larger than that of the N7-MeG formed at an isolated AGT or AGA site. For the G_3 run the average value of N7-MeG is essentially the same for an isolated base. Therefore, the methylating intermediate is not specifically attracted to the dG_3 or dG_4 runs, but rather, the reaction with DNA is focused within these regions to select guanines. Furthermore, the 6-8-fold difference between the strongest and weakest methylation site, G_{281} vs G_{282} and G_{293} vs G_{294}, indicates a substantial difference in the nucleophilicity of the guanine residues as a function of local sequence, assuming a minimal role for steric effects. This interpretation requires electrophilic selectivity by $CH_3N_2^+$. Previous reports shown k_{LiN_3}/k_{MeOH} ratios of 17.6 and 18.4 for 1-butanediazonium ion and 1-hexanediazonium ion, respectively, in aqueous CH_3OH.[22] These are values close to that found with *t*-BuBr.[22] We have found a k_{NaI}/k_{NaBr} ratio of 4.5 for the reaction of 1-propanediazonium ion in phosphate buffer.[23] On the basis of the calculated dissociation enthalpies of $CH_3N_2^+$ and 1-propanediazonium ion of 38 and 10 kcal/mol, respectively,[24] the former species should be capable of the same selectivity as the higher homologues.

Fluctuations in the nucleophilicity at N7-G sites as a consequence of subtle sequence-dependent conformational changes that alter base stacking could account for the methylation pattern. In this context, the relationship between base-stacking geometries and N7-G nucleophilicity may be related to the sequence-related electrostatic component of the stacking energy.[25] Effective intrastrand base–base electrostatic neutralization,[26] which is dependent on the helix twist angle and base roll, tilt, and propeller twist angles, would be expected to diminish nucleophilicity. The dominant effect of the 5′-base on N7-G methylation is consistent with the proposed influence of base stacking, since CPK models show that the N7-G site in B-DNA sits directly over the C4 or N1 of the 5′-purine or -pyrimidine, respectively. The influence of C on lowering the alkylation intensities at a 5′-G is also consistent with the importance of stacking since the 4-amino protons of C, which protrude into the major groove and bear a positive electrostatic charge, are situated over the N7 site of the 5′-G. The influence of a 3′-C on the alkylation of DNA by nitrogen mustards has been noted.[18] The role of base stacking in DNA and RNA alkylation, as it affects the overall proportion of adducts, has also been previously recognized. The absolute amount of *N*7-MeG is approximately 2-fold higher in poly(dG)·poly(dC) than in any dG-containing alternating heteropolymer.[12d] The enhanced methylation of guanine and adenine in RNA by MNNG under conditions favoring base stacking has also been reported.[27] Under the same conditions, DMS methylated both 5′-ribonucleotides and polyribonucleotides to a similar extent.[27]

Conclusions

The common sequence selectivity of the different methylating precursors and the salt-induced inhibition of DNA methylation, although restricted to events at N7-G, are not in agreement with the recently proposed regioselective methylation mechanism. They are congruous with the involvement of a common $CH_3N_2^+$ species that is formed external to the double helix. The origin of the sequence-dependent selectivity for certain G sites remains a topic for further investigation.

Acknowledgment. We thank Solon Rhode for making available the parvovirus plasmid. This work was funded by Public Health Service Grant CA29088 and Laboratory Cancer Research Center Support Grant CA36727 awarded by the National Cancer Institute and by American Cancer Society Center Core Grant ACS SIG-16. The support of K.M.C. by an Elizabeth Bruce Memorial Fellowship is gratefully acknowledged.

(22) Kirmse, W.; Schnurr, O.; Jendralla, H. *Chem. Ber.* **1979**, *112*, 2120-2144.
(23) Church, K.; Gold, B. *Chem. Res. Toxicol.* **1988**, *1*, 269-273.
(24) Ford, G. P. *J. Am. Chem. Soc.* **1986**, *108*, 5104-5108.
(25) Aida, M.; Nagata, C. *Int. J. Quantum Chem.* **1986**, *29*, 1253-1261.
(26) Sarai, A.; Mazur, J.; Nussinov, R.; Jernigan, R. L. *Biochemistry* **1988**, *27*, 8498-8502.
(27) Singer, B.; Fraenkel-Conrat, H. *Biochemistry* **1969**, *8*, 3260-3266.

Carboxylate–Histidine–Zinc Interactions in Protein Structure and Function

David W. Christianson* and Richard S. Alexander

Contribution from the Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323. *Received December 16, 1988*

Abstract: The three-dimensional structures of proteins contained in the Brookhaven Protein Data Bank were analyzed for bound metal ions. Well over 150 unique protein structures are available which contain seven different types of bound metal ions. Iron, calcium, and zinc are most commonly observed, and the extended coordination polyhedra of biological zinc are the subject of this study. In particular, histidine residues ligating zinc ions are often found to bridge both the zinc ion and the carboxylate side chain of a nearby aspartate (sometimes glutamate) residue. We refer to the carboxylate–histidine–zinc interaction as *indirect* carboxylate–metal coordination, and we observe this feature in all zinc enzymes of reported three-dimensional structure. Additionally, we also observe a related carbonyl–histidine–zinc interaction in some metalloproteins. We observe some *direct* carboxylate–zinc interactions, and their coordination stereochemistry is exclusively syn with respect to the carboxylate. On the basis of available protein structures and known homologues thereof, more than 30 examples of indirect carboxylate–zinc coordination across bridging histidine can be identified. The carboxylate–histidine–zinc triad may be important in the function of many zinc-containing proteins and enzymes, e.g., by strengthening metal complexation or modulating the nucleophilicity of zinc-bound water. The presence of an uncomplexed carboxylate–histidine couple (a grouping more basic than histidine alone) in a native protein can also signal a regulatory metal binding site. Indeed, the Asp^----His couple of the serine protease active site may comprise a structural, evolutionary link to the Asp^----His of the zinc protease metal coordination polyhedron.

Among the first-row transition metals, zinc is second only to iron in terms of abundance and functional importance in biological systems.[1] From an inorganic perspective, the coordination chemistry of divalent zinc cation (Zn^{2+}, hereafter zinc) might be

0002-7863/89/1511-6412$01.50/0 © 1989 American Chemical Society

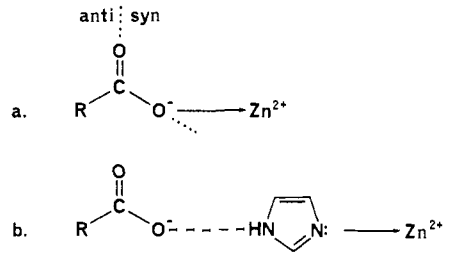

Figure 1. (a) Direct carboxylate-zinc interaction. Syn/anti designations are indicated. (b) Indirect carboxylate-zinc coordination across histidine. This mode is much more common than the direct mode illustrated in (a). The hydrogen bond from histidine is preferably received with syn stereochemistry by the carboxylate.

