X-ray Crystallographic Studies of Engineered Hydrogen Bond Networks in a Protein−Zinc Binding Site

Charles A. Lesburg and David W. Christianson*

Contribution from the Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323

Received January 24, 1995[⊗]

Abstract: Transition metal binding sites in proteins are typically comprised of 3−4 protein ligands, most of which are also embedded in hydrogen bond networks. For instance, in human carbonic anhydrase II (CAII) the carboxamide side chain of Q92 accepts a hydrogen bond from H94, the carboxylate side chain of E117 accepts a hydrogen bond from H119, and the backbone carbonyl oxygen of N244 accepts a hydrogen bond from H96. In order to probe the structural importance of these hydrogen bond networks, we have determined the three-dimensional structures of Q92A, Q92N, Q92E, Q92L, and E117A CAIIs by X-ray crystallographic methods. When interpreted in light of functional measurements (catalytic activity, protein-zinc affinity) made by Kiefer and colleagues (Kiefer, L. L.; Paterno, S. A.; Fierke, C. A. *J. Am. Chem. Soc.*, preceding paper in this issue), these high-resolution structures allow for detailed structure−function correlations which illuminate the general role of hydrogen bond networks with the second shell of residues surrounding protein−metal binding sites. Due to their structural and electrostatic contributions, these second shell residues, i.e., "indirect" metal ligands, fine-tune the pK_a and reactivity of zinc-bound solvent; additionally, these residues may contribute a factor of up to 10^4 to protein−metal affinity in a tetracoordinate metal site. It is therefore imperative that indirect metal ligands be considered in *de novo* designs of avid protein−metal binding sites. Additionally, indirect ligand−direct ligand−metal networks are important for protein−nucleic acid recognition, e.g., in the C_2H_2 class of zinc-finger transcription factors.

Introduction

Transition metal sites in proteins, such as those specific for zinc binding, are typically comprised of 3−4 direct protein ligands plus a hydrogen bond network in which these ligands are nested.[1] Hence the entire protein, in addition to the direct metal ligands, may affect the structure and function of a bound metal ion by modulating hydrogen bond−ligand−metal interactions.[2,3] For example, D124 in Cu,Zn superoxide dismutase plays a role in metal binding by forming hydrogen bonds with histidine ligands to zinc and copper.[4] When these hydrogen bonds are destroyed in the D124N and D124G variants, the enzyme loses virtually all of its zinc binding ability.[5] Similarly, D235 in cytochrome *c* peroxidase, which forms a hydrogen bond to proximal heme ligand H175, is implicated in maintaining an open coordination site on the distal side of the active site heme: D235 and its surroundings increase the liganding strength of H175 so much that the porphyrin iron lies out of the heme plane in order to optimize H175−Fe^{2+} coordination.[6,7] The neutralization of side chain charge in D235N cytochrome *c* peroxidase results in a 5-fold reduction of catalytic activity, partly due to the coordination of a solvent molecule to the distal side of the heme.[6,7] This substitution also affects the electronic coupling of the tryptophan (W191) free radical to the heme as determined in recent spectroscopic experiments.[8] Hence, hydrogen bond networks with metal ligands may be important for the structure and function of certain metalloproteins; however, there is a dearth of information regarding the systematic structural and functional dissection of such interactions.

Structure−function relationships in protein−zinc binding sites are most extensively studied in the prototypical zinc metalloenzyme carbonic anhydrase II (CAII).[9−16] The biological function of CAII is the hydration of carbon dioxide to form bicarbonate ion and a proton; the rate of this reaction approaches that of diffusion control and is limited by a proton transfer step which regenerates the zinc−hydroxide form.[1,17−21] Thus, CAII is the most appropriate paradigm for dissecting the structural and functional roles of hydrogen bond networks between "second shell" residues and protein metal ligands. The structure of CAII from human blood has been determined by X-ray crystal-

[⊗] Abstract published in *Advance ACS Abstracts*, June 15, 1995.

