Carbonic Anhydrase: Evolution of the Zinc Binding Site by Nature and by Design

DAVID W. CHRISTIANSON*

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323

CAROL A. FIERKE*

Department of Biochemistry, Box 3711, Duke University Medical Center, Durham, North Carolina 27710

Received December 18, 1995

Ever since its discovery at the University of Pennsylvania[1] and the University of Cambridge[2] more than 60 years ago, the zinc enzyme carbonic anhydrase has occupied a prominent position at the frontiers of biochemistry, medicinal chemistry, and protein engineering. Seven genetically-distinct forms of this enzyme (known as isozymes I−VII) have evolved in numerous tissues and cellular locations, and each contains a catalytically-obligatory zinc ion.[3,4] This diversity reflects the ubiquitous biological requirement for rapid hydration of carbon dioxide to yield bicarbonate ion plus a proton.[5−9] Although a deceptively simple reaction, the chemical and structural aspects of this mechanism have only recently been delineated for an isozyme found in the red blood cell, carbonic anhydrase II (CAII) (Figure 1). The three-dimensional structure of CAII[10] has stimulated research probing the determinants of the substrate association site and the pathway the product proton traverses through the enzyme active site. On the basis of phylogenetic comparisons,[3,4] the substrate and zinc binding sites are mainly conserved among all catalytic carbonic anhydrase isozymes found in mammals. However, the trajectory of catalytic proton transfer to bulk solvent has diverged during the evolution of the seven isozymes.

We begin this Account with a brief review of the CAII mechanism, emphasizing recent developments (see previous Accounts for additional information[6,7]). Notably, more X-ray crystallographic and enzymological studies have been performed on CAII and its site-specific variants than any other metalloenzyme, and these studies uniquely illuminate the molecular details of catalysis. Importantly, this work provides an elegant example of how the complementary methods of molecular biology and structural biology can be used to probe the structure, function, and stability of a zinc binding site in a metalloenzyme. Additionally, these studies set a useful foundation for understanding the evolution of carbonic anhydrase into noncatalytic biological roles, such as signal transduction. For example, in the nervous system an extracellular, CAII-like domain of receptor protein tyrosine phosphatase β binds an axonal cell recognition molecule (contactin) important for neuronal development and differentiation.[11−15] Intriguingly, the putative zinc binding site of this domain has partially evolved away from that found in CAII.

We then review recent progress in the "directed evolution"—i.e., the structure-based redesign—of the CAII zinc binding site, following Nature's example. We describe novel structural determinants of protein−metal affinity and the chemical reactivity of zinc-bound solvent. Not only does this work represent the first molecular dissection of structure−function relationships in a protein−zinc binding site (and therefore serves as a paradigm for the design of *de novo* metal sites in other proteins), it also sets the foundation for the development and optimization of CAII as a metal ion biosensor. Remarkably, this ubiquitous metalloprotein can be engineered and exploited as a sensitive tool for analytical chemistry and biotechnology.

After receiving A.B., A.M., and Ph.D. degrees in Chemistry from Harvard University, David W. Christianson moved to the University of Pennsylvania in 1988 where he is currently Associate Professor of Chemistry. While at Penn he received a Searle Scholar Award, the Young Investigator Award from the Office of Naval Research, an Alfred P. Sloan Research Fellowship, and a Camille and Henry Dreyfus Teacher-Scholar Award. As a protein crystallographer, Christianson has spent more than a decade studying structure–function relationships in metalloenzymes such as carbonic anhydrase, and a 1989 Account outlines his work on the related zinc enzyme carboxypeptidase A.

Carol A. Fierke received her B.A. and Ph.D. degrees from Carleton College and Brandeis University, respectively. After a postdoctoral fellowship in chemistry at Pennsylvania State University she joined the faculty of the Duke University Medical School where she is currently Associate Professor of Biochemistry and Chemistry as well as a member of the Duke University Cancer Center. While at Duke she received an American Heart Association Established Investigator Award, a David and Lucile Packard Foundation Fellowship, and an American Cancer Society Junior Faculty Research Award. During her career she has investigated molecular determinants of biological catalysis in metalloenzymes such as carbonic anhydrase and ribonuclease P.

(1) Stadie, W. C.; O'Brien, H. *J. Biol. Chem.* **1933**, *103*, 521−529.
(2) Meldrum, N. U.; Roughton, F. J. W. *J. Physiol.* **1933**, *80*, 113−142.
(3) Tashian, R. E. *Adv. Genet.* **1982**, *30*, 321−356.
(4) Hewett-Emmett, D.; Tashian, R. E. *Mol. Phylogenet. Evol.* **1996**, *5*, 50−77.
(5) Coleman, J. E. *J. Biol. Chem.* **1967**, *242*, 5212−5219.
(6) Bertini, I.; Luchinat, C. *Acc. Chem. Res.* **1983**, *16*, 272−279.
(7) Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30−36.
(8) Lindskog, S.; Liljas, A. *Curr. Opin. Struct. Biol.* **1993**, *3*, 915−920.
(9) Silverman, D. N. *Methods Enzymol.* **1995**, *249*, 479−503.
(10) Liljas, A.; Kannan, K. K.; Bergsten, P.-C.; Waara, I.; Fridborg, K.; Strandberg, B.; Carlbom, U.; Jarup, L.; Lovgren, S.; Petef, M. *Nature New Biol.* **1972**, *235*, 131−137.
(11) Krueger, N. X.; Saito, H. *Proc. Natl. Acad. Sci. U.S.A.* **1992,** *89*, 7417−7421.
(12) Barnea, G.; Silvennoinen, O.; Shaanan, B.; Honegger, A. M.; Canoll, P. D.; D'Eustachio, P.; Morse, B.; Levy, J. B.; LaForgia, S.; Huebner, K.; Musacchio, J. M.; Sap, J.; Schlessinger, J. *Mol. Cell. Biol.* **1993,** *13*, 1497−1506.
(13) Levy, J. B.; Cannol, P. D.; Silvennoine, O.; Barnea, G.; Morse, B.; Honegger, A. M.; Haung, J.-T.; Cannizzaro, L. A.; Park, S.-H.; Druck, T.; Huebner, K.; Sap, J.; Ehrlich, M.; Musacchio, J. M.; Schlessinger, J. *J. Biol. Chem.* **1993,** *268*, 10573−10581.
(14) Barford, D.; Jia, Z.; Tonks, N. K. *Nature Struct. Biol.* **1995**, *2*, 1043−1053.
(15) Peles, E.; Nativ, M.; Campbell, P. L.; Sakurai, T.; Martinez, R.; Lev, S.; Clary, D. O.; Schilling, J.; Barnea, G.; Plowman, G. D.; Grumet, M.; Schlessinger, J. *Cell* **1995**, *82*, 251−260.