considered colorless. However, from a biochemical perspective the coordination chemistry of zinc is indeed rich in terms of protein structure and function. On the basis of a survey of the Brookhaven Protein Data Bank (PDB),[2] the most common ligands for zinc include the side chains of histidine and cysteine. Interestingly, the basicity of histidine can be modulated by hydrogen bonds with neighboring protein residues. For instance, His-36 of myoglobin displays a high pK_a of about 8 toward the proton due to a salt link with the carboxylate of Glu-36.[3,4] If the carboxylate-histidine couple is buried (or partially buried) within a protein structure, the ΔpK_a at histidine will be greater; if the histidine is hydrogen bonded to an uncharged carbonyl rather than a charged carboxylate, the ΔpK_a will be about half as much.[5] The basicities of *syn*- and *anti*-carboxylate-imidazole couples toward the proton have been investigated recently in elegant models of the serine protease active site, and the ΔpK_a values observed in these model systems are consistent with those ascertained from protein studies.[6]

Likewise, the basicity of histidine toward a metal ion, rather than a proton, will be significantly modulated by a hydrogen-bonded salt link with a nearby aspartate or glutamate. We designate this binding mode as *indirect* carboxylate-zinc coordination across bridging histidine. If the carboxylate-histidine-zinc triad is identified with regular frequency in known protein structures, a greater understanding will be conferred upon metalloproteins of unknown structure since histidine is most often observed to ligate zinc. For example, consider the "zinc finger" motif of those proteins which exhibit zinc-dependent, sequence-specific DNA binding properties.[7] The postulated finger typically contains two histidine residues that are suspected zinc ligands. The results of a recent EXAFS study support these assignments for the *Xenopus* protein transcription factor IIIA (TFIIIA), in which two other zinc ligands are cysteine side chains.[8] Either one (or both) of the liganding histidine residues could, in turn, be hydrogen bonded to an aspartate or glutamate, and this interaction could occur within or across adjacent fingers. Hence, a carboxylate-histidine-zinc interaction could buttress each zinc finger in the proper topology for DNA binding. We find it intriguing that the zinc finger sequences identified in TFIIIA contain ample amounts of these carboxylate-containing side chains, and in one particular position (aligning with Asp-19 or Glu-48 of the first and second fingers) they are conserved as consensus aspartate or glutamate.[7a]

The direct interaction of carboxylate residues with zinc ions is observed in relatively few native zinc proteins. Nevertheless, the stereochemistry of proton or metal ion association with a carboxylate is described as syn or anti (Figure 1). Gandour[9] first noted that hydrogen bonds in proteins occur preferentially with syn orientation to the carboxylate moieties of aspartate and glutamate, and Peterson and Csizmadia[10] calculated that the *syn*-carboxylic acid is more stable than the anti isomer by 4.5 kcal/mol. Importantly, Wiberg and Laidig[11] demonstrated in subsequent molecular orbital calculations that the inherent stability of the *syn*-carboxylic acid is due to the favorable opposition of C=O and O—H bond dipoles. Extending these considerations to carboxylate-metal ion interactions in proteins, Christianson and Lipscomb[12] found that inhibitor carboxylates preferentially bind to zinc proteases with syn stereochemistry. At the same time, Rebek's group[13] has shown in studies of small molecules that powerful metal ion chelates result from rational placement of carboxylate moieties on chelating agents so that carboxylate-metal ion stereochemistry exploits favorable syn, rather than anti, stereochemistry. Glusker and colleagues[14] have recently surveyed the Cambridge small molecule crystallographic database in order to demonstrate the preferential syn stereochemistry of carboxylate-metal ion interactions.

Stereochemical considerations of indirect carboxylate-zinc interactions across histidine suggest that metal complexation will be most effective if syn hydrogen-bond geometry is achieved between carboxylate and histidine (Figure 1). However, we note that the model systems of Brown indicate that an anti-oriented carboxylate elicits the same ΔpK_a at imidazole toward the proton as does a syn-oriented carboxylate.[6c] In the current study, we evaluate the direct and indirect (i.e., across histidine) coordination behavior of biological carboxylate with zinc as observed in metalloprotein structure. This stereochemical study is relevant to the study of zinc in protein structure and enzyme catalysis.

Methods

The most recently available version of the PDB was used in the structural analyses. Three-dimensional search algorithms and geometric calculations were performed on a VAXstation 3500 with software written by R.S.A. The graphics software package FRODO, developed by Jones,[15] was utilized on an Evans and Sutherland PS300 interfaced with a VAX 11/750. There are well over 100 proteins of the PDB coordinate entries which contain metal ion heteroatoms. There are 67 entries containing iron, 41 containing calcium, 23 containing zinc, 7 containing copper, 4 containing magnesium, 1 containing manganese, 1 containing mercury, and 1 containing an ambiguous manganese or magnesium. In order to analyze and compare zinc binding sites, 11 independent zinc protein coordinate sets were extracted frim the PDB. These proteins were the following: aspartate transcarbamoylase,[16] avian pancreatic polypeptide,[17] carbonic anhydrase I,[18] carbonic anhydrase II,[19] carboxypeptidase A,[20]

(1) Vallee, B. L. In *Zinc Enzymes*; Bertini, I.; Luchinat, C.; Maret, W.; Zeppezauer, M., Eds.; Birkhauser: Boston, 1986; pp 1-15.
(2) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535-542.
(3) Carver, J. A.; Bradbury, J. H. *Biochemistry* **1984**, *23*, 4890-4905.
(4) Bradbury, J. H.; Carver, J. A. *Biochemistry* **1984**, *23*, 4905-4913.
(5) Perutz, M. F.; Gronenborn, A. M.; Clore, G. M.; Fogg, J. H.; Shih, D. T.-b. *J. Mol. Biol.* **1985**, *183*, 491-498.
(6) (a) Zimmerman, S. C.; Cramer, K. C. *J. Am. Chem. Soc.* **1988**, *110*, 5906-5908. (b) Huff, J. B.; Askew, B.; Duff, R. J.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1988**, *110*, 5908-5909. (c) Skorey, K. I.; Somayaji, V.; Brown, R. S. *J. Am. Chem. Soc.* **1989**, *111*, 1445-1452.
(7) (a) Miller, J.; McLachlan, A. D.; Klug, A. *EMBO J.* **1985**, *4*, 1609-1614. (b) Brown, R. S.; Sander, C.; Argos, P. *FEBS Lett.* **1985**, *186*, 271-274. (c) Berg, J. M. *Science* **1986**, *232*, 485-487.
(8) Diakun, G. P.; Fairall, J.; Klug, A. *Nature* **1986**, *324*, 698-699.
(9) Gandour, R. D. *Bioorg. Chem.* **1981**, *10*, 169-176.
(10) Peterson, M. R.; Csizmadia, I. G. J. *Am. Chem. Soc.* **1979**, *101*, 1076-1079.
(11) Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1987**, *109*, 5935-5943.
(12) Christianson, D. W.; Lipscomb, W. N. *J. Am. Chem. Soc.* **1988**, *110*, 5560-5565.
(13) Marshall, L.; Parris, K.; Rebek, J., Jr.; Luis, S. V.; Burguete, M. I. *J. Am. Chem. Soc.* **1988**, *110*, 5192-5193.
(14) Carrell, C. J.; Carrell, H. L.; Erlebacher, J.; Glusker, J. P. *J. Am. Chem. Soc.* **1988**, *110*, 8651-8656.
(15) Jones, T. A. In *Computational Crystallography*; Sayre, D., Ed.; Clarendon: Oxford, 1982; pp 303-317.
(16) 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*, 219-263.
(17) Blundell, T. L.; Pitts, J. E.; Tickle, I. J.; Wood, S. P.; Wu, C.-W. *Proc. Natl. Acad Sci. U.S.A.* **1981**, *78*, 4175-4179.
(18) Kannan, K. K.; Ramanadham, M.; Jones, T. A. *Ann. N.Y. Acad. Sci.* **1984**, *429*, 49-60.
(19) Kannan, K. K.; Liljas, A.; Waara, I.; Bergsten, P.-C.; Lovgren, S.; Strandberg, B.; Bengtsson, U.; Carlbom, U.; Fridborg, K.; Jarup, L.; Petef, M. *Cold Spring Harbor Symp. Quant. Biol.* **1971**, *36*, 221-231.