(1) Christianson, D. W. *Adv. Protein Chem.* **1991**, *42*, 281−355.
(2) Christianson, D. W.; Alexander, R. S. *J. Am. Chem. Soc.* **1989**, *111*, 6412−6419.
(3) Christianson, D. W.; Alexander, R. S. *Nature* **1990**, *346*, 225.
(4) Tainer, J. A.; Getzoff, E. D.; Beem, K. M.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1982**, *160*, 181−217.
(5) Banci, L.; Bertini, I.; Cabelli, D. E.; Hellewell, R. A.; Tung, J. W.; Viezzoli, M. S. *Eur. J. Biochem.* **1991**, *196*, 123−128.
(6) Wang, J.; Mauro, J. M.; Edwards, S. L.; Oatley, S. J.; Fishel, L. A.; Ashford, V. A.; Xuong, N.-h.; Kraut, J. *Biochemistry* **1990**, *29*, 7160−7173.
(7) Vitello, L. B.; Erman, J. E.; Miller, M. A.; Mauro, J. M.; Kraut, J. *Biochemistry* **1992**, *31*, 11524−35.
(8) Goodin, D. B.; McRee, D. E. *Biochemistry* **1993**, *32*, 3313−24.
(9) Alexander, R. S.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 1510−1518.
(10) Ippolito, J. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 9901−9905.
(11) Kiefer, L. L.; Ippolito, J. A.; Fierke, C. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1993**, *115*, 12581−12582.
(12) Kiefer, L. L.; Krebs, J. F.; Paterno, S. A.; Fierke, C. A. *Biochemistry* **1993**, *32*, 9896−9900.
(13) Kiefer, L. L.; Fierke, C. A. *Biochemistry* **1994**, *33*, 15233−15240.
(14) Ippolito, J. A.; Christianson, D. W. *Biochemistry* **1994**, *33*, 15241−15249.
(15) Ippolito, J. A.; Baird, T. T.; McGee, S. A.; Christianson, D. W.; Fierke, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, in press.
(16) Kiefer, L. L.; Paterno, S. A.; Fierke, C. A. **1995**, *117*, 6831−6837.
(17) Coleman, J. E. In *Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhauser: Boston, 1986; pp 49−58.
(18) Lindskog, S. In *Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhauser: Boston, 1986; pp 307−316.
(19) Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30−36.
(20) Silverman, D. N. *Can. J. Bot.* **1991**, *69*, 30−36.
(21) Lindskog, S.; Liljas, A. *Curr. Op. Struct. Biol.* **1993**, *3*, 915−920.

0002-7863/95/1517-6838$09.00/0

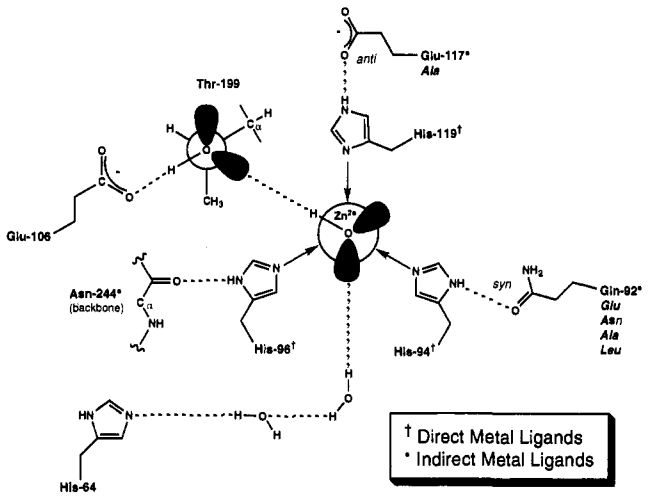

Figure 1. Schematic diagram of hydrogen bond networks in the active site of wild-type CAII. Three-dimensional structures of variants reported herein are indicated in italics at the site of substitution.

