarity of zinc enzyme families related by common ancestry.[1] Their characteristics and the identity as well as the conformations of their zinc ligands will likely prove critical to the specificity and mechanisms of action of these enzymes. On the basis of the reference structures, a combination of three His, Glu, Asp, or Cys residues creates a tridentate active zinc site (Table I). "Activated" H_2O fills and completes the coordination sphere in all such enzymatically active zinc sites, which contrasts with tetradentate structural zinc sites in which four cysteines are coordinated to zinc. Histidine is the most frequent ligand in active sites. In noncoenzyme-dependent zinc enzymes, a short spacer of 1–3 amino acids, intervening between the first two ligands L_1 and L_2, provides a zinc-binding nucleus. The third ligand, L_3, separated from L_2 by a long spacer of ~20 to ~120 amino acids, completes the coordination sphere with a long polypeptide loop that further aligns protein residues with the zinc and thereby also brings about interactions with the substrate. In our efforts to delineate the role of zinc in catalysis, we have centered on the chemistry of zinc and its effect on interacting proteins and vice versa rather than on the details of the participation of zinc in the mechanism of action of any particular zinc enzyme.

Carboxypeptidases A and B of bovine pancreas, the neutral proteases of *Bacillus thermoproteolyticus* and *Bacillus cereus*, carbonic anhydrases I and II of human red blood cells, and the dimeric alcohol dehydrogenase of equine liver are the standards of reference for those members of their respective families whose sequence is known but whose three-dimensional structure is not. Specifically, the identities of the zinc ligands and the amino acid sequences in their immediate vicinities are compared. The findings may have general implications for the elucidation of the mechanisms of related enzymes and the design of enzymatically active model systems.

[1] The National Biomedical Research Foundation and GenBank/Los Alamos database files of the Molecular Biology Computer Research Resource at Harvard Medical School were used for these searches.

(B) Structural Zinc Atoms of Enzymes

(1) Class I: Alcohol Dehydrogenase. In alcohol dehydrogenase (ADH)[2] (EC 1.1.1.1), zinc can have either a

Table I: Zinc Ligands and Their Spacing for the Catalytic and Structural Zinc[a]

enzyme	L_1	X	L_2	Y	L_3	Z	L_4	ref[b]
Class I								
alcohol dehydrogenase	Cys	20	His	106	Cys (C)		H_2O	1
alcohol dehydrogenase[c]	Cys	2	Cys	2	Cys (C)	7	Cys (C)	1
Class II								
aspartate transcarbamoylase[c]	Cys	4	Cys	22	Cys (C)	2	Cys (C)	2
Class III								
carboxypeptidase A	His	2	Glu	123	His (C)		H_2O	3
carboxypeptidase B	His	2	Glu	123	His (C)		H_2O	4
thermolysin	His	3	His	19	Glu (C)		H_2O	5
B. cereus neutral protease	His	3	His	19	Glu (C)		H_2O	6
DD carboxypeptidase	His	2	His	40	His (N)		H_2O	7
β-lactamase	His	1	His	121	His (C)		H_2O	8
phospholipase C	His	3	Glu	13	His (N)		H_2O	9
alkaline phosphatase	Asp	3	His	80	His (C)		H_2O	10
Class IV								
carbonic anhydrase I	His	1	His	22	His (C)		H_2O	11
carbonic anhydrase II	His	1	His	22	His (C)		H_2O	12

[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; Z is the number of amino acids between L_3 and L_4. L_3 is contributed by either the amino (N) or the carboxyl (C) portion of the protein. [b] References: (1) Brandén et al., (1975; (2) Honzatko et al., 1982; (3) Rees et al., 1983; (4) Schmid & Herriott, 1976; (5) Matthews et al., 1972; (6) Pauptit et al., 1988; (7) Dideberg et al., 1982; (8) Sutton et al., 1987; (9) Hough et al., 1989; (10) Kim & Wyckoff, 1989; (11) Kannan et al., 1975; (12) Liljas et al., 1972. [c] Structural zinc site; all others are in catalytic zinc sites.

structural or a catalytic (see below) role. The structural zinc atom is bound tetrahedrally to the four sulfur atoms of Cys-97, -100, -103, and -111 (Eklund & Brandén, 1983), all close to one another in the primary sequence and separated by 2, 2, and 7 amino acid residue spacers, respectively (Figure 1). Remarkably, this zinc atom, inaccessible to solvent, is close to the surface of the molecule. Its cysteine ligands are part of a lobe that projects out of the catalytic domain and has only a few side-chain interactions with the remainder of the subunit. These circumstances together with energy calculations lead to the inference that this zinc atom primarily affects local structure and conformation.

(2) *Class II*: *Aspartate Transcarbamoylase*. Aspartate transcarbamoylase (ATCase) (EC 2.1.3.2) contains but a single structural zinc atom located in the regulatory subunit, bound tetrahedrally to four cysteines, Cys-109, -114, -137, and -140, separated by 4, 22, and 2 intervening amino acid residue spacers (Honzatko et al., 1982). The polypeptide chain in this region forms two loops, which the zinc atom holds together. Stabilization of these loops, which form part of the interface between the regulatory and catalytic subunits, is thought to be responsible for stabilizing the quaternary structure. Zinc likely influences the local conformation and structure of the regulatory subunit to fine-tune its interaction with the catalytic subunit, as is apparent from the entropic contribution to the activation energy of zinc binding at such sites (Keating et al., 1988). Similar considerations and conclusions could pertain to the role of zinc atoms in DNA-binding proteins (see below). While in these two enzymes such circumstances may well affect enzymatic activity indirectly, there is no evidence of direct involvement of the zinc atom in their enzymatic—or any other—function; there is, of course, no zinc atom at the active site of the catalytic subunits of ATCase. In both instances cited, ADH and ATCase, the *structural* zinc atoms are *fully coordinated tetrahedrally* to cysteines. Neither water nor substrate has access to an open coordination site of that metal atom in these molecules, which apparently maintains the local structure in its immediate vicinity.

[2] Abbreviations: ADH, alcohol dehydrogenase; ATCase, aspartate transcarbamoylase; collag, collagenase; L_1, L_2, L_3, and L_4, the first, second, third, and fourth zinc-binding ligand, respectively; MC, mast cell; NEP, neutral endoprotease; TFIIIA, transcription factor IIIA; TL, thermolysin.

The prominence of the sulfur atoms of cysteine in the coordination chemistry of structural sites of these enzymes coincides with earlier views from inorganic chemistry and geochemistry of a predilection of zinc for sulfur ligands. The interaction of zinc with proteins that play critical roles in the regulation of gene transcription has revived such views (Klug & Rhodes, 1987).

(C) Functional Zinc Atoms of Enzymes

(1) *Class I*: *Alcohol Dehydrogenases*. The ligands of the active-site zinc atom of alcohol dehydrogenase (EC 1.1.1.1) differ significantly from those encountered thus far in all other zinc enzymes examined by X-ray crystallography. This is the only coenzyme-dependent zinc enzyme whose three-dimensional structure is known. The involvement of both zinc and NADH in the catalytic process requires suitable alignment of amino acid residues that can provide for both metal chelation and coenzyme binding sites. In ADH this has been accomplished by the use of residues 46 and 47 as zinc- and NADH-binding ligands, respectively, with two cysteines and one histidine as ligands to the active-site zinc, and by extending the short spacer between L_1 and L_2 to 20 amino acids.

The peptide backbones of all three zinc ligands are firmly anchored in secondary structural elements; none is part of a flexible loop region (Eklund & Brandén, 1983). Cys-46 (L_1) is simultaneously the first residue of a short α-helix, comprising residues 46–55, and the last residue of a β-strand. His-67 (L_2) is the first residue of a β-strand, while Cys-174 (L_3) is in the middle of an α-helix. The variation observed in residue 47, one of the two sites to which the coenzyme NAD(H) binds through its phosphate, contrasts markedly with the invariant nature of a number of other amino acids around both Cys-46 (L_1) and His-67 (L_2) (Figure 2).

A "long" spacer of 106 amino acids separates His-67 (L_2) from Cys-174 (L_3). Considering that the known sequences of dimeric and tetrameric ADHs span a remarkable range of evolution, the homology around this Cys residue is notable. In all of these ADHs a H_2O molecule completes the zinc coordination sphere. Among the enzymes whose structure is known so far, alcohol dehydrogenase is the only one in which cysteine is an active-site zinc ligand.

(2) *Class IV*: *Lyases*. The crystal structure of carbonic anhydrases I (Kannan et al., 1975) and II (Liljas et al., 1972) from human erythrocytes (EC 4.2.1.1) are known. The ac-