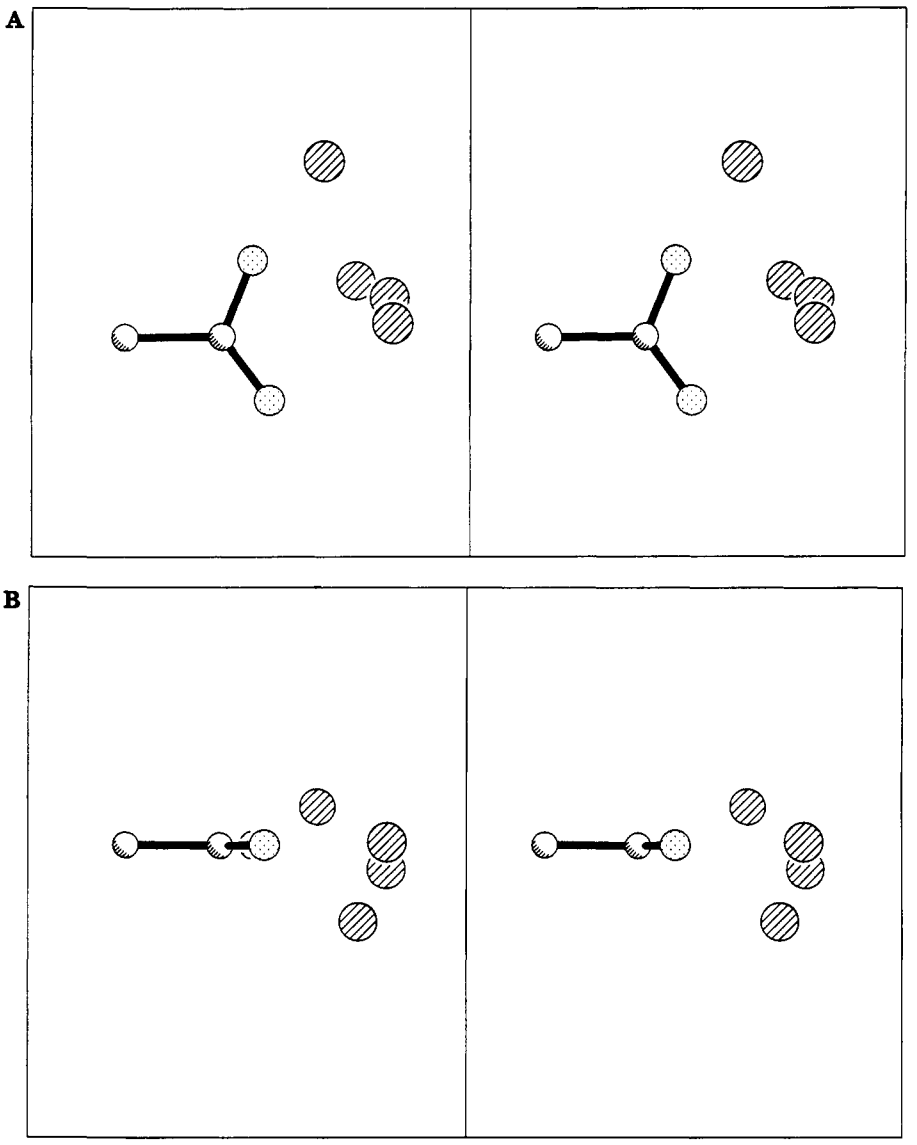

Figure 2. Stereo plots of direct carboxylate–zinc interactions (two perpendicular orientations shown). Although only four examples are currently available in the PDB (one each in avian pancreatic polypeptide, carboxypeptidase A, superoxide dismutase, and thermolysin), each displays syn stereochemistry with regard to the reference carboxylate. This figure can be compared with Figure 3 of ref 13. Atom designations throughout the figures are as follows: oxygen = dotted; nitrogen = clear; carbon = partially cross-hatched; zinc = fully cross-hatched.

insulin,[21] liver alcohol dehydrogenase,[22] metallothionein (Cd, Zn),[23] superoxide dismutase,[24] thermolysin,[25] and tonin.[26] Although there are a total of 23 coordinate entries in the PDB which contain zinc ions, 11 of these are derivatives of a parent protein (e.g., an enzyme–inhibitor complex or modified protein), and for one (DNA polymerase[27]) only C_α coordinates are reported. We analyzed the 11 independent zinc-containing proteins for the occurrence and geometries of carboxylate–zinc and histidine–zinc interactions. Histidine–zinc interactions were further analyzed for additional contacts with carboxylate side chains of aspartate or glutamate, or the amide carbonyls of asparagine, glutamine, and/or the polypeptide backbone. Unexpected stereochemical correlations were found among these zinc-containing proteins of remarkably unrelated origins.

Results and Discussion

We summarize the zinc coordination polyhedra of the 11 unique zinc protein structures in Table I. Histidine and cysteine side chains constitute the most prevalent zinc ligands, and there are only four examples of direct carboxylate–zinc interactions (each with syn stereochemistry) in native protein structures: one each in avian pancreatic polypeptide, carboxypeptidase A, superoxide dismutase, and thermolysin (Figure 2). Direct carboxylate–zinc interactions have been reported for DNA polymerase, but side-chain coordinates are not available in the PDB for this enzyme.[27] It has been noted, in examples of zinc enzyme–inhibitor complexes, that potent inhibition arises from *syn*-carboxylate–zinc coordination stereochemistry.[12]

It is significant that the carboxylate–histidine–zinc interaction occurs within several unrelated proteins (seven out of nine possible examples; see Figure 3). We find that 35% of all histidine–zinc interactions are actually indirect carboxylate–zinc interactions across bridging histidine, and it is significant that all catalytic zinc ions are liganded by such an Asp(Glu)$^-$---His couple. Interestingly, in 1974 Liljas and Rossmann noted the presence of

