Redesigning the Zinc Binding Site of Human Carbonic Anhydrase II: Structure of a His_2Asp–Zn^{2+} Metal Coordination Polyhedron

Laura L. Kiefer,‡ Joseph A. Ippolito,† Carol A. Fierke,*,‡ and David W. Christianson*,†

Department of Chemistry
University of Pennsylvania
Philadelphia, Pennsylvania 19104-6323
Department of Biochemistry
Duke University Medical Center, Box 3711
Durham, North Carolina 27710

Received September 1, 1993

The goal of protein design is to construct a stable protein molecule having predictable structure and function. Although some success toward this end has been reported,[1] the folding problem (i.e., the unknown relationship between primary and tertiary protein structure) hinders the design and construction of *de novo* proteins. Therefore, we currently focus on protein *re*design—the rational alteration of selected regions of a stable, preexisting protein scaffolding in order to confer new functional properties upon that scaffolding. Ultimately, the results of protein redesign studies will be applied toward the *de novo* design of proteins with novel functions. Implicit requirements of protein design and redesign experiments are the availabilities of both well-characterized functional properties and high-resolution protein structures determined by X-ray crystallographic methods.

Our current experiments in protein redesign focus on the chemistry and structure of transition-metal binding sites. Previously, the redesign of copper binding sites in copper–zinc superoxide dismutase[2] and blue copper proteins such as azurin[3] yielded metal sites with altered spectroscopic and catalytic properties. Here, the avid zinc binding site of human carbonic anhydrase II (CAII, K_d = 4 pM)[4] is our paradigm. The His_3–Zn^{2+} motif of this metalloenzyme is the target for *de novo* protein (re)design experiments,[1,5-7] which are particularly illuminating when accompanied by high-resolution structural analysis of the redesigned paradigm.[4,8,9] Here, we report the three-dimensional structure of the His-94→Asp (H94D) variant of CAII, which contains the first successfully-engineered tetrahedral zinc coordination polyhedron of composition His_2Asp–Zn^{2+}. In addition to its resemblance to the His_2Asp–Zn^{2+} site in alkaline phosphatase, the His_2Asp–Zn^{2+} motif engineered into CAII is chemically comparable to the His_2Glu–Zn^{2+} sites found in zinc proteases such as carboxypeptidase A and thermolysin.[10]

The H94D variant of CAII was produced using oligonucleotide-directed mutagenesis of the cloned CAII gene and purified to ≥95% homogeneity by chromatography on DEAE–Sephacel and S-Sepharose resins.[8] The effect of replacement of the histidine zinc ligand with asparate on metal binding was assessed by measuring a zinc dissociation constant (K_d) using equilibrium dialysis.[4,8] The zinc K_d = 15 ± 5 nM determined for H94D at pH 7 is increased ~10^4-fold relative to K_d = 4 ± 1 pM for wild-type CAII at the same pH. However, the zinc affinity is still significantly higher than that of other designed metal binding sites with three protein ligands.[5-7]

Wild-type CAII catalyzes the hydration of CO_2 to HCO_3^- and a H^+ with a second-order rate constant that approaches the diffusion control limit; CAII will also catalyze the hydrolysis of some esters.[11] These reactions proceed through nucleophilic attack of zinc-bound hydroxide on the carbonyl carbon of the substrate. The pK_a of the zinc-bound water increases from 6.8 in wild-type CAII to ≥9.6 in H94D CAII, as determined by measuring the pH dependence of CAII-catalyzed *p*-nitrophenyl acetate (PNPA) hydrolysis.[11,12] The second-order rate constant for PNPA hydrolysis by H94D decreases ~7-fold relative to wild-type CAII from 2500 to ≥365 M^{-1} s^{-1}. Steady-state kinetic parameters for CO_2 hydration were measured for H94D in TAPS buffer at pH 8.9[13] and are listed with the wild-type values in parentheses: k_{cat} = 8 ± 4 (1000) ms^{-1}, k_{cat}/K_M = 0.11 ± 0.01 (100) μM^{-1} s^{-1}, and K_M = 75 ± 40 (8) mM. The large decreases in activity for CO_2 hydration reflect both the decreased catalytic efficiency and the increased pK_a.

The structure of H94D CAII was determined by X-ray crystallographic methods, and crystals were prepared as described.[8,9] X-ray data from one of these crystals were collected to a limiting resolution of 2.3 Å on a Siemens multiwire area detector mounted on a Rigaku RU-200 X-ray generator. Raw data frames were processed using BUDDHA,[14] and data reduction was accomplished with PROTEIN.[15,16] The protein model was adjusted to fit electron density maps using the graphics program CHAIN[17] (maps were calculated by routines contained in X-PLOR[18]), and the model was refined with PROLSQ.[19]

* Authors to whom correspondence should be addressed.
† University of Pennsylvania. Telephone: (215) 898-5714. FAX: (215) 573-2201.
‡ Duke University Medical Center. Telephone: (919) 684-2557. FAX: (919) 684-5040.