	46						67							174						Ref.
Horse E	T	G	I	C	R	S	D	I	A	G	H	E	A	A	C	L	I	G	C	G	F	S	T	1
Human α	V	G	I	C	G	T	D	I	L	G	H	E	A	A	C	L	I	G	C	G	F	S	T	2
Human β	V	G	I	C	R	T	D	I	L	G	H	E	A	A	C	L	I	G	C	G	F	S	T	3
Human γ	A	G	I	C	R	S	D	I	L	G	H	E	A	A	C	L	I	G	C	G	F	S	T	4
Human π	T	S	L	C	H	T	D	I	V	G	H	E	A	A	C	L	L	G	C	G	F	S	T	5
Human χ	T	A	V	C	H	T	D	I	L	G	H	E	G	A	C	L	L	G	C	G	I	S	T	6
Baboon*	V	G	I	C	R	T	D	I	L	G	H	E	A	A	C	L	I	G	C	G	F	S	T	7
Horse χ	T	A	V	C	H	T	D	I	L	G	H	E	G	A	C	L	L	G	C	G	V	S	T	8
Rat χ	T	A	V	C	H	T	D	I	L	G	H	E	G	A	C	L	L	G	C	G	I	S	T	9
Rat	T	G	V	C	R	S	D	V	L	G	H	E	G	A	C	L	I	G	C	G	F	S	T	10
Mouse*	T	G	V	C	R	S	D	V	L	G	H	E	G	A	C	L	I	G	C	G	F	S	T	11
Chicken*	T	G	I	C	R	S	D	I	L	G	H	E	A	A	C	L	I	G	C	G	F	S	T	12
Quail*	T	G	I	C	R	S	D	I	L	G	H	E	A	A	C	L	I	G	C	G	F	S	T	13
Maize 1*	T	S	L	C	H	T	D	I	F	G	H	E	A	G	C	V	L	S	C	G	I	S	T	14
Maize 2*	T	A	L	C	H	T	D	I	L	G	H	E	A	G	C	I	L	S	C	G	I	S	T	15
Barley*	T	S	L	C	H	T	D	I	F	G	H	E	A	G	C	V	L	S	C	G	I	S	T	16
Potato*	T	S	L	C	H	T	D	I	L	G	H	E	A	A	C	V	L	S	C	G	I	S	T	17
Pea*	T	S	L	C	H	T	D	I	F	G	H	E	A	G	C	I	L	S	C	G	I	C	T	18
Arab.*	T	S	L	C	H	T	D	I	F	G	H	E	A	G	C	I	V	S	C	G	L	S	T	19
Asper.*	S	G	V	C	H	T	D	I	G	G	H	E	G	A	A	P	I	L	C	A	G	I	T	20
Yeast 1*	S	G	V	C	H	T	D	V	G	G	H	E	G	A	A	P	V	L	C	A	G	I	T	21
Yeast 2*	S	G	V	C	H	T	D	V	G	G	H	E	G	A	A	P	I	L	C	A	G	I	T	22
Yeast 3*	S	G	V	C	H	T	D	V	G	G	H	E	G	A	A	P	I	L	C	A	G	V	T	23
Yeast 4*	T	G	V	C	H	T	D	I	G	G	H	E	G	A	A	P	I	M	C	A	G	I	T	24

FIGURE 2: Active-site zinc ligands for alcohol dehydrogenases. For definition of abbreviations, references, and key, see Figure 1.

	94			96			119								Ref.
Human I	F	Q	F	H	F	H	W	G	S	S	A	E	L	H	V	A	H	W	1
Human II	I	Q	F	H	F	H	W	G	S	A	A	E	L	H	L	V	H	W	2
Bovine I	F	Q	F	H	F	H	W	G	I	S	A	E	L	H	L	V	H	W	3
Mouse I*	T	Q	F	H	F	H	W	G	N	S	G	E	L	H	L	V	H	W	4
Horse I*	V	Q	F	H	F	H	W	G	S	S	A	E	L	H	L	V	H	W	5
Rabbit I*	S	Q	F	H	F	H	W	G	K	S	A	E	L	H	L	V	H	W	6
Monkey*	F	Q	F	H	F	H	W	G	S	S	S	E	L	H	I	V	H	W	7
Bovine II	V	Q	F	H	F	H	W	G	S	A	A	E	L	H	L	V	H	W	8
Mouse II*	I	Q	F	H	F	H	W	G	S	A	A	E	L	H	L	V	H	W	9
Rabbit II*	I	Q	F	H	F	H	W	G	S	A	A	E	L	H	L	V	H	W	10
Sheep II*	V	Q	F	H	F	H	W	G	S	A	A	E	L	H	L	V	H	W	11
Chicken II*	V	Q	F	H	I	H	W	G	S	D	A	E	L	H	I	V	H	W	12
Human III	R	Q	F	H	L	H	W	G	S	A	A	E	L	H	L	V	H	W	13
Bovine III	R	Q	F	H	L	H	W	G	S	A	A	E	L	H	L	V	H	W	14
Horse III*	R	Q	F	H	L	H	W	G	S	A	A	E	L	H	L	V	H	W	15

FIGURE 3: Zinc ligands of carbonic anhydrases. For key to figures, see Figure 1. References: (1) Barlow, J. H., Lowe, N., Edwards, Y. H., & Butterworth, P. H. W. (1987) *Nucleic Acids Res. 15*, 2386; (2) Henderson, L. E., Henriksson, D., & Nyman, P. O. (1976) *J. Biol. Chem. 251*, 5457–5463; (3) Sciaky, M., Limozin, N., Filippi-Foveau, D., Gulian, J.-M., & Laurent-Tabusse, G. (1976) *Biochimie 58*, 1071–1082; (4) Fraser, P. J., & Curtis, P. J. (1986) *J. Mol. Evol. 23*, 294–299; (5) Jabusch, J. R., Bray, R. P., & Deutsch, H. F. (1980) *J. Biol. Chem. 255*, 9196–9204; (6) Konialis, C. P., Barlow, J. H., & Butterworth, P. H. W. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 663–667; (7) Henriksson, D., Tanis, R. J., & Tashian, R. E. (1980) *Biochem. Biophys. Res. Commun. 96*, 135–142; (8) Tashian, R. E., Hewett-Emmett, S. K., Stroup, M., Goodman, M., & Yu, Y.-S. L. (1980) in *Biophysics and Physiology of Carbon Dioxide* (Bauer, C., Gros, G., & Bartels, H., Eds.) p 165, Springer, Berlin, FRG; (9) Curtis, P. J., Withers, E., Demuth, D., Watt, R., Venta, P. J., & Tashian, R. E. (1983) *Gene 25*, 325–332; Venta, P. J., Montgomery, J. C., Hewett-Emmet, D., Wiebauer, K., & Tashian, R. E. (1985) *J. Biol. Chem. 260*, 12130–12135; (10) Ferrell, R. E., Stroup, S. K., Tanis, R. J., & Tashian, R. E. (1978) *Biochim. Biophys. Acta 533*, 1–11; (11) Tanis, R. J., Ferrell, R. E., & Tashian, R. E. (1974) *Biochim. Biophys. Acta 371*, 534–548; (12) Roger, J. H. (1987) *Eur. J. Biochem. 162*, 119–122; (13) Lloyd, J., McMillan, S., Hopkinson, D., & Edwards, Y. H. (1986) *Gene 41*, 233–239; (14) Tashian, R. E., Hewett-Emmett, S. K., Stroup, M., Goodman, M., & Yu, Y.-S. L. (1980) in *Biophysics and Physiology of Carbon Dioxide* (Bauer, C., Gros, G., & Bartels, H., Eds.) p 165, Springer, Berlin, FRG; (15) Wendorff, K. M., Nishita, T., Jabusch, J. R., & Deutsch, H. F. (1985) *J. Biol. Chem. 260*, 6129–6132.

tive-site cavity is situated in the middle of a large twisted β-sheet of the protein (Liljas et al., 1972). A β-strand encompassing residues 88–108 supplies two of the zinc ligands, His-94 (L_1) and His-96 (L_2), while another β-sheet extending from residues 113 to 126 contributes His-119 (L_3). A H_2O molecule occupies the fourth coordination site, and the resulting geometry about the zinc is a distorted tetrahedron.

The shortest possible spacer, a single amino acid, separates His-94 (L_1) from His-96 (L_2); in 15 different carbonic anhydrases, five of the seven amino acids surrounding these ligands are 95% invariant (Figure 3). A long spacer of 22 amino acids supplies His-119 (L_3). In these 15 carbonic anhydrases, four of the eight amino acids surrounding His-119 (L_3) are invariant and the other four are highly similar.

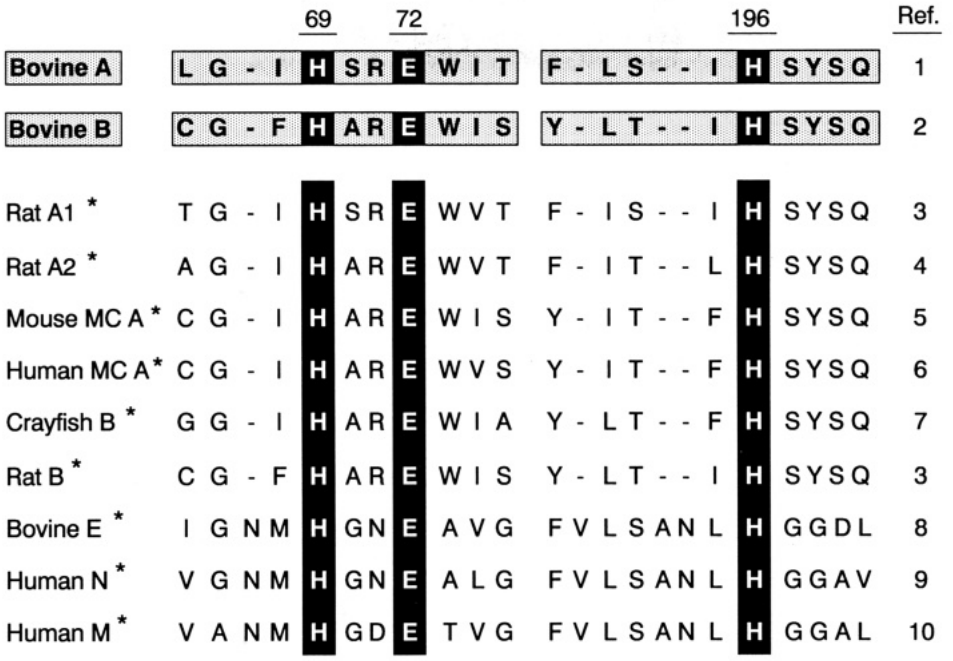

FIGURE 4: Zinc ligands of carboxypeptidases. For key to figures, see Figure 1. References: (1) Bradshaw, R. A., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1969) *Proc. Natl. Acad. Sci. U.S.A. 63*, 1369–1394; (2) Titani, K., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1666–1670; (3) Clauser, E., Gardell, S. J., Craik, C. S., MacDonald, R. J., & Rutter, W. J. (1988) *J. Biol. Chem. 263*, 17837–17845; (4) Gardell, S. J., Craik, C. S., Clauser, E., Goldsmith, E. J., Stewart, C.-B., Graf, M., & Rutter, W. J. (1988) *J. Biol. Chem. 263*, 17828–17836; (5) Reynolds, D. S., Stevens, R. L., Gurley, D. S., Lane, W. S., Austen, K. F., & Serafin, W. E. (1989) *J. Biol. Chem. 264*, 20094–20099; (6) Reynolds, D. S., Gurley, D. S., Stevens, R. L., Sugarbaker, D. J., Austen, K. F., & Serafin, W. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9480–9484; (7) Titani, K., Ericsson, L. H., Kumar, S., Jakob, F., Neurath, H., & Zwilling, R. (1984) *Biochemistry 23*, 1245–1250; (8) Fricker, L. D., Evans, C. J., Esch, F. S., & Herbert, E. (1986) *Nature 323*, 461–464; (9) Gebhard, W., Schube, M., & Eulitz, M. (1989) *Eur. J. Biochem. 178*, 603–607; (10) Tan, F., Chan, S. J., Steiner, D. F., Schilling, J. W., & Skidgel, R. A. (1989) *J. Biol. Chem. 264*, 13165–13170.