(20) Rees, D. C.; Lewis, M.; Lipscomb, W. N. *J. Mol. Biol.* **1983**, *168*, 367–387.
(21) Bordas, J.; Dodson, G. G.; Grewe, H.; Koch, M. H. J.; Krebs, B.; Randall, J. *Proc. R. Soc. London, Ser. B* **1983**, *219*, 21–39.
(22) Eklund, H.; Nordstrom, B.; Zeppezauer, E.; Soderlund, G.; Ohlsson, I.; Boiwe, T.; Soderberg, B.-O Tapia, O.; Brändén, C.-I.; Akeson, A. *J. Mol. Biol.* **1976**, *102*, 27–59.
(23) Furey, W. F.; Robbins, A. H.; Clancy, L. L.; Winge, D. R.; Wang, B. C.; Stout, C. D. *Science* **1986**, *231*, 704–710.
(24) Tainer, J. A.; Getzoff, E. D.; Beem, K. M.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1982**, *160*, 181–217.
(25) Holmes, M. A.; Matthews, B. W. *J. Mol. Biol.* **1982**, *160*, 623–639.
(26) Fujinaga, M.; James, M. N. G. *J. Mol. Biol.* **1987**, *195*, 373–396.
(27) Ollis, D. L.; Brick, P.; Hamlin, R.; Xuong, N. G.; Steitz, T. A. *Nature* **1985**, *313*, 762–766.

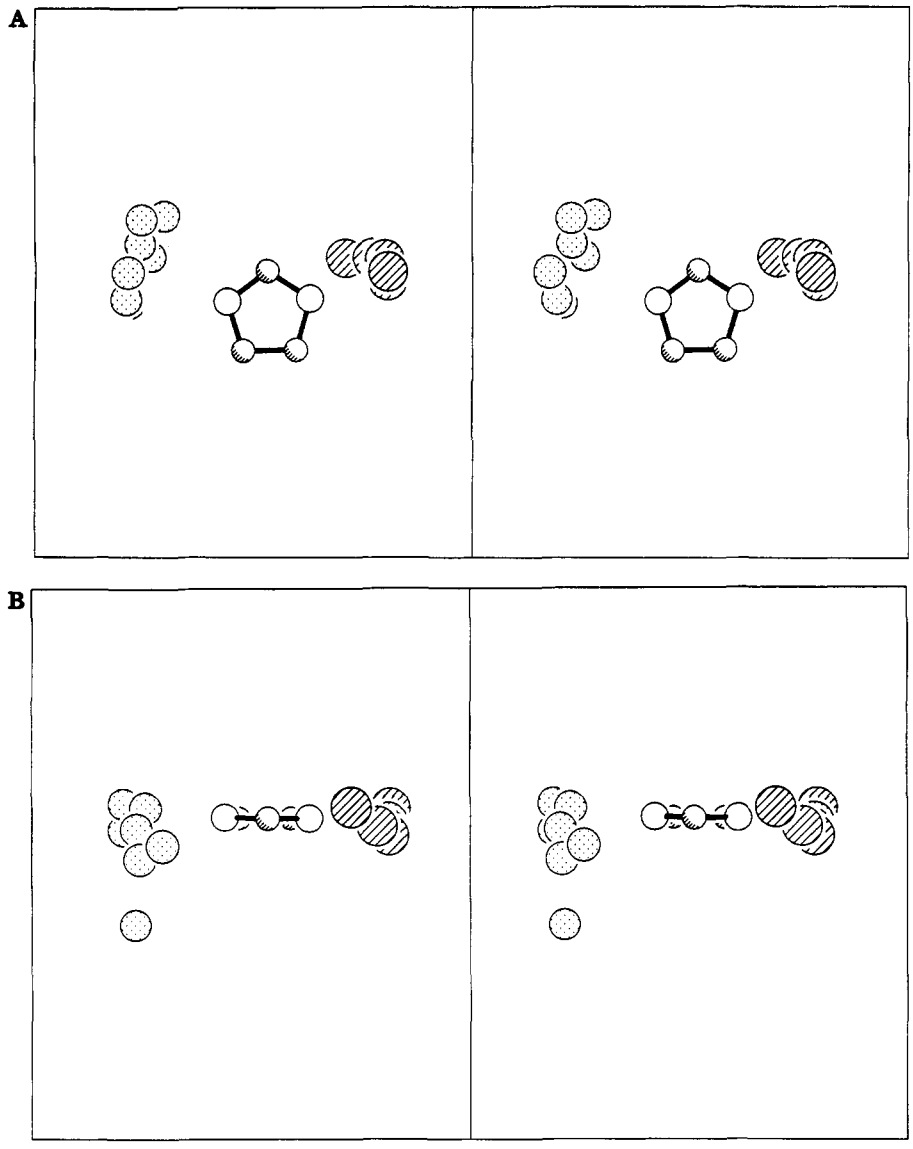

Figure 3. Stereo plots (two perpendicular orientations) of carboxylate–histidine–zinc triads extracted from the PDB are shown by a reference histidine, zinc ions, and the hydrogen-bonded oxygen atoms of the neighboring carboxylates (usually asparate; see Table I). Note the tight clustering of atomic coordinates in this potent charge–charge interaction (i.e., indirect carboxylate coordination to the metal ion) even though abstracted from seven unrelated protein structures. As expected, both oxygen and zinc prefer to be in the plane of the histidine as they intereact with this residue. Relevant statistics are recorded in Table II.

these couples in carbonic anhydrase II, carboxypeptidase A, and thermolysin, but these authors did not recognize the feature as a recurring motif in the greater family of zinc-requiring proteins.[28] The stereochemistry of the carboxylate–histidine hydrogen bond in these metalloprotein complexes is preferentially syn (although not bifurcated) with respect to the carboxylate (Figure 4), indicating a strong, charge–charge interaction across histidine between carboxylate and zinc. We do not observe a particular preference for either histidine tautomer in these interactions, which is intriguing in view of the differing basicities of N_ϵ and N_δ toward the proton.[29]

An additional 30% of histidine–zinc interactions are of a related form: a *carbonyl*–histidine–zinc interaction where the carbonyl of an amino acid side chain, usually glutamine or asparagine, receives a hydrogen bond from a histidine metal ligand (Figure 5). Recall that a carbonyl–histidine interaction is about half as effective as a carboxylate–histidine interaction in modulating the basicity of imidazole toward the proton.[5] It has been suggested that hydrogen bonds to zinc-bound histidines are important for structural reasons, i.e., in order to properly orient metal ligands.[30] However, 65% of all histidine zinc ligands are engaged in carboxylate–histidine or carbonyl–histidine interactions (Figure 5), and we propose that such triads are involved in catalysis, e.g., to electronically enhance metal complexation[28] and modulate the basicity/nucleophilicity of metal-bound water (if present). Site-directed mutagenesis could be employed in the study of zinc enzymes in order to evaluate the contribution of carboxylate–histidine–zinc and carbonyl–histidine–zinc interactions to catalysis.