lographic methods[22] and refined at 1.54 Å resolution.[23] The active site of the enzyme is a cone-shaped cleft about 15 Å deep at the base of which resides the catalytically-required zinc ion. The imidazole side chains of H94, H96, and H119 coordinate to zinc, and hydroxide ion completes the tetrahedral metal coordination polyhedron; zinc-bound hydroxide is the catalytically-active nucleophile in the hydration of carbon dioxide.[24,25] In the wild-type enzyme, the zinc ligands are fully saturated by hydrogen bond networks with second shell residues (Figure 1): H94 donates a hydrogen bond to the carboxamide side chain of Q92; H119 donates a hydrogen bond to the carboxylate side chain of E117; H96 donates a hydrogen bond to the backbone carbonyl oxygen of N244; and zinc-bound hydroxide donates a hydrogen bond to the hydroxyl side chain of T199.[23] We refer to these second shell residues as "indirect" zinc ligands.[2]

Here, we report the three-dimensional structures of enzyme variants Q92A, Q92N, Q92E, Q92L, and E117A as determined by X-ray crystallographic methods to limiting resolutions of 1.8–2.15 Å. The goal of the current study is to illuminate the structural importance of the Q92–H94 and E117–H119 hydrogen bonds in view of the functional measurements made by Kiefer and colleagues.[16] In short, these particular hydrogen bonds have a subtle but important influence on protein–zinc affinity as well as the pK_a and reactivity of zinc-bound solvent. We note that structural studies of the hydrogen bond network involving the non-protein zinc ligand (hydroxide ion) have been reported previously.[10,15,26–28]

Experimental Section

Recombinant Q92A, Q92N, Q92E, Q92L, and E117A CAIIs were generously provided by Dr. Carol Fierke, Duke University. Crystallizations were performed by the sitting drop method. Typically, a 5–10 μL drop of precipitant buffer containing 50 mM Tris–Cl or Tris–SO_4 (titrated to pH 8.0 at 20 °C) and 50–75% saturated $(NH_4)_2SO_4$ (1.75–2.5 M) was added to a 5–10 μL drop containing 0.3 mM protein and 50 mM Tris–Cl or Tris–SO_4 (titrated to pH 8.0 at 20 °C) in the crystallization well. The outer well contained 1 mL of precipitant buffer. For Q92A and Q92N, the protein solution was saturated with methylmercury acetate in order to facilitate growth of diffraction-quality parallelepipedons.[29] However, this procedure resulted in the precipitation of the remaining variants. In the absence of mercury, the remaining variants grew as thin plates; accordingly, up to 10 mM β-D-glucohexylpyranoside was included in the crystallization wells to facilitate the formation of larger crystals suitable for X-ray diffraction analysis.[30] For Q92L CAII, the addition of dithiothreitol to the protein solution was necessary to obtain satisfactory crystals. All CAII variants crystallized isomorphously with the wild-type enzyme in monoclinic space group $P2_1$ and exhibited unit cell parameters a = 42.7 Å, b = 41.7 Å, c = 73 Å, and β = 104.6°.

Crystals of each variant were mounted in 0.5 or 0.7 mm glass capillaries with a small portion of mother liquor and sealed with wax. For crystals of Q92A, Q92N, Q92E, and E117A CAIIs, a Siemens X-100A multiwire area detector mounted on a three-axis camera and equipped with Charles Supper double X-ray focusing mirrors was used for X-ray data acquisition. A Rigaku RU-200HB rotating anode X-ray generator operating at 45 kV (25–50 mA) supplied Cu Kα radiation (λ = 1.5418 Å). All data were collected at room temperature using the oscillation method. The crystal to detector distance was 10–12 cm and the detector swing angle 2θ was set at 20–30°. Data frames of 60 s (for parallelepipedons) or 120 s (for plates) duration with a 0.1° oscillation about ω were collected in total scans of at least 180° about ω. The limiting resolution of diffraction intensities for these crystals varied between 1.8 and 2.05 Å. Raw data frames were analyzed using BUDDHA[31] and replicate and symmetry-related reflections were merged using PROTEIN[32] or CCP4.[33,34] Final merging R-factors as well as other relevant data reduction statistics are reported in Table 1.