(*3*) *Class III*: *Hydrolases*. (*a*) *Metalloexoproteases*. Carboxypeptidase A (EC 3.4.17.1) is a member of a large family of zinc proteases (Vallee et al., 1983; Auld & Vallee, 1987, and references cited therein) and is one of those best characterized. The crystal structures of the bovine A and B enzymes show His-69 (L_1), Glu-72 (L_2), and His-196 (L_3) to be the three zinc ligands (Quiocho & Lipscomb, 1971; Schmid & Herriott, 1976). Two amino acid residues constitute the short amino acid spacer between His-69 (L_1) and Glu-72 (L_2), while the long spacer between Glu-72 (L_2) and His-196 (L_3) consists of 123 amino acids (Vallee & Auld, 1989). In eight carboxypeptidases A and B from human, cow, rat, mouse, and crayfish, all three ligands to the zinc atom are conserved, and very few changes occur in the residues adjacent to them (Figure 4).

His-69 (L_1) and Glu-72 (L_2) are at the ends of a reverse turn, while His-196 (L_3) is the last residue in a β pleated sheet structure extending from amino acids 191 to 196 (Rees et al., 1983). In addition, structure determination, chemical modification, and kinetic studies have identified Arg-145 and Glu-270 as functionally essential residues in the active site (Lipscomb et al., 1970; Vallee et al., 1970).

The sequences of a number of other carboxypeptidases have recently been established by means of DNA technology. Carboxypeptidase E or enkephalin convertase (Fricker et al., 1986) processes prohormones in secretory granules; carboxypeptidase N or kinase I (Gebhard et al., 1989), a blood enzyme, regulates peptide hormone activity at neutral pH; and carboxypeptidase M, a plasma membrane bound enzyme placed strategically at local tissue sites, also acts on peptide hormones (Tan et al., 1989). The sequence identity among these three proteins is approximately 41% but decreases to 15% when compared to carboxypeptidase A or B. However, suitable alignment of residues shows the preservation of all residues binding to the active-site zinc (Figure 4) as well as of Glu-270 and Arg-145 (Tan et al., 1989), which are essential to enzymatic function of carboxypeptidases A and B. A short spacer of 2 amino acid residues between L_1 and L_2 is common to all carboxypeptidases A, B, E, M, and N. The long spacer consists of 126, 130, and 138 amino acid residues for the human N, bovine E, and human M enzymes, respectively.

The most notable differences in the sequences of carboxypeptidases E, M, and N from those of carboxypeptidases A and B, which serve as standards, are the replacements of Tyr-198 and Arg-71, considered to have roles in substrate binding but not in catalysis, by Gly and Asn, respectively. Further, Ala or Thr replaces Trp-73 (Figure 4).

(*b*) *Metalloendoproteases*. The pH optimum of thermolysin (TL) (EC 3.4.24.4) from *B. thermoproteolyticus*, M_r 34 000, is near neutrality, representative of a number of neutral metalloproteinases. It contains one catalytic zinc and several structural calcium atoms[3] (Holmquist & Vallee, 1974) and has been studied extensively both in solution and in crystals. The zinc atom is bound to His-142 (L_1), His-146 (L_2), and Glu-166 (L_3) (Matthews et al., 1974). These residues are located around the zinc atom in such a manner that a water molecule completes a distorted tetrahedral coordination (Matthews et al., 1972). A short three amino acid spacer separates L_1 from L_2. Both are part of an α-helical domain behind the zinc atom extending from residues 137 to 150, approximately parallel to the cleft and traversing the entire center of the molecule. Another long internal helix, residues

[3] All four calcium sites—in which Glu and Asp are the principal ligands—are distant from the zinc atom. The calcium atoms, arranged in distorted octahedral coordination geometry, generate a complex network that may stabilize the enzyme structurally and thermally (Matthews et al., 1974).

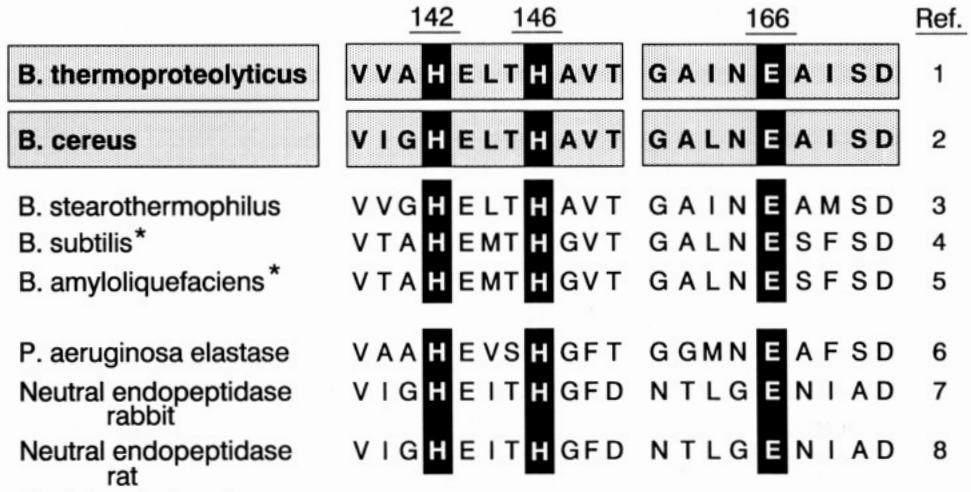

FIGURE 5: Zinc ligands of thermolysin and neutral proteases. For key to figures, see Figure 1. References: (1) Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1972) *Nature, New Biol. 238*, 35–37; (2) Sidler, W., Niederer, E., Suter, F., & Zuber, H. (1986) *Biol. Chem. Hoppe-Seyler 367*, 643–657; (3) Takagi, M., Imanaka, T., & Aiba, S. (1985) *J. Bacteriol. 163*, 824–831; (4) Levy, P. L., Pangburn, M. K., Burstein, Y., Ericsson, L. H., Neurath, H., & Walsh, K. A. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4341–4345; (5) Vasantha, N., Thompson, L. D., Rhodes, C., Banner, C., Nagle, J., & Filpula, D. (1984) *J. Bacteriol. 159*, 811–819; (6) Beuer, R. A., & Iglewski, B. H. (1988) *J. Bacteriol. 170*, 4309–4314; (7) Devault, A., Lazure, C., Nault, C., Le Moual, H., Seidah, N. G., Chrêtien, M., Kahn, P., Powell, J., Mallet, J., Beaumont, A., Roques, B. P., Crine, P., & Boileau, G. (1987) *EMBO J. 6*, 1317–1322; (8) Malfroy, B., Schofield, P. R., Kuang, W.-J., Seeburg, P. H., Mason, A. J., & Henzel, W. J. (1987) *Biochem. Biophys. Res. Commun. 144*, 59–66.

160–180, underlying the central helix and placed at about 135° to it, contributes L_3, which is separated from L_2 by a long spacer of 19 amino acids. Thus, the secondary protein structure that supplies the zinc ligands in thermolysin differs markedly from that in the carboxypeptidases (see above). Moreover, among these two enzyme families both the length of the spacers and the order of the histidines in the sequence (His, His, Glu in the carboxypeptidases versus His, Glu, His in thermolysin) differ significantly (Vallee & Auld, 1989).

The crystal structure of the closely related neutral protease from *B. cereus* has recently been defined at 3.0-Å resolution (Pauptit et al., 1988). Its sequence is 73% identical with that of thermolysin and is particularly notable for all the residues flanking the zinc-binding ligands (Figure 5). Thermolysin and three bacterial neutral proteases from *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus amyloliquefaciens* share the same specificity, zinc ligands His-142 (L_1), His-146 (L_2), and Glu-66 (L_3), near identity of the proximal amino acids, and the lengths of their short and long spacers (Figure 5). Glu-143, which has been thought to function as a general base, is present in all of the above neutral proteases.

The sequence of the elastase from *Pseudomonas aeruginosa*, an extracellular enzyme, is quite similar to that of thermolysin. While the overall amino acid identity between the mature elastase and thermolysin is 49%, their respective amino and carboxyl termini differ. Identity within the central portions of the two molecules is much higher, i.e., 67%. On that basis, His-140, His-144, and Glu-164 of this elastase have been predicted to be the equivalents of L_1, L_2, and L_3, respectively, in thermolysin (Beuer & Iglewski, 1988) (Figure 5). The long and short spacers would therefore be identical with those of thermolysin. The amino acids flanking these putative zinc ligands also closely resemble those of other members of the thermolysin family.

The specificity of neutral endopeptidase (NEP) (EC 3.4.24.11) from rabbit kidney brush border is similar to that of thermolysin; it contains 1.1 mol of zinc/mol of 93 000 protein (Kerr & Kenny, 1974) and is thought to be identical with the enkephalinase in brain membrane fractions that cleaves the Gly–Phe bond in Met- and Leu-enkephalins (Auld, 1987, and references cited therein). The amino acid sequence of NEP (749 amino acids) (Devault et al., 1987; Malfroy et al., 1987) differs greatly from that of thermolysin (316 amino acids). There are only two short, albeit highly homologous, regions of similarity (Devault et al., 1987). Yet, as judged by a hydrophobic clustering analysis (Benchetrit et al., 1988), all amino acids involved in the catalytic sites of thermolysin appear to be conserved in NEP. The His-583 and His-587 of one of these short homologous regions could correspond to His-142 (L_1) and His-146 (L_2) of thermolysin (Figure 5). Their substitution by Phe in NEP completely abolishes both the activity and interaction of the recombinant enzymes with a neutral endopeptidase metal-binding inhibitor (Devault et al., 1988). Similarly, the substitution of Val or Asp for Glu-646 results in the loss of both enzymatic activity and binding of the inhibitor (Le Moual et al., 1989), a result that could be consistent with Glu-646 being the third zinc ligand.