Although the PDB sample size of carboxylate–histidine–zinc interactions is limited in scope, this scope is immediately broadened by extending the observation to metalloproteins of unknown structure which are related, through divergence, to those of known structure. Even in proteins exhibiting sequence identities as low as 20% with those of known structure, all direct metal ion ligands *and* indirect carboxylate–histidine–metal ion interactions are conserved in the primary structure: currently, we account for 36 examples of carboxylate–histidine–zinc interactions in metalloprotein structures.[31-35] These examples consist of carbonic an-

(28) Liljas, A.; Rossmann, M. G. *Annu. Rev. Biochem.* **1974**, *3*, 475–507.
(29) Tanokura, M. *Biochim. Biophys. Acta* **1983**, *742*, 576–585.
(30) (a) Kester, W. R.; Matthews, B. W. *J. Biol. Chem.* **1977**, *252*, 7704–7710. (b) Argos, P.; Garavito, R. M.; Eventoff, W.; Rossmann, M. G.; Brändén, C.-I. *J. Mol. Biol.* **1978**, *126*, 141–158.
(31) Carbonic anhydrase homologues/isozymes, see: Hewett-Emmett, D.; Hopkins, P. J.; Tashian, R. E.; Czelusniak, J. *Ann. N.Y. Acad. Sci.* **1984**, *429*, 338–358, and references cited therein.

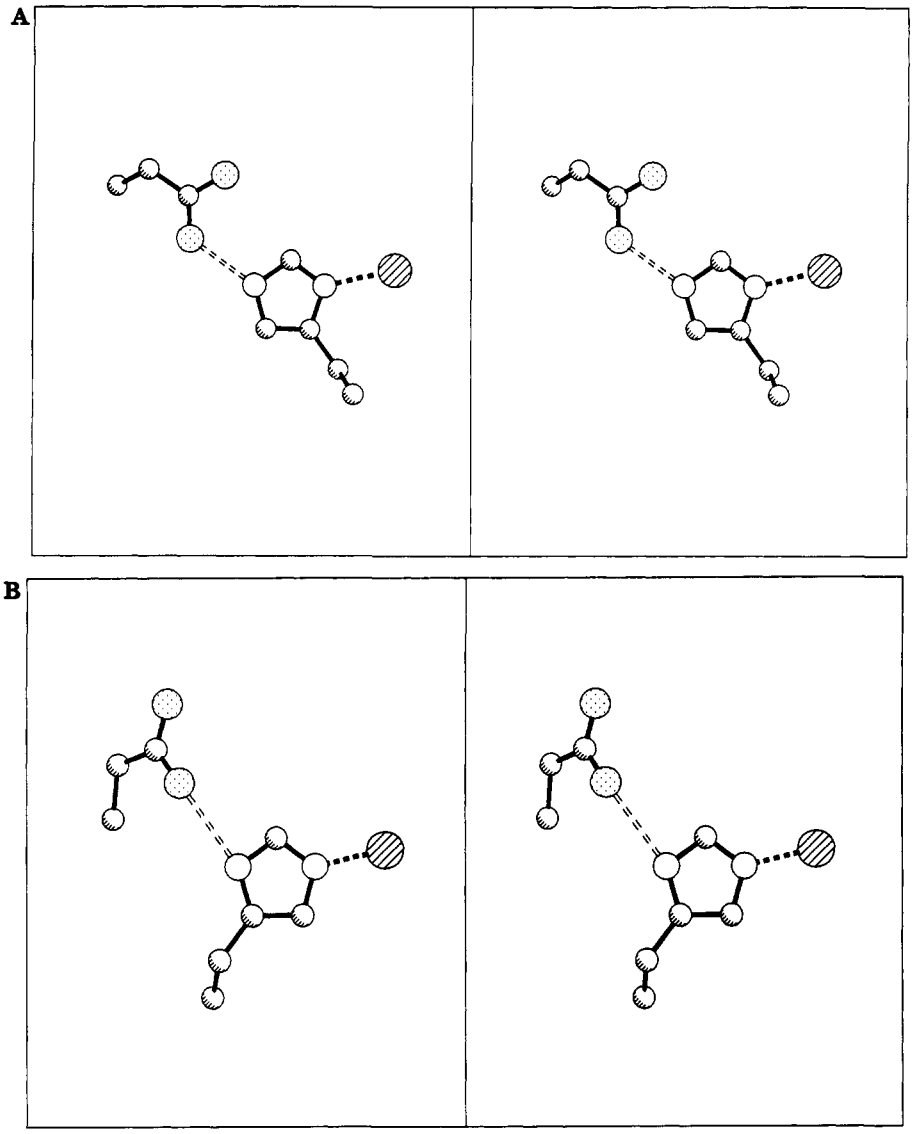

Figure 4. Carboxylate stereochemistry, as it hydrogen bonds to the histidine zinc ligand, is usually syn as shown by Asp-122/His-69 of superoxide dismutase (A). However, sometimes this stereochemistry is anti, as exemplified by Asp-170/His-142 of thermolysin (B).

hydrases and isozymes from human, horse, rhesus, ox, turtle, rabbit, mouse, and chicken sources;[31] carboxypeptidases A, B, and E from bovine and rat sources;[32] alcohol dehydrogenases from human, horse, rat, and yeast;[33] superoxide dismutases from human, bovine, horse, yeast, and bacterial sources;[34] and neutral proteases from various bacterial sources.[35] Hence, we propose that the carboxylate–histidine–zinc interaction is an important feature in the metal binding domains of zinc-requiring proteins, and particularly so where zinc is involved in catalysis. Future X-ray crystallographic studies of other independent zinc-containing proteins will no doubt provide additional examples of this recurring structural motif.

Geometric parameters for all carboxylate–histidine–zinc and carbonyl–histidine–zinc interactions observed in PDB structures are recorded in Tables II and III, respectively. For the carboxylate–histidine–zinc interactions, the average Zn^{2+}···N–C(histidine) angle of 119 ± 12° (errors represent standard deviation of sample distribution) and the average Zn^{2+}···N–C–N(histidine) dihedral angle of 183 ± 10° indicate that zinc is strongly complexed by the sp^2 lone electron pair on liganding histidine (expected values for these angles, at sp^2-hybridized nitrogen, are 120° and 180°, respectively). Likewise, the average (carboxylate)O···N–C(histidine) angle of 116 ± 19° and the average carboxylate(O)···N–C–N(histidine) dihedral angle of 169 ± 23° indicate that the hydrogen-bonded salt link between the carboxylate and histidine is similarly strong. It is interesting to compare these figures, reported in Table II, with those reported solely for the carbonyl–histidine–zinc interactions in Table III; the greater standard deviations for carbonyl–histidine stereochemistry probably reflect the weaker, indirect interaction of carbonyl dipole with zinc as compared with the stronger, indirect interaction of ionized carboxylate with zinc. The above statistics regarding metal ion complexation and hydrogen bonding about histidine in metalloproteins are well in accord with studies of hydrogen-bond geometry

(32) Carboxypeptidase homologues: (a) Titani, K.; Ericsson, L. H.; Walsh, K. A.; Neurath, H. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1666–1670. (b) Fricker, L. D.; Evans, C. J.; Esch, F. S.; Herbert, E. *Nature* **1986**, *323*, 461–464. (c) Quinto, C. A.; Quiroga, M.; Swain, W. F.; Nikovitz, W. C., Jr.; Standring, D. N.; Pictet, R. L.; Valenzuela, P.; Rutter, W. J. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 31–35.

