Volume 257, number 1, 138–140 FEB 07735 October 1989

Short and long spacer sequences and other structural features of zinc binding sites in zinc enzymes

Bert L. Vallee and David S. Auld

Center for Biochemical and Biophysical Sciences and Medicine and Department of Pathology, Harvard Medical School, Brigham and *Women's Hospital, 250 Longwood Avenue, Boston, MA 02135, USA*

Received 25 August 1989

The crystal structures of eleven zinc enzymes have served to identify common features of their Zn binding sites. Two of them have non-catalytic Zn sites, both of which contain four cysteine ligands closely spaced in the linear sequence of the protein with no bound water. In contrast, all the catalytic Zn sites have three protein ligands and, in addition, one coordinated, 'activated' water. Histidine is the predominant ligand. The spacing between the first two ligands (1–3 amino acids), the short spacer, ensures a nucleus for Zn binding. The third ligand, separated by from ~20 to ~120 amino acids, the long spacer, not only completes the coordination but also aligns protein residues for interaction with the substrate. The short and long spacing observed for catalytic zinc sites may also pertain to Fe and Cu proteins.

Metalloenzyme; Metalloproteinase; Zinc; Amino acid spacer sequence

During the last forty years, zinc enzymes have been isolated and characterized in such large numbers that they are now recognized as a distinct category within enzymology. Zinc is the only metal which is essential to the function of at least one member of each of the six categories of enzymes established by the IUB [1].

The unambiguous identification of the metal-binding ligands and the details of their mode of coordination to zinc is the consequence of X-ray crystallographic analyses performed over the last two decades. Such analyses have provided structures for a total of eleven zinc enzymes representing four of the six enzyme classes. These serve as invaluable standards of reference for others in the same class or subclass with which they are homologous. Furthermore, they give insight into and help to define common structural features of the enzyme's catalytic and non-catalytic zinc binding sites. Table 1 lists the pertinent features for ten catalytic zinc binding sites.

Histidine is by far the most common ligand present at these catalytic sites. Thus, the hydrolases, carboxypeptidase A [2], thermolysin [3], *B. cereus* neutral protease [4], phospholipase C [5] and alkaline phosphatase [6] all have two histidine residues bound to the catalytic zinc while the lyases, carbonic anhydrase I and II [7,8], and the hydrolases, β-lactamase [9] and DD-carboxypeptidase [10] all have three histidine ligands. The only catalytic zinc site to contain just one histidine residue is that of alcohol dehydrogenase [11].

Correspondence address: B.L. Vallee, Women's Hospital, 250 Longwood Avenue, Boston, MA 02135, USA

This also is unique in being the only site with cysteine ligands.

In four of the enzymes, glutamate is the third protein ligand but aspartate appears in one case. It is now known that zinc will form complexes with virtually equal readiness with N, O and S ligands in chemical systems, but in enzymes, it seems to prefer the imidazole nitrogen for catalytic zinc binding sites. On the other hand, for the two non-catalytic zinc binding sites characterized which have been in enzymes, i.e. in alcohol dehydrogenase [11] and aspartate carbamoyltransferase [12], cysteines are the sole ligands, and there are four, not three as in the case of the catalytic zinc sites.

The fourth ligand in all the catalytic sites is H_2O (table 1). The zinc-bound water is thus a critical component of the active site; it is activated for enzymatic catalysis by the nature and spatial arrangement of the protein ligands to zinc. Ultimately, it is this water molecule which, upon entering the zinc coordination sphere, becomes either activated by ionization, or polarization, or is poised for displacement, thus making it the decisive catalytic group.

The spacing between the ligands is particularly striking. In nine of the cases there are only one to three amino acids separating the first two ligands (table 1). This short spacer suggests that this arrangement of protein residues always ensures a nucleus for zinc binding. Equally characteristic is the fact that the third ligand is relatively distant from either of the first two ligands in the linear sequence, minimally ~20 and maximally ~120 residues away, the long spacer. This third protein

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

Table 1

Zinc ligands and their spacing for the catalytic zinc[a]

Enzyme	L_1	X	L_2	Y	L_3	L_4	Reference
Carbonic anhydrase I, II	His	1	His	22	His	H_2O	7,8
β-Lactamase	His	1	His	121	His	H_2O	9
DD carboxypeptidase	His	2	His	40	His	H_2O	10
Thermolysin	His	3	His	19	Glu	H_2O	3
B. cereus neutral protease	His	3	His	19	Glu	H_2O	4
Carboxypeptidase A	His	2	Glu	123	His	H_2O	2
Phospholipase C	His	3	Glu	13	His	H_2O	5
Alkaline phosphatase	Asp	3	His	80	His	H_2O	6
Alcohol dehydrogenase	Cys	20	His	106	Cys	H_2O	11

[a] The X spacing is the number of amino acids between L_1 and L_2 while the Y spacing is the number of amino acids between L_3 and its nearest zinc ligand neighbor. If L_3 comes after L_1 in the linear sequence, Y represents the number between L_3 and L_2; if L_3 comes before L_1, it is the number between L_3 and L_1

ligand thus not only secures the zinc coordination site but also brings into alignment the substrate binding groups critical for catalysis.