The amino acid sequences flanking the proposed His-583 (L_1) and His-587 (L_2) are very similar to those of thermolysin, and the 3 amino acid spacer between them is identical. However, adjacent to Glu-646, the similarity of amino acids in NEP and the thermolysin family decreases greatly. Further, the 58 amino acid long spacer of the NEP differs significantly from that of the thermolysin family (Figure 5).

(D) Zinc Enzyme Families for Which Crystal Structure Standards of Reference Do Not Exist As Yet

(1) Class III: Hydrolases. The structures of thermolysin and the neutral protease of *B. cereus*, which are the only ones known for the neutral proteases, have served for comparison with sequences of other metalloproteinases (Stöcker et al., 1988; Jongeneel et al., 1989). Compared with the above examples, in these instances it is less clear whether or not these belong to the same family of enzymes that the structural models represent. Such comparisons may or may not be warranted but are presented here as current viewpoints, clearly subject to reinterpretation once more specific structural standards become available. A persuasive case has been made for the prediction of the active zinc-binding sites in the neutral proteases from the structures, though a good many uncertainties remain. Only structure determination of a member

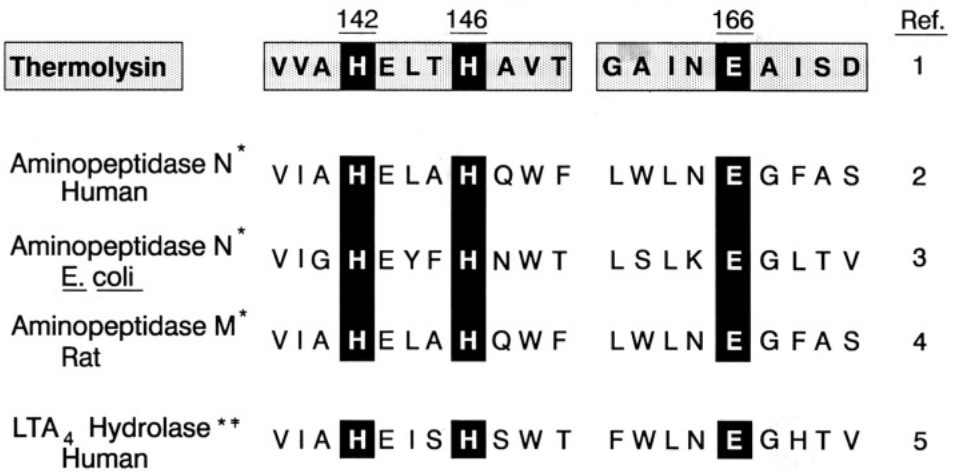

FIGURE 6: Putative zinc ligands of aminopeptidases. For ‡, see footnote 4. For key to figures, see Figure 1. References: (1) see Figure 5; (2) Olsen, J., Cowell, G. M., Kønigshøfer, E., Danielsen, E. M., Møller, J., Laustsen, L., Hansen, O. C., Welinder, K. G., Engberg, J., Hunziker, W., Spiess, M., Sjöström, H., & Norēn, O. (1988) *FEBS Lett. 238*, 307–314; (3) McCaman, M. T., & Gabe, J. D. (1986) *Gene 48*, 145–153; Foglino, M., Gharbi, S., & Lazdunski, A. (1986) *Gene 49*, 303–309; (4) Watt, V. M., & Yip, C. C. (1989) *J. Biol. Chem. 264*, 5480–5487; Malfroy, B., Kado-Fong, H., Gros, C., Giros, B., Schwartz, J.-C., & Hellmiss, R. (1989) *Biochem. Biophys. Res. Commun. 161*, 236–241; (5) Funk, C. D., Rådmark, O., Fu, J. Y., Matsumoto, T., Jörnvall, H., Shimizu, T., & Samuelsson, B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6677–6681.

of the family of enzymes can render a final judgment.

(*a*) *Metalloexopeptidases*. Aminopeptidases (EC 3.4.11.2) catalyze the hydrolysis of N-terminal amino acid residues of proteins, peptides, and amino acid amides. Several have been isolated from a wide range of tissues and bacteria and fall into two categories, but structure determinations are not on record. Human liver aminopeptidase (Garner & Behal, 1974), aminopeptidase M (Wacker et al., 1971), two bacterial enzymes, from *Bacillus licheniformis* (Rodriguez-Absi & Prescott, 1978) and *B. subtilis* (Wagner et al., 1979), and that from porcine kidney (Van Wart & Lin, 1981) each contain one zinc atom. In contrast, the bovine lens (Thompson & Carpenter, 1976) and Aeromonas aminopeptidases (Prescott et al., 1985) each contain two zinc atoms. In all of these, zinc is essential for catalytic activity.

Human intestinal aminopeptidase contains 967 amino acids (Olsen et al., 1988) and, in a domain of ~300 amino acids, is remarkably similar both to an equivalent region in the *Escherichia coli* aminopeptidase N (McCaman & Gabe, 1986; Foglino et al., 1986) and to rat kidney aminopeptidase M (Watt & Yip, 1989; Malfroy et al., 1989). In particular, a short segment in these domains contains two histidines and a glutamic acid in a linear arrangement that corresponds closely to the active zinc-binding site of thermolysin (Figures 5 and 6). If this comparison to thermolysin were to prove valid, the short spacer between His-388 (L_1) and His-392 (L_2) for the intestinal aminopeptidase would consist of 3 amino acids, identical with that of thermolysin, and the long spacer between His-392 (L_2) and Glu-411 (L_3) would be 18 instead of 19 amino acids (Vallee & Auld, 1989). If correct, this seemingly would be the first instance in which comparison of sequence identities and order of active-site zinc ligands gained from structure analysis of one enzyme family serves to predict that of another.[4]

(*b*) *Metalloendopeptidases*. The bacterial collagenases (EC 3.4.24.7) and neutral proteinases share similar zinc and calcium contents and pH activity optima, but their substrate specificity differs markedly and that of the collagenases is unique. In higher vertebrates, collagen is the most abundant protein, accounting for about a third of the total, but the triple helix of collagen renders it inert to most proteinases, including the neutral proteases. Rearrangement, synthesis, and degradation of connective tissues, as occurs in growth and development, arthritis, emphysema, lupus, tumor metastasis, osteomalacia, wound healing, bone resorption, uterine involution, etc., have all been said to involve the action of collagenases at some stage.

The purification and characterization of six collagenases from *Clostridium histolyticum* finally established them as zinc enzymes, confirming earlier speculations. Their molecular weights range from 68 000 to 125 000 (Bond & Van Wart, 1984a). All six enzymes (α, β, γ, δ, ϵ, and ζ) contain from 0.8 to 1.1 mol of zinc/mol of monomeric protein, and the calcium content varies from 1.9 to 6.8 mol/mol of protein (Bond & Van Wart, 1984b).

In one particular domain of the sequences of all known collagenases, transins, stromelysins, and human pump 1 proteinases, a short spacer of 3 amino acids separates two histidines from one another (Figure 7). In all matrix metalloproteases, a Glu is conserved juxtaposed to the potential equivalent of His-142 (L_1) of thermolysin. They have been suggested to correspond to those of the zinc-binding site of thermolysin (Figure 5) (Birkedal-Hansen, 1990, and references cited therein). The similarity of the three residues preceding and succeeding L_1 and L_2 lends weight to such a deduction, much as it does not prove it (Figure 7).

In the thermolysin family, His-146 (L_2) is separated from Glu-166 (L_3) by a 19 residue long spacer (Figure 5). However, in the matrix metalloproteinases there is no Glu at that position or anywhere near it (Figure 7). Among the other two known active-site ligands, histidines are 5, 19, 32, and 47 amino acids removed from the nearest proposed L_1 (His-218) or L_2 (His-222), and conserved aspartic acids are found after spacers of 15, 17, 21, 31, 32, 40, 45, 57, 84, and 91 amino acids, all in conserved sequences. Thus, the location and identity of L_3 in the matrix metalloproteases will remain speculative until the structure of one of the members of this family can serve as a standard of reference for the active-site zinc ligands of the others. It is not even a foregone conclusion that the two

[4] Leukotriene A_4 hydrolase, whose specificity is completely different, exhibits 20% sequence identity with aminopeptidase N (Malfroy et al., 1989). Its potential zinc binding site displays remarkable similarity to that of the aminopeptidases (Figure 6). However, neither the metal content nor the esterase or peptidase activity of this enzyme has been reported.