(33) Liver alcohol dehydrogenase homologues, see: Brändén, C.-I.; Jornvall, H.; Eklund, H.; Furugren, B. *Enzymes (3rd Ed.)* **1975**, *11*, 104–189, and references cited therein.

(34) Superoxide dismutase homologues, see: (a) Steffens, G.-J.; Bannister, J. V.; Bannister, W. H.; Flohe, L.; Gunzler, W. A.; Kim, S. A.; Otting, F. *Hoppe-Seyler's Z. Physiol. Chem.* **1983**, *364*, 675–690. (b) Fridovich, I. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1986**, *58*, 61–97, and references cited therein.

(35) Thermolysin homologues: (a) Sidler, W.; Niederer, E.; Suter, F.; Zuber, H. *Biol. Chem. Hoppe-Seyler* **1986**, *367*, 643–657. (b) Yang, M. Y.; Ferrari, E.; Henner, D. J. *J. Bacteriol.* **1984**, *160*, 15–21. (c) Takagi, M.; Imanaka, T.; Aiba, S. *J. Bacteriol.* **1985**, *163*, 824–831. (d) Vasantha, N.; Thompson, L. D.; Rhodes, C.; Banner, C.; Nagle, J.; Filpula, D. *J. Bacteriol.* **1984**, *159*, 811–819.

Table I. Coordination Polyhedra of Zinc-Containing Proteins

protein	PDB reference	resolution (Å)	refinement *R* factor	zinc ligands	carboxylate/histidine	carbonyl/histidine
alcohol dehydrogenase	4ADH	2.4	0.26	Cys-46, His-67, Cys-174[a] Cys-97, Cys-100, Cys-103, Cys-111[b]	*syn*-Asp-49/His-67 ($N_{\delta 1}$)	
aspartate transcarbamoylase	4ATC	2.6	0.24	Cys-109, Cys-114, Cys-137, Cys-140		
carbonic anhydrase II	1CAC	2.0	*c*	His-94, His-96, His-119	*anti*-Glu-117/His-119 ($N_{\varepsilon 2}$)	Asn-244 (backbone O)/His-96 ($N_{\delta 1}$) Gln-92 (amide O)/His-94 ($N_{\delta 1}$)
carbonic anhydrase I	2CAB	2.0	0.19	His-94, His-96, His-119	*anti*-Glu-117/His-119 ($N_{\varepsilon 2}$)	Gln-92 (amide O)/His-94 ($N_{\delta 1}$)
carboxypeptidase A	5CPA	1.5	0.19	His-69, Glu-72, His-196, H_2O	*syn*-Asp-142/His-69 ($N_{\varepsilon 2}$)	
insulin	1INS	1.5	0.18	$3H_2O$, 3 His-10		
metallothionein (Cd, Zn)	2MT2	2.3	0.34	Cys-13, Cys-24, Cys-26, Cys-29 Cys-5, Cys-7, Cys-21, Cys-24		
avian pancreatic hormone	1PPT	1.4	0.16	Gly-1, Asn-23, His-34, H_2O		
superoxide dismutase	2SOD	2.0	0.26	His-61, His-69, His-78, Asp-81	*syn*-Asp-122/His-69 ($N_{\varepsilon 2}$)	Asn-63 (amide O)/His-78 ($N_{\delta 1}$)
thermolysin	3TLN	1.6	0.21	His-142, His-146, Glu-166, H_2O	*anti*-Asp-170/His-142 ($N_{\delta 1}$)	Asn-165 (amide O)/His-146 ($N_{\delta 1}$)
tonin	1TON	1.8	0.20	His-57, His-97, His-99, Glu-148[d]	*syn*-Asp-102/His-57 ($N_{\delta 1}$)	Ile-95 (backbone O)/His-97 ($N_{\delta 1}$)

[a] Catalytic. [b] Structural. [c] PDB coordinates are not fully refined; the refined structure ($R = 0.19$) was reported by Eriksson, E. A.; Jones, T. A.; Liljas, A. In *Zinc Enzymes*; Bertini, I.; Luchinat, C.; Maret, W.; Zeppezauer, M., Eds.; Birkhauser: Boston, 1986; pp 317–328. The refined zinc coordination polyhedron shows improved geometry, yet all relevant interactions are identical with those determined from the unrefined structure. [d] This ligand is from a symmetry-related molecule in the crystallographic unit cell.

about sp^2 nitrogen heterocycles observed in the small-molecule crystallographic surveys of Vedani and Dunitz.[36] A Lewis acid (a hydrogen-bond donor or a metal ion) prefers a head-on and an in-plane approach to the sp^2 lone electron pair of nitrogen. Herein, we have shown that the second nitrogen (i.e., N–H) of the histidine imidazole also prefers head-on and in-plane hydrogen-bond geometry with the neighboring carboxylate as it is found in the recurring carboxylate–histidine–zinc triad.

Indirect carboxylate–metal ion complexation across histidine is observed in the PDB predominantly for zinc. We do, however, observe examples involving iron (two in myohemerythrin) and copper (one in superoxide dismutase) ligation. We conclude that the carboxylate–histidine couple, typically appearing as Asp^----His, signals a preferential binding site for transition metals. From this conclusion, for example, we can rationalize the behavior of the serine protease tonin as it sequesters zinc[26] and the platination of His-57 of chymotrypsin.[37] Additionally, Asp-226/His-231 of thermolysin is said to be a location for the binding of excess, inhibitory zinc ion.[38] Carboxypeptidase A contains two Glu^----His couples unoccupied by metal ions in the native enzyme. One of these sites, Glu-218/His-303, is observed occupied by platinum ion in recent X-ray crystallographic studies.[39] Interestingly, this site (21 Å away from the active site), or less likely the Glu-17/His-13 site (25 Å away from the active site but not geometrically optimal), may be responsible for binding excess, inhibitory zinc.[40] Asp^----His or Glu^----His sites are present at an average of about one per unique protein structure in the PDB solely on the basis of distance criteria, but some of these sites exhibit less-than-optimal geometry for hydrogen bonding. Nevertheless, histidine residues of these sites additionally hydrogen bond to a water molecule, a backbone carbonyl oxygen, or a serine hydroxyl group (exclusively in the serine proteases). Such Lewis base rich sites imply an ability on the part of the protein to complex regulatory metal ions as activators or inhibitors. Perhaps the Asp^----His couple of the serine protease active site is evolutionarily linked, via some primordial ancestor, to the Asp^----His couple ligating the metal ion of the zinc protease. This speculation is intriguing in view of the X-ray structure of the serine protease tonin, where the Asp-102/His-57 couple unexpectedly complexes adventitious, inhibitory zinc.[26]

In a model of the zinc coordination polyhedron of carboxypeptidase A, Nakagawa and colleagues have demonstrated in molecular orbital calculations that the presence of a carboxylate–histidine–zinc interaction facilitates proton transfer from histidine to the carboxylate of aspartate.[41] These investigators find that the forms CO_2^-···H–His and CO_2H···His^- are approximately isoenergetic in the presence of zinc, but in the absence of zinc they find that the latter form is destabilized by 24 kcal/mol. The result for an isolated CO_2^-···H–His interaction is consistent with other molecular orbital calculations (in models of the serine protease active site) where this form is considered by far to be the energetically favorable species.[42] These theoretical studies concur that the presence of the carboxylate enhances the basicity of histidine. The suggestion of Nakagawa and colleagues that zinc facilitates the formation of the histidinate anion, however, may rest upon the in vacuo nature of molecular orbital calculations and requires experimental verification. However, if the results of calculation are correct, then the carboxylate–histidine–zinc triad