S0001-4842(95)00123-3 CCC: $12.00

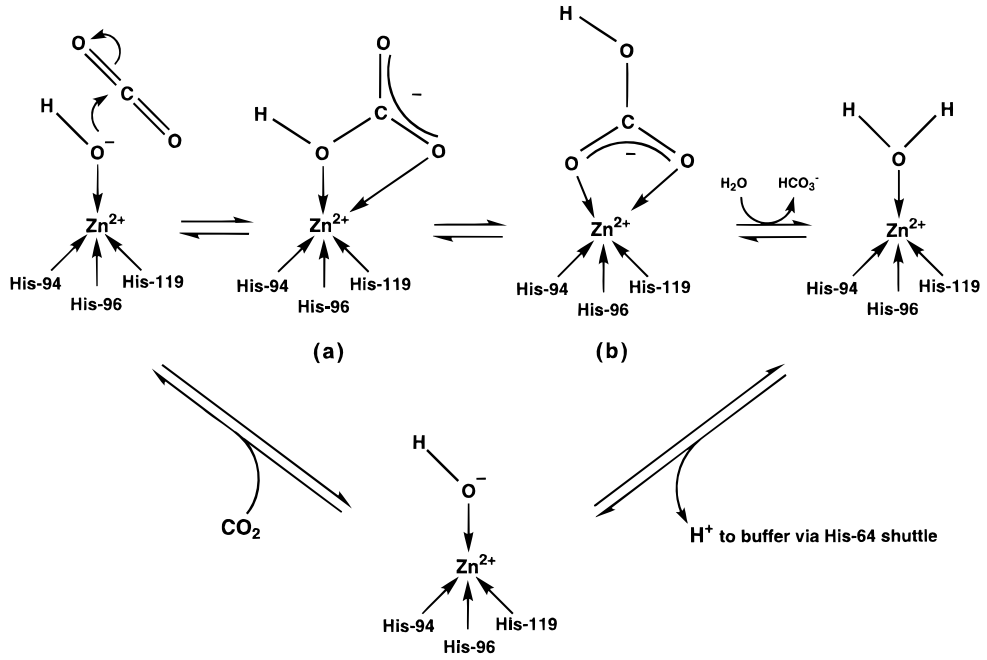

Figure 1. Summary of the CAII mechanism. Zinc-bound hydroxide attacks the carbonyl carbon of CO_2 to form zinc-bound bicarbonate. The initial mode of bicarbonate binding (a) may reflect the structure of either a discrete intermediate or the transition state. Bicarbonate may then isomerize (b), representing either a productive or a nonproductive complex. Following the exchange of a water molecule for zinc-bound bicarbonate, a proton is transferred from zinc-bound water to solvent via His-64 to regenerate the zinc hydroxide species.

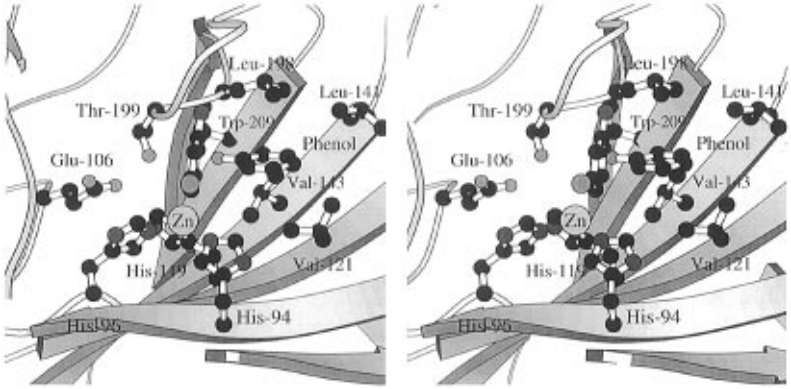

Figure 2. CAII active site, complexed with the competitive inhibitor phenol bound in the hydrophobic pocket.[31] Important active-site residues are labeled. Note that inhibitor (and therefore substrate) binding does not require the displacement of zinc-bound solvent (unlabeled gray sphere). Figure generated with MOLSCRIPT.[89]

Mechanism of Catalysis: Substrate Binding and Proton Transfer

The substrate association site of CAII is a hydrophobic pocket adjacent to zinc-bound hydroxide,[10,16] formed in large part by residues Val-143 at its base and Val-121, Trp-209, and Leu-198 at its neck (Figure 2). This pocket is highly conserved among all active isozymes on the basis of phylogenetic comparisons,[3,4] although in isozyme III the pocket is somewhat constricted by the bulky side chain of Phe-198.[17] The structure-based dissection of this pocket in CAII delineates the minimum size and shape of the pocket required for enzymatic activity; for instance, the occlusion of the pocket by the radical Val-143 → Phe substitution obliterates catalytic activity due to the loss of the substrate association site.[18,19] More moderate amino acid substitutions in the pocket (e.g., Val-143 → Gly or Val-121 → Leu) modestly affect the stability of the transition state for CO_2 hydration (rather than substrate affinity) and alter the structure of active site solvent, leading to slight changes in the rate constant for proton transfer from zinc-bound water to His-64 (see Figure 1). The hydrophobic pocket therefore has a minimum width[20−23] and depth[18,19] for efficient catalysis, and linear free energy relationships indicate that the volume of the amino

(16) Lindskog, S. In *Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhauser: Boston, 1986; pp 307−316.
(17) Eriksson, A. E.; Liljas, A. *Proteins: Struct., Funct. Genet.* **1993**, *16*, 29−42.
(18) Alexander, R. S.; Nair, S. K.; Christianson, D. W. *Biochemistry* **1991**, *30* 11064−11072.
(19) Fierke, C. A.; Calderone, T. L.; Krebs, J. F. *Biochemistry* **1991**, *30*, 11054−11063.
(20) Nair, S. K.; Calderone, T. L.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* **1991**, *266*, 17320−17325.
(21) Nair, S. K.; Christianson, D. W. *Biochemistry* **1993**, *32*, 4506−4514.
(22) Krebs, J. F.; Rana, F.; Dluhy, R. A.; Fierke, C. A. *Biochemistry* **1993**, *32*, 4496−4505.
(23) Nair, S. K.; Krebs, J. F.; Christianson, D. W.; Fierke, C. A. *Biochemistry* **1995**, *34*, 3981−3989.

acid residue at the base of the pocket is critical for activity,[19] whereas the hydrophobicity of residues at the neck of the pocket is likewise important.[20,22,23]

These results are consistent with molecular dynamics simulations that point to the hydrophobic pocket as a substrate association site.[24−26] Experimental evidence supporting the catalytic role of the hydrophobic pocket comes from a variety of spectroscopic studies,[27,28] including Fourier transform infrared spectroscopy.[22,29] The shift of the asymmetric stretching vibration of CO_2 to a lower wavenumber upon association with CAII is consistent with the transfer of CO_2 from aqueous solution to a hydrophobic environment. These experiments also indicate that the affinity of CAII for CO_2 is low (~0.1 M), as required by the high turnover number of the enzyme which necessitates a rapid product dissociation rate constant. Finally, although the precatalytic enzyme−substrate complex is too short-lived to be observed by traditional X-ray crystallographic methods, the structure of the only known competitive inhibitor of CAII-catalyzed CO_2 hydration, phenol,[30] has been solved in complex with the enzyme:[31] phenol binds in the hydrophobic pocket and makes van der Waals contacts with Val-121, Val-143, Leu-198, and Trp-209 while its hydroxyl group hydrogen bonds with zinc-bound hydroxide (Figure 2). Since a competitive inhibitor must bind in the same location as the substrate, CO_2 must therefore bind in the hydrophobic pocket prior to catalysis. Importantly, CO_2 binding does not displace zinc-bound hydroxide, although long-range (i.e., >3 Å) weakly-polar interactions with zinc may contribute to substrate orientation.