(1) Handel, T. M.; Williams, S. A.; DeGrado, W. F. *Science* **1993**, *261*, 879–885.
(2) (a) Lu, Y.; Gralla, E. B.; Roe, J. A.; Valentine, J. S. *J. Am. Chem. Soc.* **1992**, *114*, 3560–3562. (b) Lu, Y.; LaCroix, L. B.; Lowery, M. D.; Solomon, E. I.; Bender, C. J.; Peisach, J.; Roe, J. A.; Gralla, E. B.; Valentine, J. S. *J. Am. Chem. Soc.* **1993**, *115*, 5907–5918.
(3) (a) Murphy, L. M.; Strange, R. W.; Karlsson, B. G.; Lundberg, L. G.; Pascher, T.; Reinhammer, B.; Hasnain, S. S. *Biochemistry* **1993**, *32*, 1965–1975. (b) Karlsson, B. G.; Nordling, M.; Pascher, T.; Tsai, L.-C.; Sjölin, L.; Lundberg, L. G. *Protein Eng.* **1991**, *4*, 343–349. (c) Chang, T. K.; Iverson, S. A.; Rodrigues, C. G.; Kiser, C. N.; Lew, A. Y. C.; Germanas, J. P.; Richards, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1325–1329. (d) den Blaauwen, T.; Canters, G. W. *J. Am. Chem. Soc.* **1993**, *115*, 1121–1129. (e) Romero, A.; Hoitink, C. W. G.; Nar, H.; Huber, R.; Messerschmidt, A.; Canters, G. W. *J. Mol. Biol.* **1993**, *229*, 1007–1021.
(4) Kiefer, L. L.; Krebs, J. F.; Paterno, S. A.; Fierke, C. A. *Biochemistry* **1993**, *32*, 9896–9900.
(5) (a) Iverson, B. L.; Iverson, S. A.; Roberts, V. A.; Getzoff, E. D.; Tainer, J. A.; Benkovic, S. J.; Lerner, R. A. *Science* **1990**, *249*, 659–662. (b) Roberts, V. A.; Iverson, B. L.; Iverson, S. A.; Benkovic, S. J.; Lerner, R. A.; Getzoff, E. D.; Tainer, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6654–6658. (c) Wade, W. S.; Koh, J. S.; Han, N.; Hoekstra, D. M.; Lerner, R. A. *J. Am. Chem. Soc.* **1993**, *115*, 4449–4456. (d) Pessi, A.; Bianchi, E.; Crameri, A.; Venturini, S.; Tramontano, A.; Sollazzo, M. *Nature* **1993**, *362*, 367–369.
(6) (a) Ghadiri, M. R.; Choi, C. *J. Am. Chem. Soc.* **1990**, *112*, 1630–1632. (b) Ruan, F.; Chen, Y.; Hopkins, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 9403–9404. (c) Suh, S.-S.; Haymore, B. L.; Arnold, F. H. *Protein Eng.* **1991**, *4*, 301–305.
(7) (a) Handel, T.; DeGrado, W. F. *J. Am. Chem. Soc.* **1990**, *112*, 6710–6711. (b) Regan, L.; Clarke, N. D. *Biochemistry* **1990**, *29*, 10878–10883.
(8) (a) Alexander, R. S.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 1510–1518. (b) Kiefer, L. L.; Ippolito, J. A.; Fierke, C. A.; Christianson, D. W., manuscript in preparation.
(9) Ippolito, J. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 9901–9905.
(10) (a) Kim, E. E.; Wyckoff, H. W. *J. Mol. Biol.* **1991**, *218*, 449–464. (b) Rees, D. C.; Lewis, M.; Lipscomb, W. N. *J. Mol. Biol.* **1983**, *168*, 367–387. (c) Holmes, M. A.; Matthews, B. W. *J. Mol. Biol.* **1982**, *160*, 623–639.
(11) (a) Pocker, Y.; Sarkanen, S. *Adv. Enzymol.* **1978**, *47*, 149–274. (b) Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30–36.
(12) Krebs, J. F.; Ippolito, J. A.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* in press.
(13) Khalifah, R. G. *J. Biol. Chem.* **1971**, *246*, 2561–2573.
(14) Durbin, R. M.; Burns, R.; Moulai, J.; Metcalf, P.; Freymann, D.; Blum, M.; Anderson, J. E.; Harrison, S. C.; Wiley, D. C. *Science* **1986**, *232*, 1127–1132.
(15) Steigemann, W. Ph.D. Thesis, Max-Planck-Institut für Biochemie, 8033 Martinsreid bei München, Germany, 1974.
(16) The *R*-factor for scaling and merging the X-ray data from one crystal was 0.086, where $R = \sum|I_{hi} - \langle I_h\rangle|/\sum\langle I_h\rangle$; I_{hi} = scaled intensity for reflection *h* in data set *i*, $\langle I_h\rangle$ = average intensity for reflection *h* calculated from replicate data.
(17) Sack, J. S. *J. Mol. Graphics* **1988**, *6*, 224–225.
(18) Brünger, A. T.; Kuriyan, J.; Karplus, J. *Science* **1987**, *235*, 458–460.
(19) Hendrickson, W. A. *Methods Enzymol.* **1985**, *115*, 252–270.