The presence of closely spaced ligands at the structural zinc sites would seem to be consistent with their role in stabilizing protein conformation. Such an arrangement of ligands would likely confirm rigidity to that region of the molecule and, hence, function in much the same way that calcium, for example, is thought to stabilize protein structures. In contrast, longer spacings seen in catalytic sites would imply much greater flexibility in metal coordination geometry. Presumably this would be required in order to poise the active site for catalysis and also to allow for changes in coordination chemistry during catalysis. In addition, it would bring together those amino acid side chains that are to function in the catalytic process including an outer sphere ligand whose coordination would activate water. Equally important is the creation of a substrate binding pocket as well as the potential for conformational change that flexible coordination would provide. Variation in spacing length could therefore reflect differences in substrate specificity, function of water and catalytic mechanism.

It is important to point out that when an X-ray structure is not available as a reference for a given category of zinc enzymes, it will be difficult to infer all three ligands to the zinc. However, if sequences of amino acids can be found where a histidyl residue is separated by one to three amino acids from another histidyl, glutamyl or aspartyl residue, then a reasonable measure of success can be expected for identifying two of the ligands [13,14].

The characteristically short (1–3 amino acid) and long (~20–120 amino acid) spacers observed for the catalytic zinc sites also pertains to a number of zinc sites of as yet undefined function in the dimetal complexes of superoxide dismutase (His 61, His 69, His 78, Asp 81) [15], phospholipase C (Asp 55, His 69, His 118, Asp 122) [5], and alkaline phosphatase (Asp 51, Asp 369, His 370) [6].

A number of copper and iron containing proteins also exhibit these spacing characteristics. Thus, the Fe-S cluster complex of aconitase is coordinated to His 37, Cys 359, Cys 422, and Cys 425 [16]. Furthermore, in the ascorbate oxidase from zucchini [17], the mononuclear copper is ligated to His 446, Cys 508, His 513, and Met 518, and the trinuclear copper cluster contains the ligands His 62, His 64, His 106, His 108, His 449, His 451, His 507, and His 509. In these cases, the amino acid spacing likely provides the scaffolding for interaction with the catalytic metal ion.

The data strongly imply that the identity of the ligands to zinc, their sequential alignment, proximity or distance from one another, and the presence or absence of water in the coordination sphere relates to the enzymatic function of the zinc site. This is reflected in terms of the reaction specificity of the enzyme and ultimately, no doubt, to the mechanism(s) by which catalysis occurs.

REFERENCES

[1] Vallee, B.L. and Galdes, A. (1984) Adv. Enzymol. Relat. Areas Mol. Biol. 56, 283–430.

[2] Rees, D.C., Lewis, M. and Lipscomb, W.N. (1983) J. Mol. Biol. 168, 367–387.

[3] Matthews, B.W., Jansonius, J.N., Colman, P.M., Schoenborn, B.P. and Dupourque, D. (1972) Nature New Biol. 238, 37–41.

[4] Pauptit, R.A., Karlsson, R., Picot, D., Jenkins, J.A., Niklaus-Reimer, A.-S. and Jansonius, J.N. (1988) J. Mol. Biol. 199, 525–537.

[5] Hough, E., Hansen, L.K., Birknes, B., Jynge, K., Hansen, S., Horvik, A., Little, C., Dodson, E. and Derewenda, Z. (1989) Nature 338, 357–360.

[6] Sowadski, J.M., Handschumacher, M.D., Krishna Murthy, H.M., Foster, B.A. and Wyckoff, H.W. (1985) J. Mol. Biol. 186, 417–433.

[7] Liljas, A., Kannan, K.K., Bergsten, P.C., Waara, I., Fridborg, K., Strandberg, B., Carlbom, V., Järup, L., Lövgren, S. and Petef, M. (1972) Nature 235, 131–137.

[8] Kannan, K.K., Notstrand, B., Fridborg, K., Lövgren, S., Orlsson, A. and Petef, M. (1975) Proc. Natl. Acad. Sci. USA 72, 51–55.

[9] Sutton, B.J., Artymiuk, P.J., Cordero-Borboa, A.E., Little, C., Phillips, D.C. and Waley, S.G. (1987) Biochem. J. 248, 181–188.

[10] Dideberg, O., Charlier, P., Dive, G., Joris, B., Frère, J.M. and Ghuysen, J.M. (1982) Nature 299, 469–470.

[11] Branden, C.I., Jörnvall, M., Eklund, M. and Furugren, B. (1975) in: Enzymes, vol.11, 3rd edn (Boyer, P.D. ed.) p.103, Academic Press, New York.

[12] Honazatko, 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. and Lipscomb, W.N. (1982) Mol. Biol. 160, 219–263.

[13] Stöcker, W., Wolz, R., Zwilling, R., Strydom, D.J. and Auld, D.S. (1988) Biochemistry 27, 5026.
[14] Jongeneel, C.V., Bouvier, J. and Bairoch, A. (1989) FEBS Lett. 242, 211–214.
[15] Tainer, J.A., Getzoff, E.D., Beem, K.M., Richardson, J.S. and Richardson, D.C. (1982) J. Mol. Biol. 160, 181–217.
[16] Robbins, A.M. and Stout, C.D. (1989) J. Inorg. Biochem. 36, 301.
[17] Messerschmidt, A., Rossi, A., Lodenstein, R., Huber, R., Bolognesi, M., Gatti, G., Marchesini, A., Petruzzelli, R. and Finazzi-Agro, A. (1989) J. Mol. Biol. 203, 513–529.