Product bicarbonate ion is formed by the nucleophilic attack of zinc-bound hydroxide at CO_2 immobilized in the hydrophobic pocket, and the binding mode of bicarbonate ion to zinc has been the subject of some controversy. Central to this controversy is the possible role of Thr-199 as a "doorkeeper".[32−34] That is to say, the side chain hydroxyl group of this residue is proposed to allow only anions capable of donating a hydrogen bond to Thr-199 access to zinc binding. X-ray crystallographic structures of Co^{2+}-substituted and Thr-199 → Ala CAIIs in complex with bicarbonate are consistent with this interpretation.[35,36] However, the binding of azide anion, a competitive inhibitor of bicarbonate dehydration, demonstrates that Thr-199 is *not* a doorkeeper all the time:[37,38] azide anion binds to zinc but does not hydrogen bond with Thr-199. Instead, the zinc-bound azide nitrogen makes a non-hydrogen-bonded, van der Waals interaction with the Thr-199 hydroxyl group. Such a binding mode is likewise possible for product bicarbonate, as illustrated in Figure 1. Although such an interaction does not contribute to enzyme-product affinity, it is nevertheless likely to contribute to efficient catalysis by facilitating rapid product dissociation. Consistent with this interpretation are studies of Thr-199 variants showing that deletion of the Thr-199 hydroxyl group stabilizes bicarbonate binding[39,40] and alters the structure of the bound bicarbonate ion.[36] Therefore, the Thr-199 side chain promotes maximum catalytic efficiency by destabilizing the product complex to provide for rapid dissociation and by selecting for a catalytically competent bicarbonate−zinc complex (Figure 1). Thus, it is clear that the zinc binding site and its environment have evolved for optimal catalytic activity.

Following bicarbonate dissociation is the rate-determining step of proton transfer to regenerate the active catalyst, zinc-bound hydroxide (Figure 1).[5−9] The product proton is not transferred from zinc-bound water directly to bulk solvent; instead, it is first transferred to an intermediate "shuttle" residue, and then transferred to bulk solvent. Because of its greater exposure to solvent, the shuttle residue can more efficiently transfer a proton to a variety of acceptors with pK_a values higher than that of water.[7] Interestingly, the proton transfer pathway has divergently evolved among the carbonic anhydrase isozymes. In isozyme II, His-64 is the catalytic proton shuttle,[41,42] and it exhibits significant conformational mobility which accompanies its function.[43−45] This residue is too far from zinc-bound solvent to allow for direct proton transfer; instead, proton transfer is achieved across two intervening, hydrogen-bonded solvent molecules in the native enzyme (Figure 3). Isozymes IV, VI, and VII also contain His-64, which may similarly function as a proton shuttle (isozyme I contains His-64, but His-200 is the major proton shuttle group in this isozyme[46]). However, isozyme III contains Lys-64, and the main proton transfer pathway is direct transfer to bulk solvent.[47] Intriguingly, isozyme V contains Tyr-64, which plays no major role in proton transfer[48] due in part to steric effects arising from the bulky adjacent side chain of Phe-65.[49] The three-

(24) Merz, K. M. *J. Mol. Biol.* **1990**, *214*, 799−802.
(25) Liang, J.-Y.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3675−3679.
(26) Merz, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 406−411.
(27) Williams, T. J.; Henkens, R. W. *Biochemistry* **1985**, *24*, 2459−2462.
(28) Bertini, T.; Luchinat, C.; Monnanni, R.; Roelens, S.; Moratal, J. M. *J. Am. Chem. Soc.* **1987**, *109*, 7855−7856.
(29) Riepe, M. E.; Wang, J. H. *J. Biol. Chem.* **1968**, *243*, 2779−2787.
(30) Simonsson, I.; Jonsson B.-H.; Lindskog, S. *Biochem. Biophys. Res. Commun.* **1982**, *108*, 1406−1412.
(31) Nair, S. K.; Ludwig, P. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1994**, *116*, 3659-3660.
(32) Håkansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. *J. Mol. Biol.* **1992**, *227*, 1192-1204.
(33) Liljas, A.; Håkansson, K.; Jonsson, B. H.; Xue, Y. *Eur. J. Biochem.* **1994**, *219*, 1−10.
(34) Lindahl, M.; Svensson, L. A.; Liljas, A. *Proteins: Struct., Funct. Genet.* **1993**, *15*, 177-182.
(35) Håkansson, K.; Wehnert, A. *J. Mol. Biol.* **1992**, *228*, 1212−1218.
(36) Xue, Y.; Liljas, A.; Jonsson, B.-H. *Proteins: Struct., Funct. Genet.* **1993**, *17*, 93−106.
(37) Nair, S. K.; Christianson, D. W. *Eur. J. Biochem.* **1993**, *213*, 507−515.
(38) Jonsson, B. M.; Håkansson, K.; Liljas, A. *FEBS Lett.* **1993**, *322*, 186−190.
(39) Krebs, J. F.; Ippolito, J. A.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* **1993**, *268*, 27458−27466.
(40) Liang, Z.; Xue, Y.; Behravan, G.; Jonnson, B.-H.; Lindskog, S. *Eur. J. Biochem.* **1993**, *211*, 821−827.
(41) Steiner, H.; Jonsson, B.-H.; Lindskog, S. *Eur. J. Biochem.* **1975**, *59*, 253−259.
(42) Tu, C. K.; Silverman, D. N.; Forsman, C.; Jonsson, B.-H.; Lindskog, S. *Biochemistry* **1989**, *28*, 7913−7918.
(43) Krebs, J. F.; Fierke, C. A.; Alexander, R. S.; Christianson, D. W. *Biochemistry* **1991**, *30*, 9153−9160.
(44) Nair, S. K.; Christianson, D. W. *J. Am. Chem. Soc.* **1991**, *113*, 9455−9458.
(45) Taoka, S.; Tu, C.; Kistler, K. A.; Silverman; D. N. *J. Biol. Chem.* **1994**, *269*, 17988-17992.
(46) Engstrand, C.; Johnson, B. H.; Lindskog, S. *Eur. J. Biochem.* **1995**, *229*, 696−702.
(47) Jewell, D. A.; Tu, C.; Paranawithana, S. R.; Tanhauser, S. M.; LoGrasso, P. V.; Laipis, P. J.; Silverman, D. N. *Biochemistry* **1991**, *30*, 1484−1490.
(48) Heck, R. W.; Tanhauser, S. M.; Manda, R.; Tu, C. K.; Laipis, P. J.; Silverman, D. N. *J. Biol. Chem.* **1994**, *269*, 24742−24746.
(49) Boriack-Sjodin, P. A.; Heck, R. W.; Laipis, P. J.; Silverman, D. N; Christianson, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10955−10959.