Although a sufficiently large crystal of Q92L CAII was grown, the intensity distribution statistics on data collected with the Siemens detector were unacceptable. However, intensity data for Q92L CAII collected on an R-AXIS IIC image plate detector proved to be excellent. Individual ϕ oscillation images spanning 3° were collected for 30 min each for a total sweep of 180°. A Rigaku RU-200HB rotating anode X-ray generator furnished Cu Kα radiation at 124 kV/40 mA (λ = 1.5418 Å). The crystal-to-detector distance was set to 10 cm and the swing angle 2θ was set to 0°. The crystal orientation matrix was initially obtained using REFIX[35] and reflections subsequently integrated with MOSFLM.[36] Data reduction was completed with CCP4.[33,34]

The starting coordinate set for the refinement of each variant was that of wild-type CAII[37] with the atoms of the variant side chain and active site water molecules within 10 Å of the bound metal ion deleted from the model. The structure of each CAII variant was refined by simulated annealing with energy minimization as implemented in X-PLOR.[38] When the crystallographic R-factor dropped below 0.20 in each refinement, the variant side chain as well as active site water

(22) Liljas, A.; Kannan, K. K.; Bergsten, P.-C.; Waara, I.; Fridborg, K.; Strandberg, B.; Carlbom, U.; Jarup, L.; Lovgren, S.; Petef, M. *Nature (London), New Biol.* **1972**, *235*, 131–137.
(23) Håkansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. *J. Mol. Biol.* **1992**, *227*, 186–190.
(24) Coleman, J. E. *J. Biol. Chem.* **1967**, *242*, 5212–5219.
(25) Lindskog, S.; Coleman, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 2505–2508.
(26) Krebs, J. F.; Ippolito, J. A.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* **1993**, *268*, 27458–66.
(27) Liang, Z.; Xue, Y.; Behravan, G.; Jonsson, B. H.; Lindskog, S. *Eur. J. Biochem.* **1993**, *211*, 821–7.
(28) Xue, Y.; Liljas, A.; Jonsson, B. H.; Lindskog, S. *Proteins: Struct., Funct., and Genet.* **1993b**, *17*, 93–106.
(29) Tilander, B.; Strandberg, B.; Fridborg, K. *J. Mol. Biol.* **1965**, *12*, 740–760.
(30) McPherson, A.; Koszelak, S.; Axelrod, H.; Day, J.; Williams, R.; Robinson, L.; McGrath, M.; Cascio, D. *J. Biol. Chem.* **1986**, *1986*, 1969–1975.
(31) 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.
(32) Steigemann, W. Thesis, Techniche Universität, München, Germany, 1974.
(33) Collaborative Computing Project, Number 4. *Acta Crystallogr.* **1994**, *D50*, 760–763.
(34) French, G. S.; Wilson, K. S. *Acta Crystallogr.* **1978**, *A34*, 517–525.
(35) Kabsch, W. *J. Appl. Crystallogr.* **1993**, *24*, 795–800.
(36) Nyborg, J.; Wonacott, A. J. In *The Rotation Method in Crystallography*; Arndt, U. W., Wonacott, A. J., Eds.; North-Holland: Amsterdam, 1977; pp 139–152.
(37) Alexander, R. S.; Nair, S. K.; Christianson, D. W. *Biochemistry* **1991**, *30*, 11064–11072.
(38) Brünger, A. T.; Kuriyan, J.; Karplus, M. *Science* **1987**, *235*, 458–460.

Table 1. Crystallographic Data Collection and Refinement Statistics for Carbonic Anhydrase II Variants