0002-7863/93/1515-12581$04.00/0 © 1993 American Chemical Society

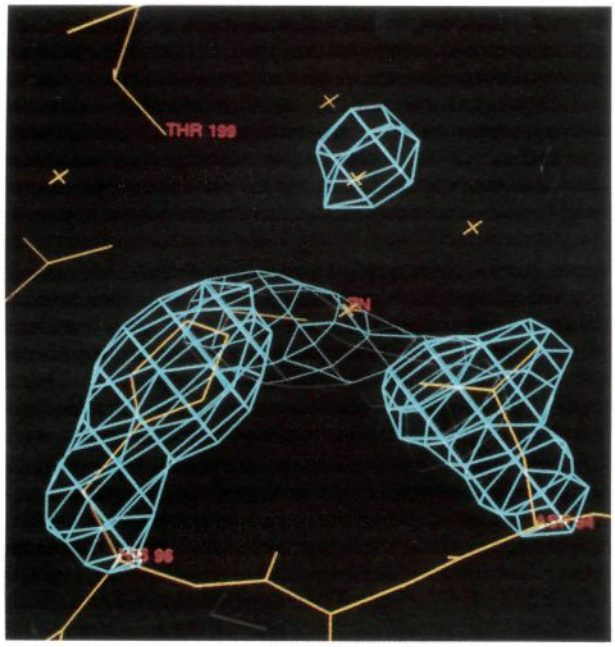

Figure 1. Difference electron density map generated with Fourier coefficients $|F_o| - |F_c|$ and phases calculated from the final H94D CAII model less the side chains of Asp-94, His-96, His-119, and the non-protein zinc ligand. The map is contoured at 3.5σ, and refined atomic coordinates are superimposed; Asp-94, His-96, Thr-199, and Zn^{2+} are indicated.

Refinement converged smoothly to a final crystallographic *R*-factor of 0.170.[20]

The electron density map of Figure 1 reveals tetrahedral zinc coordination by Asp-94, His-96, His-119, and a non-protein ligand. The engineered carboxylate group of Asp-94 coordinates to zinc with *syn*-unidentate stereochemistry ($O_\delta1$–Zn^{2+} separation = 2.1 Å), which is consistent with expectations based on analyses of small molecule[21] and protein[22,23] crystal structures. The C_γ–$O_\delta1$–Zn^{2+} angle is 133°, and the interaction is almost precisely within the plane of the carboxylate group. In wild-type CAII, His-94 donates a hydrogen bond to the side-chain carbonyl of Gln-92. The $O_\delta2$ atom of Asp-94 is within hydrogen-bonding distance (2.9 Å) of Gln-92; additionally, $O_\delta2$ also forms a strong *syn* hydrogen bond (2.6 Å) with an active-site water molecule, which in turn hydrogen bonds to the non-protein zinc ligand.

Apart from the zinc coordination polyhedron, the structure of H94D CAII does not differ significantly from that of wild-type CAII[24] (root mean square deviation of C_α coordinates = 0.2 Å). The superposition of H94D and wild-type CAII shows that the position of the engineered Asp-94 side chain is isosteric with that of the wild-type His-94 side chain, and the β-sheet superstructure containing Asp-94 is not disrupted. However, because the His→Asp substitution is not completely isosteric with respect to zinc-coordinating atoms (the N_ϵ atom of His-94 coordinates to zinc in wild-type CAII, whereas the $O_\delta1$ of Asp-94 coordinates to zinc in the variant), the zinc ion moves nearly 1 Å toward Asp-94. Consequently, His-96 and His-119 rotate slightly to maintain metal coordination, and the His-96 N_ϵ–Zn^{2+} separation increases from 2.1 to 2.4 Å. These structural changes undoubtedly contribute to decreased protein–zinc affinity. Structural changes in the apoenzyme may also contribute, and these effects incur a net free energy cost of 4.8 kcal mol^{-1}.