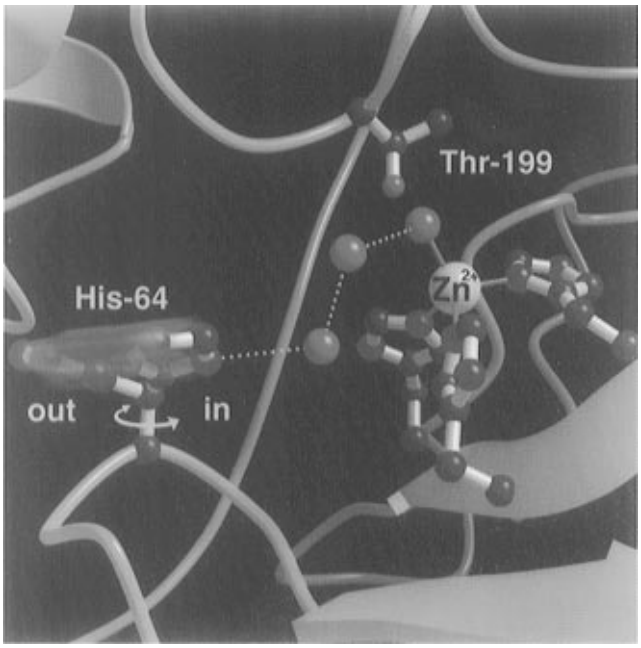

Figure 3. Proton shuttle His-64 adopts the "in" conformation in native CAII and engages in a hydrogen-bonded solvent network (red) with zinc-bound solvent at pH 8.5.[32] Proton transfer across this network regenerates the reactive zinc-bound hydroxide species in catalysis (Figure 1). Upon protonation at lower pH values, His-64 swings away from the active site to the "out" conformation.[44] Thus, His-64 is conformationally mobile, as indicated. Figure generated with MOLSCRIPT.[89]

dimensional structure of isozyme V suggests that a residue in the vicinity of Tyr-131 (e.g., the phenolic side chain of this residue or one of three flanking lysines) may serve as a proton shuttle group,[49] but further experiments are necessary to confirm this speculation. Rate constants for proton transfer in the carbonic anhydrases are dependent on the difference in pK_a between zinc-bound water and the proton shuttle group, and are limited by organization of the intervening hydrogen-bonded solvent molecules.[9,45,50] Individual components of the zinc binding site may therefore affect catalytic proton transfer by modulating the pK_a of zinc-bound water.

Structure-Based Protein Engineering: Redesigning the Zinc Binding Site

As summarized in the preceding section and in Figure 1, the zinc ion in the CAII active site plays a central role in the mechanism of CO_2 hydration. The principal role of zinc is that of an electrostatic catalyst, since it stabilizes the negatively-charged transition state leading to bicarbonate formation. It also depresses the pK_a of its bound water molecule to provide an ample supply of nucleophilic hydroxide ion for catalysis at neutral pH values. Logically, the protein environment of zinc is critical for optimizing the electrostatic properties of the metal ion for catalysis, and this environment includes not only the direct metal ligands but also the residues hydrogen bonding with these ligands (i.e., "indirect" or "second shell" ligands) (Figure 4). By combining the techniques of molecular and structural biology, we have dissected the determinants of affinity and catalysis in the zinc binding site of CAII.

Lessons learned in the structure-based redesign of the CAII zinc binding site are also relevant to the design and construction of *de novo* transition metal binding sites in other proteins,[51−57] since only a few examples are structurally characterized by atomic resolution X-ray crystallographic methods to provide ultimate proof-of-design.[58,59] Indeed, examples characterized by other biophysical techniques sometimes illustrate the limitations of current structure-based design approaches despite the best attempts of computational structure analysis—for example, where "designed" metal ligands do not coordinate to the target metal ion,[60,61] or where zinc affinity is at least 4 orders of magnitude weaker[57] than that measured for CAII.[62,63] Furthermore, none of the designed sites function as efficient catalysts. Hence, there is much yet to be learned regarding structure−activity and structure−stability relationships in protein−metal binding sites, and the zinc binding site of CAII is a universally-recognized and easily-characterized paradigm. We note that there is a novel application of redesigned CAII zinc binding site variants in analytical chemistry and biotechnology: a CAII-based metal ion biosensor.[64] Following Nature's example in the evolution of carbonic anhydrase for different biological functions, the "directed evolution", or structure-based redesign, of the CAII zinc binding site allows for the optimization of its properties[65] in the development of a zinc biosensor.

Altering the Direct Zinc Ligands. A thorough understanding of three factors must precede the molecular dissection and redesign of direct metal ligands in any protein−metal binding site. First, the chemical nature of the target metal must be considered. It could be "hard", like the small, highly-charged magnesium ion, it could be "soft", like the large, highly-polarizable mercury ion, or it could be intermediate between these two extremes (i.e., "borderline").[66] An optimized metal binding site in a protein molecule will exhibit hardness complementary to that of its target metal ion. Since zinc is a metal of borderline hardness, it is satisfactorily coordinated by the harder ligands aspartate and glutamate, the borderline ligand histidine, and the softer ligand cysteine. Second, the conformations of the amino acid side chains that comprise the metal coordination site

(50) Silverman, D. N.; Tu, C.; Chen, X.; Tanhauser, S. M.; Kresge, A. J.; Laipis, P. J. *Biochemistry* **1993**, *32*, 10757−10762.
(51) Tainer, J. A.; Roberts, V. A.; Getzoff, E. D. *Curr. Opin. Biotechnol.* **1992**, *3*, 378−387.
(52) Regan, L. *Trends Biochem. Sci.* **1995**, *20*, 280−285.
(53) Christianson, D. W. *Adv. Protein Chem.* **1991**, *42*, 281−355.
(54) Higaki, J. N.; Fletterick, R. J.; Craik, C. S. *Trends Biochem. Sci.* **1992**, *17*, 100−104.
(55) Klemba, M.; Gardner, K. H.; Marino, S.; Clarke, N. D.; Regan, L. *Nature Struct. Biol.* **1995**, *2*, 368−373.
(56) Elling, C. E.; Nielsen, S. M.; Schwartz, T. W. *Nature* **1995**, *374*, 74−77.
(57) Müller, H. N.; Skerra, A. *Biochemistry* **1994**, *33*, 14126−14135.
(58) McGrath, M. E.; Haymore, B. L.; Summers, N. L.; Craik, C. S.; Fletterick, R. J. *Biochemistry* **1993**, *32*, 1914−1919.
(59) Browner, M. F.; Hackos, D.; Fletterick, R. J. *Nature Struct. Biol.* **1994**, *1*, 327−333.
(60) Hellinga, H. W.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 763−785.
(61) Hellinga, H. W.; Caradonna, J. P.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 787−803.
(62) Lindskog, S.; Nyman, P. O. *Biochim. Biophys. Acta* **1964**, *85*, 462−474.
(63) Kiefer, L. L.; Krebs, J. F.; Paterno, S. A.; Fierke, C. A. *Biochemistry* **1993**, *32*, 9896-9900.
(64) Thompson, R. B.; Jones, E. R. *Anal. Chem.* **1993**, *65*, 730−734.
(65) Kiefer, L. L.; Paterno, S. A.; Fierke, C. A. *J. Am. Chem. Soc.* **1995**, *117*, 6831−6837.
(66) Pearson, R. G. *J. Am. Chem. Soc.* **1963** *85*, 3533−3539.