(36) Vedani, A.; Dunitz, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7653–765; see Figures 1 and 2.
(37) Brothers, H. M.; Kostić, N. M. *Biochemistry* **1989**, *28*, 1944.
(38) (a) Pangburn, M. K.; Walsh, K. A. *Biochemistry* **1975**, *14*, 4050–4054. (b) Holmquist, B.; Vallee, B. L. *J. Biol. Chem.* **1974**, *249*, 4601–4607.
(39) Shoham, G.; Christianson, D. W., unpublished results.
(40) Hirose, J.; Ando, S.; Kidani, Y. *Biochemistry* **1987**, *26*, 6561–6565.
(41) Nakagawa, S.; Umeyama, H.; Kitaura, K.; Morokuma, K. *Chem. Pharm. Bull.* **1981**, *29*, 1–6.
(42) (a) Kollman, P. A.; Hayes, D. M. *J. Am. Chem. Soc.* **1981**, *103*, 2955–2961. (b) Umeyama, H.; Nakagawa, S.; Kudo, T. *J. Mol. Biol.* **1981**, *150*, 409–421. (c) Umeyama, H.; Hirono, S.; Nakagawa, S. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6266–6270. (d) Nakagawa, S.; Umeyama, H. *J. Mol. Biol.* **1984**, *179*, 103–123.

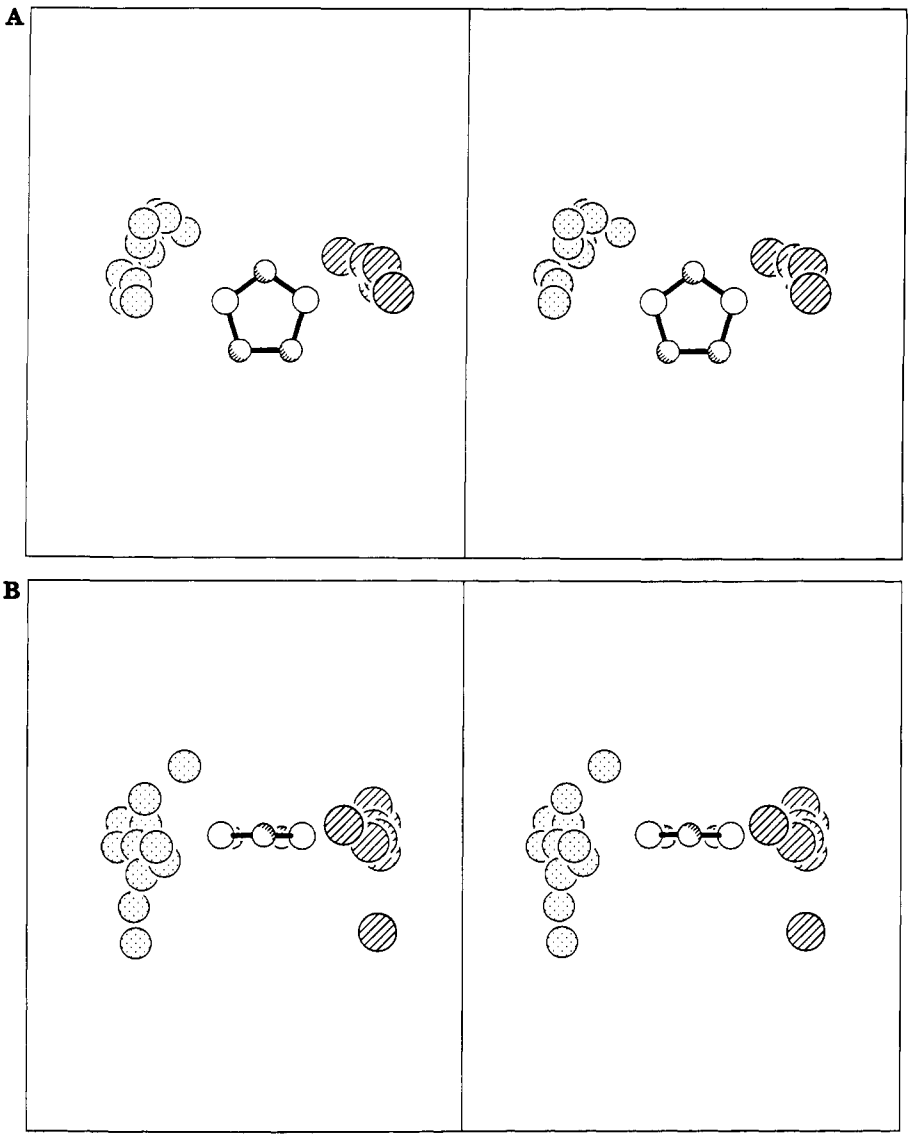

Figure 5. Stereo plots of carboxylate–histidine–zinc and carbonyl–histidine–zinc interactions extracted from the PDB. Two perpendicular orientations are shown. Note the wider scatter of oxygen atoms, corresponding to carbonyl and carboxylate oxygens, which hydrogen bond to the histidine zinc ligand. Relevant statistics for carbonyl–histidine–zinc interactions are recorded in Table II.

Table II. Stereochemistry of Carboxylate–Histidine–Zinc Interactions

protein	distance (Å) Zn–N	angle (deg) Zn–N–C	dihedral angle (deg) Zn–N–C–N	distance (Å) O–N	angle (deg) O–N–C	dihedral angle (deg) O–N–C–N
alcohol dehydrogenase	2.2	115	190	3.1	99	163
carbonic anhydrase I	1.9	119	189	2.6	137	172
carbonic anhydrase II	1.5	96	168	3.6	117	120
carboxypeptidase A	2.1	127	179	2.6	144	186
superoxide dismutase	2.1	129	196	2.4	111	188
thermolysin	2.1	131	173	2.7	91	173
tonin	2.0	119	188	2.7	110	179
average	2.0	119	183	2.8	116	169
σ	0.2	12	10	0.4	19	23

Table III. Stereochemistry of Carbonyl–Histidine–Zinc Interactions

protein	distance (Å) Zn–N	angle (deg) Zn–N–C	dihedral angle (deg) Zn–N–C–N	distance (Å) O–N	angle (deg) O–N–C	dihedral angle (deg) O–N–C–N
carbonic anhydrase I	1.9	129	174	2.9	124	177
	1.9	135	185	2.7	105	201
carbonic anhydrase II	3.3	114	243	2.7	87	219
superoxide dismutase	2.0	126	201	2.8	146	252
thermolysin	2.1	116	184	2.9	96	174
tonin	2.1	129	198	2.7	126	206
average	2.2	125	198	2.8	114	205
σ	0.5	8	24	0.1	22	29

is more properly described as a carboxylic acid-histidinate-zinc triad.