The electron density map of the H94D CAII active site reveals changes in the binding mode of the non-protein zinc ligand which is modeled and refined satisfactorily as a water molecule, although we cannot exclude the possibility of ammonia or disordered Tris binding.[25] Although Tris does not coordinate to the zinc ion of wild-type CAII, other nitrogen ligands are known to coordinate by displacing the zinc-bound solvent molecule.[26,27]

In H94D CAII, the non-protein zinc ligand forms a weak hydrogen bond with the side chain of Thr-199 (O–O separation = 3.3 Å); this interaction normally orients the zinc-bound hydroxyl group for substrate attack and maintains the low pK_a of zinc–H_2O.[12,28,29] Based on a comparison between the properties of T199A[12,29] and H94D CAIIs, we conclude that a majority of the decrease in catalytic efficiency and a portion of the increase in pK_a arises from a weakened hydrogen bond between Thr-199 and zinc-bound water in H94D CAII. It is likely that the decreased electrostatic stabilization of both the ground-state zinc–hydroxide and the transition state, caused by substitution of a negatively charged aspartate for a neutral histidine ligand, also affects these properties.

In summary, we have successfully replaced a naturally occurring histidine zinc ligand with an aspartate residue in the redesign of the CAII zinc binding site. The properties of H94D CAII demonstrate that: (1) the hydrogen bond network involving zinc-bound solvent is crucial for catalysis; (2) the chemical nature of the protein zinc ligands dictates the pK_a of zinc-bound solvent with little effect on its reactivity; and (3) the protein structure is sufficiently plastic to allow subtle rearrangements which accommodate metal binding with optimal stereochemistry. Furthermore, this work represents a first step toward conferring new catalytic functions on CAII as well as illuminating structure–function relationships that will facilitate the design of *de novo* metalloproteins with specific catalytic functions.[30,31]

(20) $R = \sum||F_o| - |F_c||/\sum|F_o|$; $|F_o|$ and $|F_c|$ are the observed and calculated structure factors, respectively.

(21) Carrell, C. J.; Carrell, H. L.; Erlebacher, J.; Glusker, J. P. *J. Am. Chem. Soc.* **1988**, *110*, 8651–8656.
(22) Christianson, D. W.; Alexander, R. S. *J. Am. Chem. Soc.* **1989**, *111*, 6412–6419.
(23) Chakrabarti, P. *Protein Eng.* **1990**, *4*, 49–56.
(24) Alexander, R. S.; Nair, S. K.; Christianson, D. W. *Biochemistry* **1991**, *30*, 11064–11072.
(25) (a) Higaki, J. N.; Haymore, B. L.; Chen, S.; Fletterick, R. J.; Craik, C. S. *Biochemistry* **1990**, *29*, 8582–8586. (b) McGrath, M. E.; Haymore, B. L.; Summers, N. L.; Craik, C. S.; Fletterick, R. J. *Biochemistry* **1993**, *32*, 1914–1919.
(26) Bertini, I.; Luchinat, C.; Scozzafava, A. *J. Am. Chem. Soc.* **1977**, *99*, 581–584.
(27) (a) Bertini, I.; Luchinat, C. *Acc. Chem. Res.* **1983**, *16*, 272–279. (b) Nair, S. K.; Christianson, D. W. *Eur. J. Biochem.* **1993**, *213*, 507–515. (c) Mangani, S.; Liljas, A. *J. Mol. Biol.* **1993**, *232*, 9–14.
(28) Merz, K. M., Jr. *J. Mol. Biol.* **1990**, *214*, 799–802.
(29) (a) Xue, Y.; Liljas, A.; Jonsson, B.-H.; Lindskog, S. *Proteins: Struct., Funct. Genet.* **1993**, *17*, 93–106. (b) Liang, Z.; Xue, Y.; Behravan, G.; Jonsson, B.-H.; Lindskog, S. *Eur. J. Biochem.* **1993**, *211*, 821–827.
(30) Atomic coordinates of H94D CAII have been deposited into the Brookhaven Protein Data Bank with the accession code 1CVC; see: 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.
(31) This work was supported by the Office of Naval Research (D.W.C.) and the NIH (Grant GM40602 to C.A.F.). C.A.F. acknowledges an American Heart Association Established Investigatorship and a David and Lucile Packard Foundation Fellowship, and D.W.C. acknowledges an Alfred P. Sloan Research Fellowship and a Camille and Henry Dreyfus Teacher-Scholar Award. J.A.I. is supported in part by NIH Training Grant GM08275.