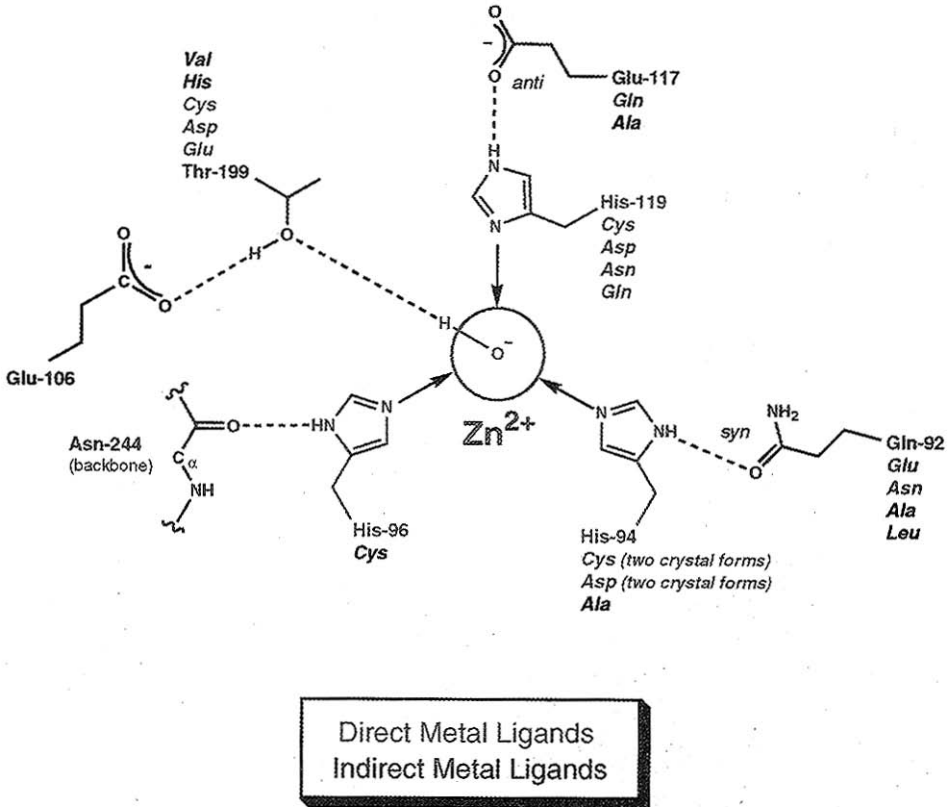

Figure 4. Scheme of the CAII zinc binding site. Variants with characterized properties and three-dimensional structures are indicated by italics at the site of substitution.

Table 1. Properties of Selected CA Variants with Altered Zinc Binding Sites[a]

variant	k_{cat}/K_M (μM^{-1} s^{-1})	pK_a	zinc K_d (nM)
wild type	110	6.8	0.004
H94D	0.11	≥9.6	15
H94C	0.11	≥9.5	33
H96C	0.073	8.5	60
H119C	0.11	nd	50
H119D	3.8	8.6	25

[a] Data from ref 75. nd = no data.

must be favorable.[54,67] If an engineered ligand must incur too high an energetic cost to coordinate to a metal ion, it will not do so. Even so, certain regions of a protein structure—e.g., loops, α-helices, or β-strands—can be sufficiently pliable to allow for modest conformational changes or segmental shifts which optimize metal coordination by an engineered ligand. Finally, and perhaps most importantly, the separation and stereochemistry of ligand−metal coordination must be reasonable. The geometric preferences of carboxylate,[68,69] imidazole,[70] and thiolate[71] ligands for protein−metal ion coordination have been outlined, and the results of these studies provide structural criteria by which metal binding site designs can be evaluated and optimized.

The alteration of direct zinc ligands in CAII yields important structural and functional insight as the properties of selected variants (Table 1) are interpreted in light of their three-dimensional structures.[72−76] As a point of reference, recall that the protein ligands to zinc in wild-type CAII are His-94, His-96, and His-119 (Figures 2 and 4). A notable feature in the CAII zinc binding site is that metal binding to the histidine ligands is very cooperative: the deletion of any one protein ligand by a histidine → alanine substitution decreases the zinc affinity[75] by a factor of ~10^5. This is much larger than the 10−100-fold decrease in metal affinity observed for removing one ligand from Cys_2His_2 zinc binding sites, including a zinc finger peptide[77] and a *de novo* designed zinc coordination polyhedron in the 4α-helical bundle protein designated $Z\alpha_4$.[78]

Similarly, substitution of a neutral histidine ligand by the negatively-charged side chains of aspartate, glutamate, or cysteine decreases the zinc affinity ~10^4-fold.[75] The structure of His-94 → Asp CAII reveals that this loss of protein−metal affinity arises primarily from the 0.9 Å movement of zinc that accommodates the shorter side chain of the ligand. Intriguingly, the structure of His-94 → Cys CAII demonstrates that the zinc ion occupies the same site in the His-94 → Asp and His-94 → Cys variants;[75,76] additionally, a 1 Å deformation of the β-strand containing Cys-94 ensures that this shortest ligand side chain coordinates to zinc (Figure 5). Notably, this β-strand

(67) Ponder, J. W.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 775−791.
(68) Christianson, D. W.; Alexander, R. S. *J. Am. Chem. Soc.* **1989**, *111*, 6412−6419.
(69) Chakrabarti, P. *Protein Eng.* **1990**, *4*, 49−56.
(70) Chakrabarti, P. *Protein Eng.* **1990**, *4*, 57−63.
(71) Chakrabarti, P. *Biochemistry* **1989**, *28*, 6081−6085.
(72) Alexander, R. S.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 1510−1518.
(73) Ippolito, J. A.; Nair, S. K.; Alexander, R. S.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Protein Eng.* **1995**, *8*, 975−980.
(74) Kiefer, L. L.; Ippolito, J. A.; Fierke, C. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1993**, *115*, 12581−12582.
(75) Kiefer, L. L.; Fierke, C. A. *Biochemistry* **1994**, *33*, 15233−15240.
(76) Ippolito, J. A.; Christianson, D. W. *Biochemistry* **1994**, *33*, 15241−15249.
(77) Merkle, D. L., Schmidt, M. H.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 5450−5451.
(78) Klemba, M.; Regan, L. *Biochemistry* **1995**, *34*, 10094−10100.

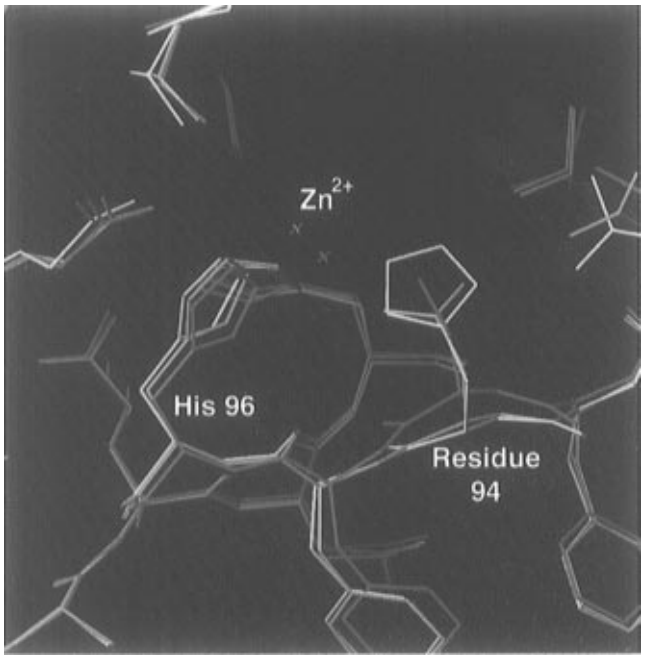

Figure 5. Superposition of wild-type (yellow), His-94 → Asp (cyan), and His-94 → Cys (red) CAIIs. Note the plasticity of the β-strand containing residue 94, as well as the movement of Zn^{2+} which accommodates amino acid substitutions. Zn^{2+} moves to an identical location in the His-94 → Asp and His-94 → Cys variants. Reprinted with permission from ref 76. Copyright 1994 American Chemical Society.

deformation results in only an additional 2-fold loss of protein−metal affinity relative to that of His-94 → Asp CAII. Thus, the β-strand containing metal ligands His-94 and His-96 is sufficiently pliable to accommodate various metal ligand residues at position 94. However, it is the movement of zinc from its naturally-evolved wild-type position—and *not* the compensatory plasticity of the β-strand containing the metal ligands—that most seriously compromises protein−metal affinity. It is intriguing that β-sheet structure and stability can be regulated by the construction of transition metal binding sites, and this feature can be exploited in protein engineering experiments with many other systems.