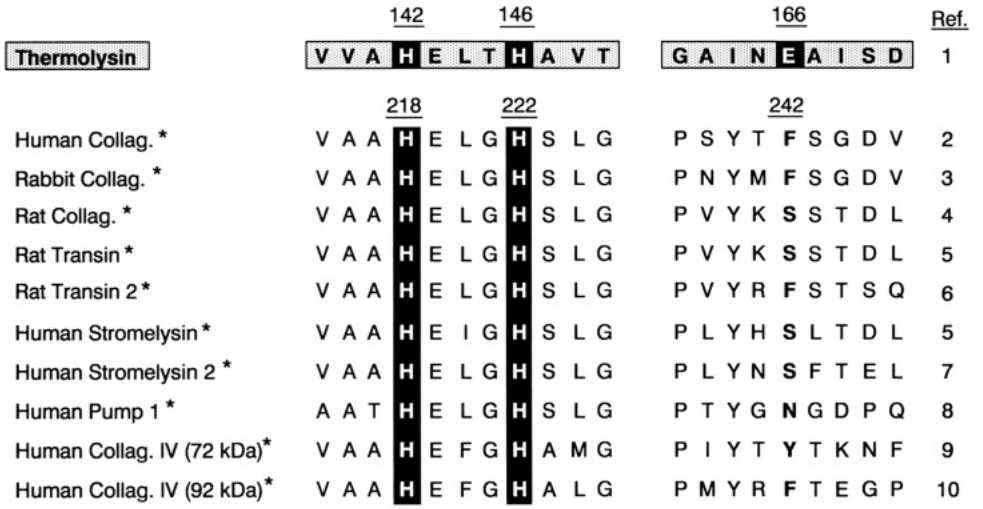

FIGURE 7: Putative zinc ligands of collagenase. For key to figures, see Figure 1. The second sequence begins at 16 amino acids beyond the putative second zinc ligand. References: (1) see Figure 5; (2) Goldberg, G. I., Wilhelm, S. M., Kronberger, A., Bauer, E. A., Grant, G. A., & Eisen, A. Z. (1986) *J. Biol. Chem. 261*, 6600–6605; (3) Fini, M. E., Plucinska, I. M., Mayer, A. S., Gross, R. H., & Brinckerhoff, C. E. (1987) *Biochemistry 26*, 6156–6165; (4) Matrisian, L. M., Glaichenhaus, N., Gesnel, M. C., & Breathnach, R. (1985) *EMBO J. 4*, 1435–1440; (5) Whitham, S. E., Murphy, G., Angel, P., Rahmsdorf, H. J., Smith, B. J., Lyons, A., Harris, T. J. R., Reynolds, J. J., Herrlich, P., & Docherty, A. J. P. (1986) *Biochem. J. 240*, 913–916; (6) Breathnach, R., Matrisian, L. M., Gesnel, M. C., Staub, A., & Leroy, P. (1987) *Nucleic Acids Res. 15*, 1139–1151; (7) Muller, D., Quantin, B., Gesnel, M. C., Millon-Collard, R., Abecossis, J., & Breathnach, R. (1988) *Biochem. J. 253*, 187–192; (8) Quantin, B., Murphy, G., & Breathnach, R. (1989) *Biochemistry 28*, 5327–5334; (9) Collier, I. E., Wilhelm, S. M., Eisen, A. Z., Marmer, B. L., Grant, G. A., Seltzer, J. L., Kronberger, A., He, C., Bauer, E. A., & Goldberg, G. I. (1988) *J. Biol. Chem. 263*, 6579–6587; (10) Wilhelm, S. M., Collier, I. E., Marmer, B. L., Eisen, A. Z., Grant, G. A., & Goldberg, G. I. (1989) *J. Biol. Chem. 264*, 17213–17221.

particular His residues that have been thought to be involved in binding zinc are, indeed, those that do. Clearly there are several additional His residues that might be the ones; in addition, there are Asp residues that could be involved as well.

(E) "Velcro" Mechanism of the Activation of the Matrix Prometalloproteinases

The metalloproteinases (Figure 7) that catalyze the hydrolysis of the major components of the extracellular matrix are synthesized as zymogen-like inactive or latent precursors and are converted subsequently to the active form. They are another example of an ever-growing number of physiological processes initiated by selective enzymatic cleavage of peptide bonds in enzymes and hormone precursors, vasoactive products, proteins involved in growth and development, blood coagulation, fibrinolysis, digestion, complement activation, and yet others (Neurath, 1986, 1989). Linderstrøm-Lang (1952) proposed long ago that limited proteolysis proceeds by either a "one-by-one" or a "zipper" mechanism.

The activation of procollagenases and progelatinases proceeds by a different mechanism, which we choose to call the "Velcro" mechanism. Their propeptides contain a solitary cysteine residue at position 92 (fibroblast collagenase numbering) (Figure 8), in a highly conserved region, PRCGVPDV (Witham et al., 1986; Sanchez-Lopez et al., 1988). It seems established that this forms a mercaptide with the sole zinc atom of what will become the mature enzyme. A number of studies on fibroblast procollagenase show that it can be activated by trypsin (Vaes, 1972; Birkedal-Hansen, 1976; Stricklin et al., 1983), organomercurials (Werb & Burleigh, 1974; Sellers et al., 1977; Grant et al., 1987), salts such as NaI and NaSCN (Shinkai & Nagai, 1977), detergents (Birkedal-Hansen & Taylor, 1982), and thiol exchange reactions (Macartney & Tschesche, 1983; Springman et al., 1990). These results suggest that it is the dissociation and/or displacement of that cysteine from the zinc atom that results in activity by zymogen activation (Springman et al., 1990). The cysteine, sticking to the zinc atom through its SH group and acting like Velcro,

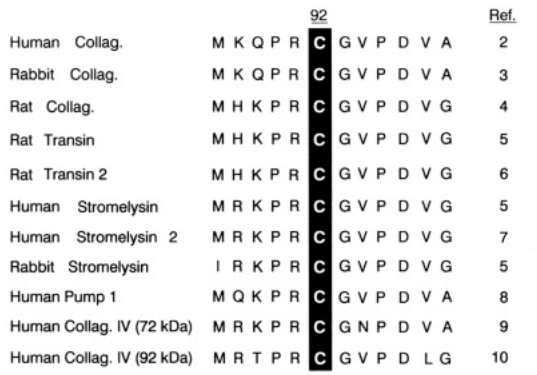

FIGURE 8: Propeptide of matrix metalloproteinase precursors. For key to figures, see Figure 1. For references, see Figure 7.

prevents the zinc atom from becoming enzymatically active until the cysteine is removed.

While the details of the process still require definition, there seems no doubt that the principal chemical event which induces activity is the removal of the SH ligand of cysteine from the zinc (Figure 9). This is seemingly the first instance in which a fully coordinated, tetradentate (i.e., structural type) zinc atom is converted into a tridentate (i.e., enzymatically functional) one through the displacement of one ligand, i.e., cysteine, which is then replaced by water. However, the identity of the remainder of the ligands is not yet known (see above).

The dissociation of the zinc monothiolate complex of the inactive matrix prometalloproteinases into its constituents, i.e., cysteine plus a tridentate zinc complex of the active enzyme forms, represents a new activation mechanism that is based on zinc coordination chemistry. Since activated water is only found in relation to tridentate sites, while the cysteine is coordinated to the zinc, the protein will be inactive. The sole cysteinyl residue of the activation peptide apparently blocks the active-site zinc atom, thereby preventing its participation in catalysis. Its removal through physiological or pathological processes constitutes the activation process, allowing the entry of H_2O or substrate. This mechanism, hitherto unknown,

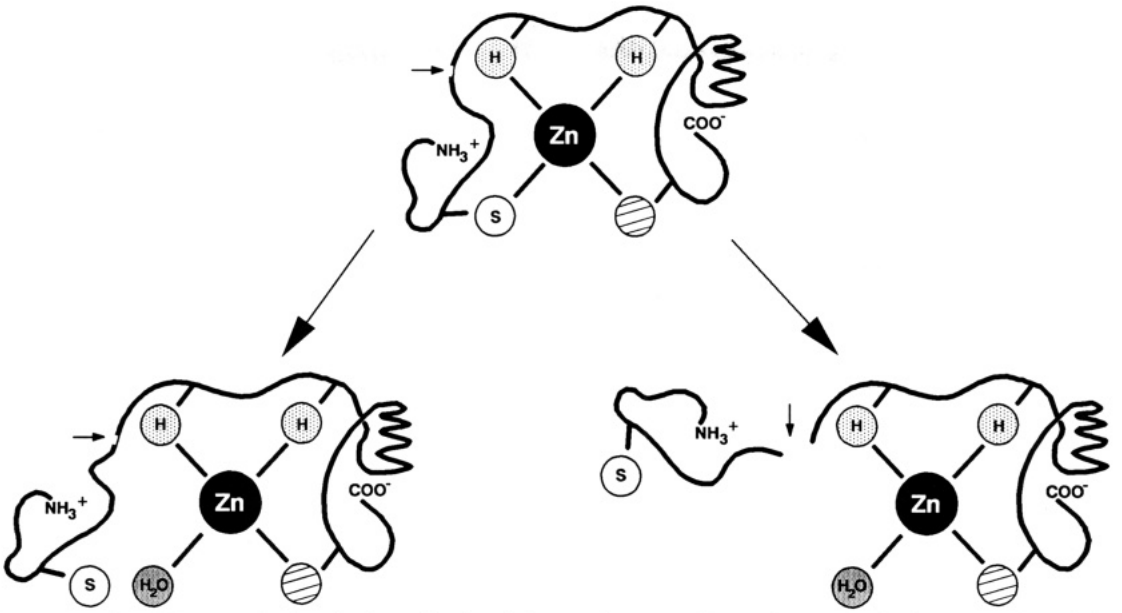

FIGURE 9: Schematic of the Velcro mechanism for the activation of the matrix prometalloproteinases. Activation occurs through the conversion of the zinc coordinated to four amino acid residues (tetradentate) into one that is coordinated to only three residues (tridentate) through the removal of one ligand, cysteine, S, which is then replaced by water. The displacement of cysteine is induced either by proteolytic cleavage and/or by conformational changes of the propeptide. The small arrow indicates the peptide region of the presumably autocatalytic cleavage site. Ligands to the functional zinc are thought to be two histidines, H; a third ligand is as yet unidentified and is symbolized by the crosshatched circle.

apparently represents yet another example of the versatility of zinc chemistry in biological reactions.

(F) Zinc Coordination Chemistry Relevant to the Structure and Function of Zinc Enzymes

Cysteine residues of zinc enzymes form tetradentate zinc complexes with very high stability constants, which ensure both overall structures and local conformations akin to those provided by disulfides. Tridentate combinations of histidine, glutamic and aspartic acid, and cysteine side chains bind zinc firmly and have proved to be characteristic of catalytic function by providing coordination sites open to water and/or substrate complexes and their transition-state intermediates. Which features of zinc chemistry may account for these findings?

The inherent chemical potential and reactivities of zinc are not exceptional. Oxidoreduction, characteristic of the neighboring transition elements, is a major source of changes in coordination geometries, rates of ligand substitution, and amphoteric properties of these elements. Importantly, however, zinc is both stable and inert to oxidoreduction. Indeed, it has been emphasized that generally the divalent state is more stable than are higher oxidation states (Cotton & Wilkinson, 1988).