The carboxylate-histidine-zinc interaction is found in enzymes of different origins and markedly different functions, yet it may serve a unified function in catalysis. Indeed, this triad may distinguish catalytic zinc from structural zinc in native metalloprotein structures. In addition to promoting metal ion complexation, one possible role for the carboxylate-histidine-zinc interaction (or the weaker carbonyl-histidine-zinc interaction) is the modulation of the nucleophilicity of zinc-bound water. Nature can "fine tune" the nucleophilicity of zinc-bound water not only by suitable selection of direct metal ligands such as the side chains of histidine, cysteine, and glutamate but also by modulating the contacts of these side chains with neighboring enzyme residues. Histidine, because of its potent, dual Lewis acid/base character, is most susceptible to such modulation. Although cysteine or glutamate can make hydrogen-bond contacts while coordinated to metal ions, these contacts are not favorable due to electronic or stereochemical reasons: sulfur is not preferred in hydrogen bonding, and a carboxylate in syn coordination to a metal ion can receive a hydrogen bond only in the less favorable anti orientation. Hence, histidine is most susceptible to wide-range modulation by the protein as to the strength and "hardness" of metal ion complexation.

We conclude our discussion of carboxylate-histidine-zinc interactions by returning to the exemplar of the zinc finger. Although the zinc of TFIIIA presumably does not interact directly with the DNA molecule, it might do so indirectly. For instance, a phosphate-histidine-zinc interaction, rather than the analogous carboxylate-histidine-zinc interaction proposed earlier, may be important in the zinc-dependent DNA binding behavior of the postulated zinc fingers. Either possibility might be consistent with the structure of an isolated zinc finger recently proposed from a two-dimensional NMR study,[43] and our continuing studies of phosphate-Lewis acid stereochemistry may allow us to evaluate these possibilities.

Summary

We have identified a recurring structural feature in the extended coordination polyhedron of biological zinc, the carboxylate-histidine-zinc interaction. This triad is usually found as Asp^----His $\rightarrow Zn^{2+}$, and we find that syn stereochemistry is preferred between the carboxylate and histidine. All interacting atoms of the triad are held firmly in the plane of histidine's imidazole ring, and we find this recurring triad in all zinc enzymes of known three-dimensional structure. Importantly, the carboxylate-histidine-zinc interaction, as so far observed, may distinguish catalytic zinc from structural zinc in native metalloprotein structures. We currently account for 36 examples of this triad in proteins of known structure and their homologues with sequence identities as low as 20%. Interestingly, the Asp^----His $\rightarrow Zn^{2+}$ triad of the zinc protease may bear an evolutionary resemblance to the Asp^----His couple of the serine protease. Additionally, a carboxylate-histidine couple, where unoccupied by metal ions in protein structures (on average, about one per unique protein structure deposited in the PDB), may signal a regulatory location for the binding of transition metal ions.

Acknowledgment. We thank the National Institutes of Health for Grant GM 08275. Additionally, we thank the Research Foundation of the University of Pennsylvania for its support of this research, and we are grateful to Prof. M. G. Rossmann for helpful comments.

Registry No. His, 71-00-1; Asp, 56-84-8; Zn, 7440-66-6.

(43) Parraga, G.; Horvath, S. J.; Eisen, A.; Taylor, W. E.; Hood, L.; Young, E. T.; Klevit, R. E. *Science* **1988**, *241*, 1489-1492.

Determination of the Microscopic Rate Constants for the Base-Catalyzed Conjugation of 5-Androstene-3,17-dione

Ralph M. Pollack,* Baifei Zeng, Joseph P. G. Mack,[1] and Sherif Eldin

Contribution from the Laboratory for Chemical Dynamics, University of Maryland Baltimore County, Baltimore, Maryland 21228. *Received January 20, 1989*

Abstract: The hydroxide ion catalyzed isomerization of 5-androstene-3,17-dione (**1**) to 4-androstene-3,17-dione (**2**) proceeds through the formation of an intermediate dienolate ion (1^-). This dienolate ion has been observed in the ultraviolet spectrum (λ_{max} ca. 256 nm) during the isomerization reaction. Rate constants for the formation of the dienolate ion and both its reversion to reactant (**1**) and its conversion to product (**2**) in aqueous solution were measured. In addition, the rate of exchange of the C-6 protons of **2** in D_2O/MeOD was determined. These results enable a complete description of the reaction profile to be made, including all rate constants and the pK_a values for **1** (12.7) and **2** (16.1). The possible relevance of these results to the mechanism of action of the enzyme 3-oxo-Δ^5-steroid isomerase is briefly discussed.

The isomerization of β,γ-unsaturated ketones to their α,β-unsaturated isomers is a simple example of a class of reactions involving 1,3 proton shifts.[2] This reaction has proven to be a useful vehicle for the examination of several general phenomena, including stereoelectronic effects on proton transfer,[3] electrostatic catalysis,[3b] substituent effects on double bond protonation,[4] and nucleophilic catalysis.[5] In addition, much effort has been expended to try to elucidate the mechanism of the enzymatic reaction of 5-androstene-3,17-dione (**1**) to 4-androstene-3,17-dione (**2**) catalyzed by 3-oxo-Δ^5-steroid isomerase.[6]

(1) Current address: Crystallography Laboratory, FRRF-NCI, BRI Basic Research Program, P.O. Box B, Frederick, MD 21701.

(2) For recent reviews of relevant material see: (a) Pollack, R. M.; Bounds, P. L.; Bevins, C. L. In *The Chemistry of Functional Groups, Enones*; Patai, S.; Rappoport, Z., Eds.; Wiley: New York, 1989; p 559. (b) Capon, B., *Ibid.* **1989**; p 1063.

(3) (a) Malhotra, S. K.; Ringold, H. J. *J. Am. Chem. Soc.* **1965**, *87*, 3228. (b) Whalen, D. L.; Weimaster, J. F.; Ross, A. M.; Radhe, R. *J. Am. Chem. Soc.* **1976**, *98*, 7319. (c) Pollack, R. M. *Tetrahedron*, in press.

(4) (a) Noyce, D. S.; Evett, M. *J. Am. Chem. Soc.* **1972**, *87*, 394. (b) Noyce, D. S.; Evett, M. *J. Am. Chem. Soc.* **1972**, *87*, 397.

(5) (a) Kayser, R. H.; Pollack, R. M. *J. Am. Chem. Soc.* **1975**, *97*, 952. (b) Pollack, R. M.; Kayser, R. H. *J. Am. Chem. Soc.* **1976**, *98*, 4174. (c) Benisek, W. F.; Jacobson, A. *Bioorg. Chem.* **1975**, *4*, 41. (d) Okuyama, T.; Kitada, A.; Fueno, T. *Bull. Chem. Soc. Jpn.* **1977**, *50*, 2358. (e) Perera, S. K.; Dunn, W. A.; Fedor, L. R. *J. Org. Chem.* **1980**, *45*, 2816.

(6) For reviews see ref 2a and Batzhold, F. H.; Benson, A. M.; Covey, D. F.; Talalay, P. *Adv. Enzyme Regul.* **1976**, *14*, 243.

0002-7863/89/1511-6419$01.50/0 © 1989 American Chemical Society