The introduction of a negative charge into the zinc coordination polyhedron yields a significant increase in the pK_a of zinc-bound water as assayed by the pH dependence of esterase activity (in addition to catalyzing CO_2 hydration, CAII also catalyzes ester hydrolysis; Table 1).[72,74,75] This effect must arise from simple electrostatics: by interacting with a newly-introduced negative charge from the protein, zinc cannot stabilize a bound hydroxide ion as effectively as in the wild-type metalloenzyme. These results are consistent with *ab initio* calculations indicating that the pK_a of zinc-bound water is dependent on the charge of the ligands.[79] In general, ΔpK_a values are reliable and sensitive indicators of changes in the electrostatic environment of zinc.

Remarkably, the hydrase and esterase activities of CAII variants with substitutions in either the direct or indirect ligands are inversely proportional to the pK_a of the zinc−water ligand (Figure 6).[43,65,75] This is contrary to structure−activity relationships for a

(79) Bertini, I.; Luchinat, C.; Rosi, M.; Sgamellotti, A.; Tarantelli, F. *Inorg. Chem.* **1990**, *29*, 1460−1463.

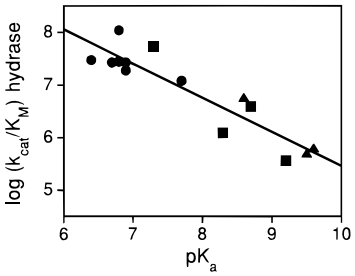

Figure 6. log of the pH-independent second-order rate constant for CO_2 hydration versus the zinc−water pK_a for wild-type and variants of CAII,[43,65,75] including indirect ligand substitutions at position 92 (Gln → Ala, Leu, Asn, Glu) and position 117 (Glu → Asp, Ala) (●), direct ligand substitutions (His-94 → Cys, His-94 → Asp, His-119 → Asp) (▲), and position 199 variants (Thr → Ser, Ala, Val, Pro) (■). The data are fit to a line yielding R = 0.92 and slope −0.65.

small molecule nucleophile attacking an amide, ester, or carbonyl carbon where nucleophilicity increases with basicity.[80] The negative slope observed for CAII variants indicates that electrostatic stabilization of the transition state by the positively-charged zinc ion, rather than the nucleophilicity of the zinc hydroxide, is a dominant factor in catalysis. One possible interpretation of these data is that the protein ligand−zinc separation changes to facilitate electrostatic stabilization of the transition state by maintaining the bond valence sum of zinc.[81,82] Therefore, any diminution of the positive charge in the zinc binding site has dire consequences for catalysis; however, variants with a substituted neutral ligand, such as His-119 → Gln, retain reasonable catalytic activity (C.-c. Huang and C. A. Fierke, unpublished results) and maintain active site structure (C. A. Lesburg and D. W. Christianson, unpublished results). The neutral histidine zinc ligands in CAII are essential for attaining high catalytic efficiency by enhancing the net positive charge of the metal ion site.[75,79] Intriguingly, the CAII-like domains of receptor protein tyrosine phosphatases contain histidine → glutamine substitutions, but neither zinc binding properties nor catalytic activity has been documented for these systems.[11−15]

Amino acid substitutions elsewhere in the direct metal coordination polyhedron of CAII reveal that not all metal ligands reside on pliable segments of active site β-strands: the His-119 → Cys substitution yields an anomalously long zinc−sulfur separation of 2.8 Å, and the His-96 → Cys substitution leads to an unexpected conformational change for the Cys-96 side chain that places it 6.1 Å away from zinc (a water molecule moves into the intended coordination position of Cys-96, thereby maintaining the tetrahedral geometry).[76] Each of these amino acid substitutions also yields a 10^4-fold loss of protein−metal affinity.[75] Despite the unexpected consequences of the His-96 → Cys and His-119 → Cys substitutions, it is clear that (1) the less buried an engineered metal ligand is in the protein scaffolding, the more capable it will be of achieving metal coordination through a plastic structural change and (2) no matter what amino acid substitution is made in the protein−zinc coordination polyhedron, the coordination geometry of zinc remains

(80) Fersht, A. R. *Enzyme Structure and Mechanism*, 2nd ed.; W. H. Freeman & Co.: New York, 1985; pp 83-85.
(81) Thorp, H. H. *Inorg. Chem.* **1992**, *31*, 1585−1588.
(82) Xiang, S. B.; Short, S. A.; Wolfenden, R.; Carter, C. W. *Biochemistry* **1996**, *35*, 1335−1341.

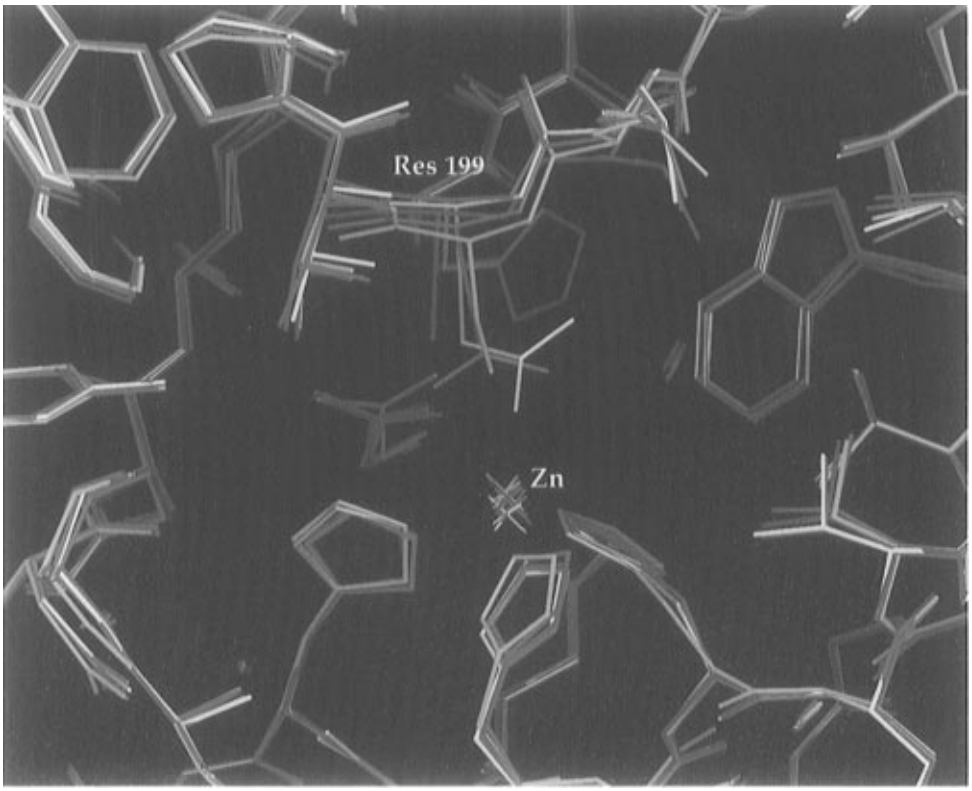

Figure 7. Superposition of wild-type (magenta), Thr-199 → Cys (cyan), Thr-199 → Asp (red), Thr-199 → Glu (yellow), and Thr-199 → His (green) CAIIs. Note the compensatory plasticity of the polypeptide backbone flanking residue 199 which facilitates zinc coordination by the engineered ligand (except in Thr-199 → His CAII); the protein−zinc dissociation constant for Thr-199 → Glu CAII is 20 fM, the highest-affinity site ever engineered into a protein. Reprinted with permission from ref 85. Copyright 1995 National Academy of Sciences.