The lack of redox changes makes zinc stable in a biological medium whose potential is in constant flux. Furthermore, some of its physical–chemical qualities present important advantages in biology. Zinc is amphoteric and exists in both metal hydrate and hydroxide forms at pH values near neutrality. The coordination sphere has proven exceptionally flexible. The stereochemical adaptability of zinc coordination complexes in enzymes is unusual and constitutes one of the striking features of its coordination complexes; the multiplicity of coordination numbers and geometries denotes that zinc submits readily to the demands of its ligands. In fact, it is through these properties that proteins and other biological macromolecules alter the reactivities of zinc. Thus, proteins and enzymes affect the chemistry of zinc, much as it, in turn, adapts to these macromolecules. Collectively, these physicochemical features are important means for the translation of chemical structure into multiple biological functions. Zinc thereby becomes a versatile interactant for different donor groups of varying ligand types resulting in a broad range of stability constants, reactivities, and functions.

(G) Zinc Thiolate Clusters: Metallothioneins

Zinc enzymes and metallothionein were discovered almost contemporaneously. While efforts to understand the role of zinc in enzyme mechanisms rapidly became an important biochemical topic, the existence of metallothionein went virtually unnoticed for many years, and its precise function(s) remain(s) unknown to the present. It required 25 years of work to establish its structure, which revealed a remarkable coordination complex that may well prove as important to an extension of zinc coordination chemistry itself and to its implications for biochemistry as turned out to be the case for the examination of zinc enzymes in relation to the mechanism of enzyme action.

Metallothionein, first isolated from equine kidney cortex in 1957 (Margoshes & Vallee, 1957), has been found in the animal kingdom wherever it has been sought (Kägi & Kojima, 1987). It has a molecular weight of 6700 and is composed of 62 (or 61) amino acids, including 20 cysteines, but cystine and heterocyclic and aromatic amino acids are absent. It contains 7 mol of zinc and/or cadmium/mol of protein.[5] In some instances copper, iron, or mercury has also been detected.

In multiple species, a series of metals, hormones, and other organic molecules induce its formation, but the functions of metallothionein itself remain unknown (Kägi & Kojima, 1987; Kägi & Schäffer, 1988). In native mammalian metallothionein, the 7 mol of zinc and/or cadmium/mol of protein, together with the presence of 20 cysteinyl residues, absorption maxima at 215 and 248 mm, and the resultant CD and MCD

[5] Plants and fungi contain cadmium(II) thiolate polypeptides with the primary structure $(\gamma\text{-Glu-Cys})_n$ or $(\gamma\text{-Glu-Cys})_n\text{-}\beta\text{-Ala}$, where $n = 2$–11. They have variously been given trivial names, i.e., phytochelatins, homophytochelatins, or cadystins. These polypeptides are mentioned to complete the inventory of Zn and Cd chemistry in biology even though native zinc complexes of this type have not been reported thus far (Rauser, 1990).

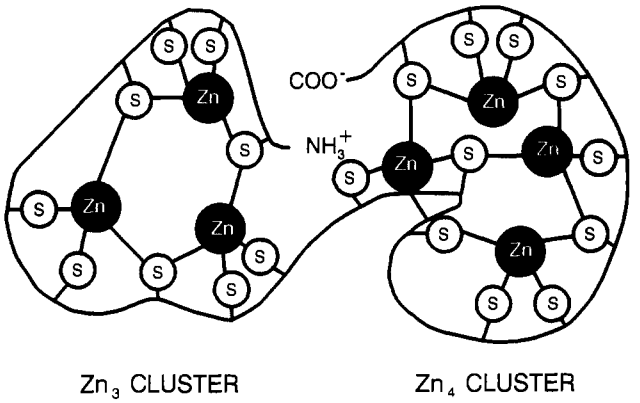

FIGURE 10: Zinc thiolate clusters of metallothioneins (Kägi & Kojima, 1987).

spectra are consistent with an average ratio of approximately 3 SH/Me atom. This was thought to reflect a negatively charged complex owing to the formation of three or four mercaptide bonds (Kägi & Vallee, 1961). Subsequent structural studies of metallothionein by ^{113}Cd NMR and 2D ^{1}H NMR and X-ray diffraction have demonstrated that cadmium and zinc are the nuclei of metal clusters in metallothioneins (Otvos & Armitage, 1980; Braun et al., 1986; Furey et al., 1986; Kägi & Schäffer, 1988).

The seven zinc atoms form two distinct and separate metal clusters, Zn_4Cys_{11} (residues 33–60) and Zn_3Cys_9 (residues 5–29), with five and three cysteine residues acting as bridging ligands between two metal ions in each cluster, respectively (Figure 10). Such zinc thiolate clusters had neither been anticipated inductively nor could they have been deduced, since at that time neither complex ions nor proteins were known to form them. They are reminiscent of the ferredoxin, rubredoxin, or molybdenum clusters, much as those participate in oxidoreductive reactions. Awareness of the existence of zinc and cadmium clusters is timely given the structural role of zinc in some zinc enzymes and its role in the expression and transcription of the genetic message. Whatever the function(s) of metallothionein will turn out to be, it (they) will likely involve its interactions with other molecules. Hence, this zinc and cadmium thiolate cluster structure, until now unique to metallothionein, may ultimately be found to telltale specific metabolic functions. This view is supported by the fact that the zinc thiolate cluster structure has most recently been observed in DNA-binding proteins (see below). Furthermore, discovery of such a structure in a biochemical system has provided the impetus to attempt to design, synthesize, and inspect the properties of simpler inorganic adamantane structures formed with the phenylthiols, $Zn_4(SR)_{10}$ (Dance, 1981; Hencher et al., 1981; Hagen et al., 1982).

(H) Zinc in the Control of Gene Expression

(1) Xenopus Transcription Factor IIIA. Evidence that zinc is crucial to DNA and RNA synthesis and cell division emerged in the 1970s (Vallee, 1977a,b; Auld, 1979; Vallee & Falchuk, 1981). It was reinforced by Wu's demonstration that transcription factor IIIA (TFIIIA), which activates the transcription of the *Xenopus* 5S RNA gene, is a zinc protein containing from 2 to 3 mol of zinc/mol of protein (Hanas et al., 1983). Subsequently, TFIIIA was shown to consist of 9 repeat sequences of about 30 amino acids, each containing 2 Cys residues separated by a short spacer of 2–5 amino acid residues and 2 His residues separated by two short spacers of 3–4 amino acid residues. A long spacer between the putative inner Cys and His ligands consists of 12 amino acid residues (Ginsberg et al., 1984; Brown et al., 1985). Each 7S particle was reported to contain 7–11 zinc atoms/mol (Miller et al., 1985).

It was proposed that 2 Cys and 2 His residues of each repeat unit form a tetrahedral coordination complex with each of nine zinc atoms, thereby generating a peptide domain postulated to interact with DNA. The results of limited proteolytic degradation, EXAFS measurements (Miller et al., 1985), organic synthesis of small peptide domains and NMR of their zinc complexes (Lee et al., 1989), and absorption spectroscopy of their cobalt complexes (Frankel et al., 1987; Green & Berg, 1989) were thought to be consistent with the above findings and interpretation of their functional significance.

However, a very recent extension of the earlier studies renewedly claims and affirms that the isolated protein and its RNA complex (the 7S particle) contain only two—not nine—firmly bound, intrinsic zinc atoms essential to the transcription factor IIIA activation of the transcription of the 5S RNA gene. No evidence for the presence of additional loosely bound zinc ions was found. Addition of zinc to the protein containing only the two intrinsic zinc atoms enhanced neither the affinity of protein binding to DNA or 5S RNA nor transcription activation (Shang et al., 1989).

(2) Other DNA-Binding Putative Zinc Proteins. Subsequent to the initial reports regarding zinc in TFIIIA, a large number and variety of sequences of DNA-binding proteins were shown to contain Cys and His residues thought to be analogous and separated by short and long intervening amino acid spacer sequences (Rosenberg et al., 1986; Vincent et al., 1985; Hartshorne et al., 1986; Schuh et al., 1986; Chowdhury et al., 1987). They were identified by computer searches intended to single out proteins, loosely defined as not homologous with but containing TFIIIA-like sequences, that bind to nucleic acids. It was inferred further that these reflected putative metal-binding domains which might participate in DNA binding and, hence, gene regulation. Such searches have now revealed well in excess of 150 such proteins, suspected—but not shown—to contain zinc. The initial postulate was that one pair each of spatially juxtaposed Cys and His ligands in such sequences, each of them separated by short amino acid spacers, would prove characteristic of zinc sites in the domains of DNA-binding proteins. On the basis of analytical findings demonstrating the presence of zinc in the glucocorticoid receptor protein (Freedman et al., 1988), this proposition was subsequently amended to accommodate an alternative ligand arrangement consisting of two pairs of Cys zinc ligands but no His. A total of six classes of proteins with putative metal-binding sites and participating in binding to nucleic acids have been enumerated: (1) low molecular weight nucleic acid binding or gene regulatory proteins, (2) adenovirus E1A gene products, (3) aminoacyl-tRNA synthetases, (4) large T antigens, (5) bacteriophage proteins, and (6) hormone receptors (Berg, 1986, 1988). Additional computer searches have resulted in an increasing number of such sequences featuring cysteine and histidine residues in locations analogous to and homologous with those stipulated above (Sunderman & Barber, 1988).

In these, the evidence for the presence of zinc-binding sites and their relevance to these hypotheses is quite variable but may imply that the "putative" zinc-binding sites are tantamount to both the presence and a biological function of zinc in a given instance. A large number of such studies ignored the fact that once proteins which contain putative metal-binding domains are recognized, the presence, mode of binding, and role of zinc ions in such systems must still be verified. In some instances, the induction or enhancement of an activity

Table II: Transcription Proteins Established To Contain Zinc by Quantitative Analysis

protein	MW × 10^{-3}	Zn (mol/mol of protein)	ligands	ref[a]
TFIIIA	40	2	2 Zn { 4 Cys, 4 His	1
TFIIIA	40	7–11	9 Zn { 18 Cys, 18 His	2
Glu Rec	19	2	2 Zn { 9 Cys, 1 His	3
GAL4	17	2	2 Zn { 6 Cys	4
g32P	35	1	1 Zn { 3 Cys, 1 His	5

[a] References: (1) Hanas et al., 1983; Shang et al., 1989; (2) Miller et al., 1985; (3) Freedman et al., 1988; (4) Hollenberg et al., 1987, and Pan & Coleman, 1989, 1990; (5) Giedroc et al., 1986.

by addition of Zn^{2+} ions and/or its diminution or abolition by chelating agents are cited to either confirm or prove this. However, in the past, the results of such experiments have not proven to be reliable criteria for either the presence or potential function of zinc or other metals (Vallee & Wacker, 1970; Vallee & Auld, 1990a).