tetrahedral, even if an exogenous ligand must be recruited from the solvent. This is particularly notable, since the coordination number of zinc in small molecule complexes can be 4 (tetrahedral) or 5 or 6 (octahedral).[83]

Augmenting the Direct Zinc Ligands. Since the alteration of a wild-type protein ligand typically yields a 10^4-fold loss of protein−metal affinity (Table 1), a more viable route toward improving protein−metal affinity might be to augment the metal coordination polyhedron with a fourth direct ligand from the protein. We have successfully achieved the introduction of an additional protein ligand to zinc by substituting cysteine, aspartate, or glutamate for Thr-199.[63,84,85] This residue hydrogen bonds to zinc-bound hydroxide in the wild-type enzyme (Figures 2 and 4), and its location on the relatively pliable[18] Ser-197−Cys-206 loop suggested that we could exploit the compensatory plasticity of this loop to achieve metal coordination by ligand residues of various sizes. Thr-199 variants exhibit greatly diminished catalytic activity,[39,40,63,85] in some cases due to the displacement of the zinc-bound hydroxide by the substituted side chain.[84,85] Additionally, the zinc coordination polyhedron remains tetrahedral in all Thr-199 variants.

As expected, certain Thr-199 variants exhibit enhanced protein−zinc affinity (Table 2). The structural basis for enhanced affinity appears to be the chelate effect: a tetradentate protein−zinc complex is tighter than a tridentate protein−zinc complex, presuming that the engineered protein ligand does not impose unfavorable structural or chemical constraints on its environment (Figure 7). For instance, a modest 4-fold enhancement in zinc affinity is achieved in the Thr-199 → Cys variant, and a greater enhancement is probably precluded by the conformational changes that the engineered side chain and its flanking polypeptide backbone undergo to achieve zinc coordination.[63,84] However, conformational changes accompanying zinc coordination appear to be of smaller magnitude in the Thr-199 → Glu variant, and this may contribute to its 200-fold enhanced zinc affinity.[85] With a zinc dissociation constant of 20 fM, this represents the highest affinity zinc binding site ever engineered in a protein.

Table 2. Properties of Selected CA Variants with Augmented Zinc Binding Sites[a]

variant	k_{cat}/K_M (μM^{-1} s^{-1})	zinc K_d (pM)	variant	k_{cat}/K_M (μM^{-1} s^{-1})	zinc K_d (pM)
wild type	110	4	T199E	0.04	0.02
T199C	0.11	1.1	T199H	0.02	80
T199D	0.04	4			

[a] Data from refs 63 and 84.

These results demonstrate the importance of zinc−ligand separation compared to zinc−ligand stereochemistry for metal affinity; although the Thr-199 → Asp variant exhibits favorable *syn*-bidentate zinc coordination stereochemistry, zinc−oxygen separations of 2.5 Å are moderately long and therefore yield no net enhancement of zinc affinity for this variant.

(83) Cotton, F. A.; Wilkinson, G. *Advanced Inorganic Chemistry*, 5th ed.; John Wiley & Sons: New York, 1988; pp 597−599.
(84) Ippolito, J. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 9901−9905.
(85) Ippolito, J. A.; Baird, T. T.; McGee, S. A.; Christianson, D. W.; Fierke, C.A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5017−5021.
(86) Lesburg, C. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1995**, *117*, 6838−6844.

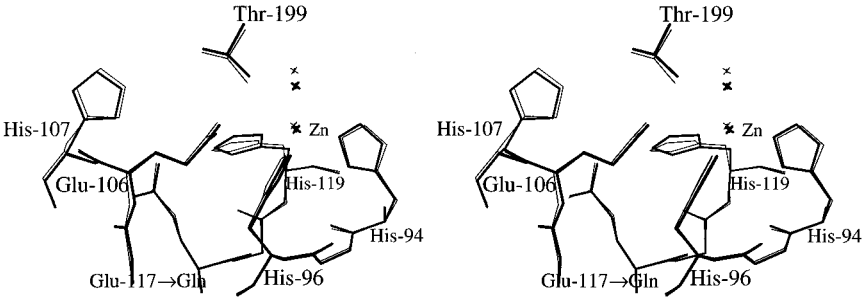

Figure 8. Least-squares superposition of Glu-117 → Gln CAII (thick bonds) and wild-type CAII (thin bonds). Even though the Glu-117 → Gln substitution is essentially isosteric, it results in a severe loss of catalytic activity and a dramatic increase in the kinetics of protein−zinc association.

The zinc−oxygen separation in the Thr-199 → Glu variant is 2.2 Å, which yields significantly enhanced affinity despite less-favorable *anti* coordination stereochemistry.

Notably, a 20-fold decrease in zinc affinity is measured in Thr-199 → His CAII; the three-dimensional structure of this variant reveals that the engineered side chain does not coordinate to zinc.[85] The loss of the Thr-199 hydrogen bond to zinc-bound hydroxide (i.e., the loss of an indirect ligand) therefore causes the 20-fold loss of zinc affinity. Thus, there are limits to the compensatory plasticity that facilitates the coordination of engineered ligands to zinc, and for His-199 there appears to be a steric barrier to achieving metal coordination with optimal ligand−metal separation and stereochemistry.

Altering the Indirect Zinc Ligands. The direct protein ligands of transition-metal sites in proteins are typically nested in a hydrogen bond network.[53,68] Hence, the entire protein, in addition to the direct metal ligands, may affect the stability and chemical properties of the protein−metal complex by modulating hydrogen bond−ligand−metal interactions. In wild-type CAII, zinc has four such indirect or second shell ligands (Figure 4): His-94 donates a hydrogen bond to the carboxamide side chain of Gln-92, His-119 donates a hydrogen bond to the carboxylate side chain of Glu-117, His-96 donates a hydrogen bond to the backbone carbonyl oxygen of Asn-244, and zinc-bound hydroxide donates a hydrogen bond to the hydroxyl side chain of Thr-199.

In the preceding section we have just seen that the deletion of the hydrogen bond between zinc-bound solvent and Thr-199 results in a 20-fold diminution of protein−zinc affinity. Correspondingly, the dissection of hydrogen bonds involving Gln-92 and Glu-117 reveals that hydrogen bonds to direct metal ligands exert a subtle yet important influence on protein−zinc affinity, and the pK_a and reactivity of zinc-bound solvent.[65,86,87] Surprisingly, these hydrogen bonds also have a profound influence on protein−zinc equilibration kinetics, and these experiments have yielded CAII variants that equilibrate with zinc rapidly enough for use in a real-time zinc biosensor.