Considering the vast number of articles whose titles refer to "zinc fingers", it is important to realize that the presence of zinc has been confirmed analytically in only four instances (Table II). Moreover, in each of those, the combination and identity of the putative zinc ligands differ. The 2 Cys and 2 His of TFIIIA have been mentioned already. In the glucocorticoid receptor, there are 9 Cys and 1 His together with at least two zinc atoms (Freedman et al., 1988). The transcription factor of GAL4 contains 6 Cys and two zinc atoms (Hollenberg et al., 1987; Pan & Coleman, 1989), and the gene 32 protein required for DNA replication in bacteriophage T4 contains one zinc as well as 3 Cys and 1 His ligands (Giedroc et al., 1986). Thus, in these DNA-binding proteins where the presence of zinc is documented, the relevant ligands are very variable and seemingly not predictable on the basis of the hypotheses that have been suggested (Berg, 1986). Yet other combinations of ligands in such binding sites may exist, but first and foremost, the presence of zinc should be established for each. X-ray crystallographic or NMR data, once available as structural standards of reference for members of these protein families, will no doubt answer most of the questions yet remaining.

(*3*) *GAL4*: *Zinc Thiolate Cluster*. The DNA-binding domain of the transcription factor GAL4 consists of the 62 N-terminal residues, denoted by GAL4(62*). Its study by Cd NMR was timely (Pan & Coleman, 1990). ^{1}H–^{113}Cd heteronuclear multiple quantum NMR spectroscopy and phase-sensitive double-quantum-filtered ^{1}H COSY of the ^{112}Cd- and ^{113}Cd-substituted GAL4(62*) derivatives provide direct evidence that the two bound ^{113}Cd ions are coordinated by *six* cysteines, two of which form bridging ligands between the ^{113}Cd ions (Figure 11). The overall arrangement is that of a zinc thiolate cluster structure akin to that of metallothionein (Kägi & Kojima, 1987), not that of a zinc finger. The highly conserved arrangement of Cys in GAL4 and other fungal transcription factors almost certainly predicts the occurrence of such binuclear zinc clusters in other similar DNA-binding proteins (Pan & Coleman, 1990).

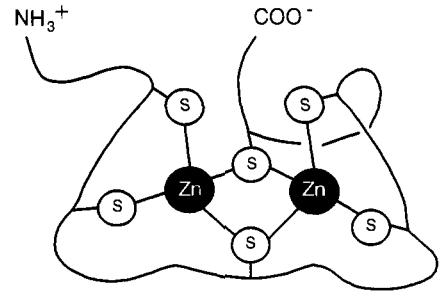

FIGURE 11: Zinc thiolate cluster of GAL4 (Pan & Coleman, 1990).

SUMMARY

The last decade has seen remarkable developments in establishing zinc-dependent interrelationships between enzyme activities, protein structures and folding, and products of gene expression. Different permutations of zinc chemistry seem pivotal both for the expression of catalytic function of zinc enzymes and their local structure and for DNA-binding proteins and their interactions that result in gene expression. Conversely, the structure, conformation, and folding of the protein determine the response of the metal. This reciprocity expresses itself in particular through the number and nature of the protein ligands forming the coordination complex. Tridentate zinc sites (plus activated water) are characteristic of catalysis. In enzymes, tetradentate zinc sites are inaccessible to solvent, critically controlling local protein folding, structure, and conformation. The extent to which they may regulate protein–protein and protein–DNA interactions remains to be defined. In enzymes, histidines predominate in catalytic zinc sites while cysteines are the exclusive zinc ligands in structural sites. Tetradentate zinc coordination in procollagenase is convertible and becomes tridentate in collagenase through the removal of the cysteine ligand of the propeptide from the zinc atom. This conversion yields a catalytically active zinc enzyme site in which a water molecule replaces the cysteine ligand.

Biology has availed itself of zinc coordination chemistry for specific biological objectives. The zinc clusters of metallothionein seem to go beyond this, representing an example of "inorganic natural products chemistry", since such a zinc structure had never been observed in nonbiological zinc chemistry prior to discovery of metallothionein. While the function of metallothionein is still unknown, its zinc thiolate cluster, a structural motif unique to biology, is perhaps trying to telltale function. Ultimately such zinc cluster structures, if proven to exist widely, might, in fact, be characteristic of and synonymous with functions yet to emerge—much as is the case for active zinc site enzymes. The zinc thiolate cluster of the transcription factor GAL4 may just be the first of its kind in DNA-binding proteins.

REFERENCES

Auld, D. S. (1979) *Adv. Chem. Ser. 172*, 112–133.

Auld, D. S. (1987) in *Enzyme Mechanisms* (Page, M. I., Ed.) Royal Society of Chemistry, Letchworth, U.K.

Auld, D. S., & Vallee, B. L. (1987) in *Hydrolytic Enzymes* (Neuberger, A., & Brocklehurst, K., Eds.) pp 201–255, Elsevier, New York.

Benchetrit, T., Bissery, V., Mornon, J. P., Devault, A., Crine, P., & Roques, B. P. (1988) *Biochemistry 27*, 592–596.

Berg, J. M. (1986) *Science 232*, 485–487.

Berg, J. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 99–102.

Beuer, R. A., & Iglewski, B. H. (1988) *J. Bacteriol. 170*, 4309–4314.

Birkedal-Hansen, H. (1990) in *Matrix Metalloproteinases and Inhibitors* (Birkedal-Hansen, H., Werb, Z., Welgus, H., &

Van Wart, H., Eds.) Gustav Fischer Verlag, Stuttgart, West Germany (in press).
Birkedal-Hansen, H., & Taylor, R. E. (1982) *Biochem. Biophys. Res. Commun. 107*, 1173–1178.
Birkedal-Hansen, H., Cobb, C. M., Taylor, R. E., & Fullmer, H. W. (1976) *J. Biol. Chem. 251*, 3162–3168.
Bond, M. D., & Van Wart, H. E. (1984a) *Biochemistry 23*, 3077–3085.
Bond, M. D., & Van Wart, H. E. (1984b) *Biochemistry 23*, 3085–3091.
Brandén, C. I., Jörnvall, H., Eklund, M., & Furugren, B. (1975) in *Enzymes, 3rd Ed.* (Boyer, P. D., Ed.) Vol. 11, p 103, Academic Press, New York.
Braun, W., Wagner, G., Wörgötter, E., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 125–129.
Brown, R. S., Sander, C., & Argos, P. (1985) *FEBS Lett. 186*, 271–274.
Chowdhury, K., Deutsch, U., & Gruss, P. (1987) *Cell 48*, 771–778.
Cotton, F. A., & Wilkinson, G. (1988) in *Advanced Inorganic Chemistry*, 5th ed., p 599, Wiley-Interscience, New York.
Dance, I. G. (1981) *Inorg. Chem. 20*, 2155.
Devault, A., Lazure, C., Nault, C., Le Moual, H., Seidah, N. G., Chrétien, M., Kahn, P., Powell, J., Mallet, J., Beaumont, A., Roques, B. P., Crine, P., & Boileau, G. (1987) *EMBO J. 6*, 1317–1322.
Devault, A., Nault, C., Zollinger, M., Fournie-Zaluski, M.-C., Roques, B. P., Crine, P., & Boileau, G. (1988) *J. Biol. Chem. 263*, 4033–4040.
Dideberg, O., Charlier, P., Dive, G., Joris, B., Frère, J. M., & Ghuysen, J. M. (1982) *Nature 299*, 469–470.
Eklund, H., & Brandén, C.-I. (1983) in *Zinc Enzymes* (Spiro, T. G., Ed.) pp 123–152, Wiley-Interscience, New York.
Foglino, M., Gharbi, S., & Lazdunski, A. (1986) *Gene 49*, 303–309.
Frankel, A. D., Berg, J. M., & Pabo, C. O. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4841–4845.
Freedman, L. P., Luisi, B. F., Korszun, Z. R., Basavappa, R., Sigler, P. B., & Yamamoto, K. R. (1988) *Nature 334*, 543–546.
Fricker, L. D., Evans, C. J., Esch, F. S., & Herbert, E. (1986) *Nature 323*, 461–464.
Furey, W. F., Robbins, A. H., Clancy, L. L., Winge, D. R., Wang, B. C., & Stout, C. D. (1986) *Science 231*, 704–710.
Garner, C. W., Jr., & Behal, F. J. (1974) *Biochemistry 13*, 3227–3233.
Gebhard, W., Schube, M., & Eulitz, M. (1989) *Eur. J. Biochem. 178*, 603–607.
Giedroc, D. P., Keating, K. M., Williams, K. R., Konigsberg, W. H., & Coleman, J. E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8452–8456.
Ginsberg, A. M., King, B. O., & Roeder, R. G. (1984) *Cell 39*, 479.
Grant, G. A., Eisen, A. Z., Marmer, B. L., Roswit, W. T., & Goldberg, G. I. (1987) *J. Biol. Chem. 262*, 5886–5889.
Green, L. M., & Berg, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4047–4051.
Hagen, K. S., Stephan, D. W., & Holm, R. H. (1982) *Inorg. Chem. 21*, 3928.
Hanas, J. S., Hazuda, D. J., Bogenhagen, D. F., Wu, F. Y.-H., & Wu, C.-W. (1983) *J. Biol. Chem. 258*, 14120–14125.
Hartshorne, T. A., Blumberg, H., & Young, E. T. (1986) *Nature 320*, 283–287.
Hencher, J. L., Khan, M., Said, F. F., & Tuck, D. G. (1981) *Inorg. Nucl. Chem. Lett. 17*, 187.
Hollenberg, S. M., Giguère, V., Segui, P., & Evans, R. M. (1987) *Cell 49*, 39–46.
Holmquist, B., & Vallee, B. L. (1974) *J. Biol. Chem. 249*, 4601–4607.
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. (1982) *J. Mol. Biol. 160*, 219–263.
Hough, E., Hansen, L. K., Birknes, B., Jynge, K., Hansen, S., Horvik, A., Little, C., Dodson, E., & Derewenda, Z. (1989) *Nature 338*, 357–360.
Jongeneel, C. V., Bouvier, J., & Bairoch, A. (1989) *FEBS Lett. 242*, 211–214.
Kägi, J. H. R., & Vallee, B. L. (1961) *J. Biol. Chem. 236*, 2435–2442.
Kägi, J. H. R., & Kojima, Y., Eds. (1987) *Metallothionein II*, Proceedings of the Second International Meeting, Zurich, Switzerland, Birkhäuser Verlag, Basel, Switzerland.
Kägi, J. H. R., & Schäffer, A. (1988) *Biochemistry 27*, 8509–8515.
Kannan, K. K., Notstrand, B., Fridborg, K., Lövgren, S., Orlsson, A., & Petef, M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 51–55.
Keating, K. M., Ghosaini, L. R., Giedroc, D. P., Williams, K. R., Coleman, J. E., & Sturtevant, J. M. (1988) *Biochemistry 27*, 5240–5245.
Kerr, M. A., & Kenny, A. J. (1974) *Biochem. J. 137*, 489–495.
Kim, E. E., & Wyckoff, H. W. (1989) *Clin. Chim. Acta 186*, 175–188.
Klug, A., & Rhodes, D. (1987) *Trends Biochem. Sci. 12*, 464–469.
Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., & Wright, P. E. (1989) *Science 245*, 635–637.
Le Moual, H., Crine, P., & Boileau, G. (1989) in *Abstract of Matrix Metalloproteinase Conference* (Birkedal-Hansen, H., Werb, Z., Welgus, H., & Van Wart, H., Eds.) Destin, FL.
Liljas, A., Kannan, K. K., Bergstén, P. C., Waara, I., Fridborg, B., Strandberg, B., Carlbom, U., Järup, L., Lövgren, S., & Petef, M. (1972) *Nature, New Biol. 235*, 131–137.
Linderstrøm-Lang, K. (1952) in Lane Medical Lectures, *Stanford Univ. Publ., Univ. Ser., Med. Sci. 6*, 1–115.
Lipscomb, W. N., Reeke, G. N., Hartsuck, J. A., Quiocho, F. A., & Bethge, P. H. (1970) *Philos. Trans. R. Soc. London, B 257*, 177–214.
Macartney, H. W., & Tschesche, H. (1983) *Eur. J. Biochem. 130*, 71–78.
Malfroy, B., Schofield, P. R., Kuang, W.-J., Seeburg, P. H., Mason, A. J., & Henzel, W. J. (1987) *Biochem. Biophys. Res. Commun. 102*, 206–214.
Malfroy, B., Kado-Fong, H., Gros, C., Giros, B., Schwartz, J.-C., & Hellmiss, R. (1989) *Biochem. Biophys. Res. Commun. 161*, 236–241.
Margoshes, M., & Vallee, B. L. (1957) *J. Am. Chem. Soc. 79*, 4813.
Matthews, B. W., Jansonius, J. N., Colman, P. M., Schoenborn, B. P., & Dupourque, D. (1972) *Nature, New Biol. 238*, 37–41.
Matthews, B. W., Weaver, L. H., & Kester, W. R. (1974) *J. Biol. Chem. 249*, 8030–8044.
McCaman, M. T., & Gabe, J. D. (1986) *Gene 48*, 145–153.
Miller, J., McLachlan, A. D., & Klug, A. (1985) *EMBO J. 4*, 1609–1614.
Neurath, H. (1986) *Chem. Scr. 27B*, 221–229.

Neurath, H. (1989) *Trends Biochem. Sci. 14*, 268–271.
Olsen, J., Cowell, G. M., Kønigshøfer, E., Danielsen, E. M., Møller, J., Laustsen, L., Hansen, O. C., Welinder, K. G., Engberg, J., Hunziker, W., Spiess, M., Sjöström, H., & Norēn, O. (1988) *FEBS Lett. 238*, 307–314.
Otvos, J. D., & Armitage, I. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7094–7098.
Pan, T., & Coleman, J. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3145–3149.
Pan, T., & Coleman, J. E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2077–2081.
Pauptit, R. A., Karlsson, R., Picot, D., Jenkins, J. A., Niklaus-Reimer, A.-S., & Jansonius, J. N. (1988) *J. Mol. Biol. 199*, 525–537.
Prescott, J. M., Wagner, F. W., Holmquist, B., & Vallee, B. L. (1985) *Biochemistry 24*, 5350–5356.
Quiocho, F. A., & Lipscomb, W. N. (1971) *Adv. Protein Chem. 25*, 1–58.
Rauser, W. E. (1990) *Annu. Rev. Biochem. 59*, 61.
Rees, D. C., Lewis, M., & Lipscomb, W. N. (1983) *J. Mol. Biol. 168*, 367–387.
Riordan, J. F., & Vallee, B. L. (1988) *Methods Enzymol. 158*, 3–6.
Rodriguez-Absi, J., & Prescott, J. M. (1978) *Arch. Biochem. Biophys. 186*, 383–391.
Rosenberg, U. B., Schröder, C., Preiss, A., Kienlin, A., Côté, S., Riede, I., & Jäckle, H. (1986) *Nature 319*, 336–339.
Sanchez-Lopez, R., Nicholson, R., Gesnel, M.-C., Matrisian, L. M., & Breathnach, R. (1988) *J. Biol. Chem. 263*, 11892–11899.
Schmid, M. F., & Herriott, J. R. (1976) *J. Mol. Biol. 103*, 175–190.
Schuh, R., Aicher, W., Gaul, U., Côté, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., Kemler, R., & Jäckle, H. (1986) *Cell 47*, 1025.
Sellers, A., Cartwright, E., Murphy, G., & Reynolds, J. J. (1977) *Biochem. J. 163*, 303.
Shang, Z., Liao, Y.-D., Wu, F. Y.-H. & Wu, C.-W. (1989) *Biochemistry 28*, 9790–9795.
Shinkai, H., & Nagai, Y. (1977) *J. Biochem. (Tokyo) 81*, 1261–1268.
Springman, E. B., Angleton, E. L., Birkedal-Hansen, H., & Van Wart, H. W. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 364–368.
Stöcker, W., Wolz, R. L., Zwilling, R., Strydom, D. J., & Auld, D. S. (1988) *Biochemistry 27*, 5026–5032.
Stricklin, G. P., Jeffrey, J. J., Roswit, W. T., & Eisen, A. Z. (1983) *Biochemistry 22*, 61–68.
Sunderman, F. W., Jr., & Barber, A. M. (1988) *Ann. Clin. Lab. Sci. 18*, 267–288.
Sutton, B. J., Artymiuk, P. J., Cordero-Borboa, A. E., Little, C., Phillips, D. C., & Waley, S. G. (1987) *Biochem. J. 248*, 181–188.
Tan, F., Chan, S. J., Steiner, D. F., Schilling, J. W., & Skidgel, R. A. (1989) *J. Biol. Chem. 264*, 13165–13170.
Thompson, G. A., & Carpenter, F. H. (1976) *J. Biol. Chem. 251*, 53–60.
Vaes, G. (1972) *Biochem. J. 126*, 275–289.
Vallee, B. L. (1977a) *Experientia 33*, 600.
Vallee, B. L. (1977b) in *Biological Aspects of Inorganic Chemistry* (Dolphin, D., Ed.) pp 37–70, Wiley, New York.
Vallee, B. L. (1986) in *Zinc Enzymes* (Bertini, I., Luchinat, C., Maret, W., & Zeppezauer, M., Eds.) pp 1–15, Birkhäuser, Boston, MA.
Vallee, B. L. (1988) *Biofactors 1*, 31–36.
Vallee, B. L., & Wacker, W. E. C. (1970) *Proteins (2nd Ed.) 5*, 1–192.
Vallee, B. L., & Falchuk, K. F. (1981) *Philos. Trans. R. Soc. London, B 294*, 185–197.
Vallee, B. L., & Auld, D. S. (1989) *FEBS Lett. 257*, 138–140.
Vallee, B. L., & Auld, D. S. (1990a) in *Matrix Metalloproteinases and Inhibitors* (Birkedal-Hansen, H., Werb, Z., Welgus, H., & Van Wart, H., Eds.) Gustav Fischer Verlag, Stuttgart, West Germany (in press).
Vallee, B. L., & Auld, D. S. (1990b) *Proc. Natl. Acad. Sci. U.S.A. 87*, 220–224.
Vallee, B. L., Riordan, J. F., Auld, D. S., & Latt, S. A. (1970) *Philos. Trans. R. Soc. London, B 257*, 215–230.
Vallee, B. L., Galdes, A., Auld, D. S., & Riordan, J. F. (1983) in *Zinc Enzymes* (Spiro, T. G., Ed.) pp 26–75, Wiley, New York.
Van Wart, H. E., & Lin, S. H. (1981) *Biochemistry 20*, 5682–5689.
Vincent, A., Colot, H. V., & Rosbash, M. (1985) *J. Mol. Biol. 186*, 149–166.
Wacker, H., Lehky, P., Fischer, E. H., & Stein, E. H. (1971) *Helv. Chim. Acta 54*, 473–484.
Wagner, F. W., Ray, L. E., Ajabnoor, M., Ziemba, P. E., & Hall, R. L. (1979) *Arch. Biochem. Biophys. 197*, 63–72.
Watt, V. M., & Yip, C. C. (1989) *J. Biol. Chem. 264*, 5480–5487.
Werb, Z., & Burleigh, M. C. (1974) *Biochem. J. 137*, 373.
Witham, S. E., Murphy, G., Angel, P., Rahmsdorf, H. J., Smith, B. J., Lyons, A., Harris, T. J. R., Reynolds, J. J., Herrlich, P., & Docherty, A. J. P. (1986) *Biochem. J. 240*, 913–916.