Substitution of various amino acid residues for Gln-92 and Glu-117 reveals that the pK_a (and hence the reactivity) of zinc-bound solvent is adjustable within ±1 pK_a unit depending upon the charge of the indirect ligands (Table 3). In general, the effects of indirect ligands on the pK_a of zinc-bound water are more subtle than effects observed for direct ligands (Table 1). The more that zinc interacts with negative charge from its direct *and* indirect ligands, the less it is capable of stabilizing zinc-bound hydroxide. This results in an elevated pK_a for zinc-bound water.

Table 3. Properties of Selected CA Variants with Altered Indirect Zinc Ligands[a]

variant	k_{cat}/K_M (μM^{-1} s^{-1})	pK_a	zinc K_d (pM)
wild type	110	6.8	4
Q92A	29	6.8	18
Q92L	30	6.4	30
Q92N	27	6.9	5
Q92E	12	7.7	5
E117A	19	6.9	40
E117Q	0.002	≥9.9	4,400
T199A	1.1	8.3	60

[a] Data from refs 43 and 65.

Intriguingly, the direct metal ligands recruit hydrogen-bonding partners from other protein sites, or even solvent, upon the deletion of their wild-type hydrogen bond partners.[86] This suggests that direct metal ligands structurally and electrostatically require hydrogen bond partners for optimal protein−metal affinity, and measurements by Kiefer and colleagues[65] suggest that indirect protein−metal interactions via hydrogen bonds to metal ligands may contribute a factor of up to 10^4 to protein−metal affinity. Indirect ligands preorient the metal binding site for metal complexation and thereby minimize the conformational entropic cost of fixing the metal ligands in the metal binding step. Additionally, indirect ligands enhance the electrostatic interaction between the protein and its direct ligands.

Indirect ligands also play an important role in modulating protein−zinc association kinetics; remarkably, substitutions for Glu-117 (Glu-117 → Ala, Glu-117 → Asp, and Glu-117 → Gln) increase the zinc association rate constant 10^2−10^4-fold.[65,87] X-ray crystallographic studies of these variants reveal that they are nevertheless isostructural with the wild-type enzyme (Figure 8).[86,87] However, the Glu-117 → Gln variant has an elevated pK_a for the zinc−water ligand as well as drastically reduced catalytic activity and acetazolamide affinity, even though the structure of the complex with acetazolamide, a transition state analogue, is identical to that of the wild type.[87] Given that the wild-type and variant enzymes are isostructural, we propose that the loss of catalytic activity arises from the reversed polarity of the hydrogen bond between residue 117 and zinc ligand His-119, which requires that the side chain of His-119 be stabilized as the negatively-charged imidazolate (Figure 9).

(87) Huang, C.-c.; Lesburg, C. A.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **1996**, *35*, 3439−3446.

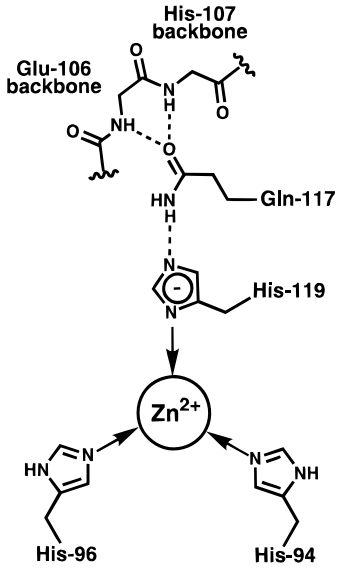

Figure 9. Hydrogen bond scheme proposed for Glu-117 → Gln CAII.

Concluding Remarks

Structure−function relationships are now better defined for the carbonic anhydrases than for any other family of enzymes, and applications of techniques in molecular and structural biology uniquely illuminate structural aspects of the catalytic mechanism. It is intriguing that the protein scaffolding of CAII is also readily adaptable to functions other than CO_2 hydration.[15,64] However, a detailed understanding of the structure, function, and stability of the CAII zinc binding site is a prerequisite for understanding (or predicting) its evolution by Nature or by design.

The molecular dissection of the CAII zinc binding site yields five key lessons pertinent to the evolution of structure, affinity, and function in a protein−zinc binding site: (1) Despite a chemically-diverse variety of alterations and augmentations to the direct metal ligands, the zinc coordination polyhedron usually remains tetrahedral even if nonprotein ligands must be recruited from solvent,[76] or if a nonprotein ligand must be displaced by an engineered protein ligand.[84,85] This behavior is consistent with the facile tetrahedral coordination of Zn^{2+}.[6,85,88,89] (2) Significant determinants of ligand−zinc affinity are optimal ligand−metal separation and favorable side chain torsion angles for the engineered ligand. Ligand−metal stereochemistry appears to be an important secondary factor on the basis of structure−affinity relationships in Thr-199 variants.[85] Favorable ligand−metal stereochemistry, as universally observed in naturally-evolved zinc binding sites, nevertheless must "fine-tune" the stability and function of the bound metal ion.[53] (3) Conformational changes in the protein may accompany the alteration of direct ligands, and these changes may facilitate ligand−metal coordination if the energetic cost of such changes is less than the energetic benefit of the coordination interaction.[75,76] The magnitude of such structural changes will depend on the net change in size of the altered ligand as well as its environment (i.e., buried, partially-buried, solvent-exposed). Conversely, it is possible that a metal binding site can be constructed in a protein to achieve a conformational change which triggers a desired chemical or regulatory effect.[58,59] (4) The entire protein should be considered a "ligand" to its bound metal ion(s), including both direct *and* indirect ligands. Indirect ligands, such as carboxylate or carboxamide groups which accept hydrogen bonds from zinc-coordinated histidines (e.g., Glu-117−His-119 and Glu-92−His-94 pairs in CAII), play modest yet important roles in modulating protein−zinc affinity (a factor of ∼10 per indirect ligand[65]) and the pK_a of zinc-bound water ($\Delta pK_a \approx 0.1-0.2$ for structurally conservative alterations and 0.5−1.0 for alterations in the net charge of indirect ligands[65]). It is likely that analogous hydrogen bonds between the phosphate backbone of DNA and Cys_2His_2 (and possibly Cys_3His) zinc finger transcription factors similarly contribute to the stability of the protein−DNA complex.[68,86] (5) The neutral ligands[65,79,87] and the hydrogen-bonding network with the zinc-bound solvent molecule[24,39,40] modulate the electrostatic environment of the active site, as assayed by the pK_a, and are thus essential for the high catalytic activity of carbonic anhydrase.

In conclusion, we predict that the consideration of these five lessons in the design of *de novo* zinc binding sites should lead to the construction of high-affinity metal coordination polyhedra with desired properties, including high catalytic activity.

We thank the National Institutes of Health (Grants GM45614 to D.W.C. and GM40602 to C.A.F.) and the Office of Naval Research for their generous support of this work. Additionally, we thank our laboratory personnel for their outstanding achievements, which are specifically acknowledged throughout the text in our references to the primary literature. C. A. Lesburg and S. K. Nair are acknowledged for their assistance with the figures. Finally, we thank Drs. Ivano Bertini, Michael Klemba, Steve Marino, Lynne Regan, and David Silverman for critical comments.

AR9501232

(88) Berg, J. M.; Merkle, D. L. *J. Am. Chem. Soc.* **1989**, *111*, 3759−3761.
(89) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, 24, 946−950.