variant	Q92A	Q92N	Q92E	Q92L	E117A (Cl^-)	E117A (Br^-)
no. of crystals	1	1	1	1	1	1
no. of measured reflcns	42585	52142	29234	63267	65391	24157
no. of unique reflcns	19242	18858	15319	18287	17518	12926
max resolution (Å)	1.8	1.8	2.05	2.15	1.9	2.05
min resolution (Å)	8.0	6.5	8.0	8.0	8.0	8.0
R_{merge}[a]	0.10	0.10	0.071	0.067	0.065	0.087
no. of reflcns used in refinement	18151	17556	14543	12513	17196	12508
completeness of data (%)	79	77	87	93	88	80
R_{cryst}[b]	0.161	0.169	0.158	0.176	0.176	0.187
no. of water molecules in final cycle of refinement	230	203	168	169	175	175
root-mean-square deviation from ideal bond lengths (Å)	0.007	0.006	0.009	0.011	0.009	0.010
root-mean-square deviation from ideal bond angles (degree)	1.5	1.4	1.6	1.7	1.6	1.5
root-mean-square deviation from ideal dihedral angles (degree)	25.4	25.4	25.5	25.9	25.3	25.4
root-mean-square deviation from ideal improper angles (degree)	1.3	1.2	1.3	1.5	1.3	1.4

[a] R_{merge} for replicate reflections, $R = \sum|I_{hi} - \langle I_h\rangle|/\sum\langle I_h\rangle$; I_{hi} = intensity measured for reflection h in data set i, $\langle I_h\rangle$ = average intensity for reflection $\boldsymbol{h}$ calculated from replicate data. [b] Crystallographic R factor, $R = \sum||F_o| - |F_c||/\sum|F_o|$; F_o and F_c are the observed and calculated structure factors, respectively.

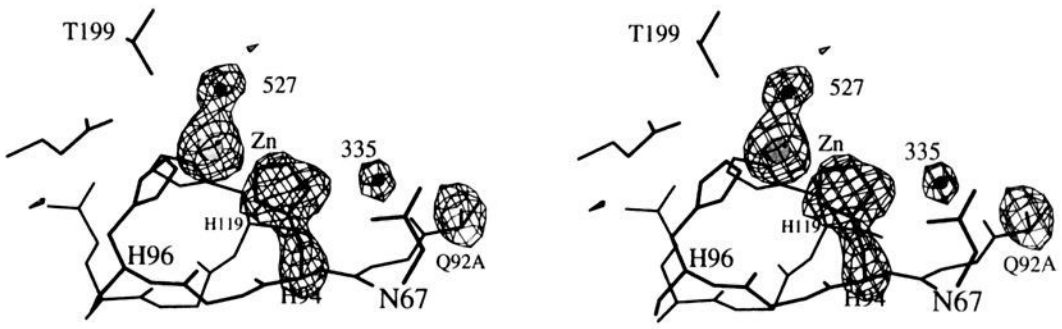

Figure 2. Q92A CAII: Difference Fourier ($|F_o| - |F_c|$) electron density map contoured at 3.5σ in which the C_α atoms and side chains of residues H94 and A92, Zn^{2+}, zinc-bound solvent (#527), and solvent molecule (#335) were omitted from the structure factor calculation.

molecules were modeled into electron density maps generated with Fourier coefficients $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ and phases calculated from the in-progress atomic model. The graphics software CHAIN,[39] installed on a Silicon Graphics Indigo workstation, was utilized for all map-fitting procedures. Final electron density maps and coordinate superpositions shown in Figures 2 through 9 were generated using MINIMAGE[40] and/or MOLSCRIPT.[41] Final coordinates have been submitted to the Protein Data Bank.[42]

During the refinement of each CAII variant, electron density maps were periodically calculated from the in-progress atomic model, and only minimal manual adjustments of atomic coordinates were found to be necessary. Refinement ultimately yielded structures with final crystallographic R-factors of 0.158−0.176 and excellent stereochemistry (Table 1). The rms error in atomic positions for all CAII variants was estimated to be 0.15−0.20 Å.[43]

Results

Overall, the refined structure of each CAII variant is very similar to that of the wild-type enzyme. Due to the higher quality and resolution of diffraction data measured from certain CAII variants, approximately twice as many ordered solvent molecules are observed in the corresponding electron density maps relative to the wild-type structure.[37] In some structures good electron density characterizes H4 and/or K261; these residues are included in the corresponding final models. However, if only the backbone atoms are interpretable in electron density maps, H4 and/or K261 are included in the final coordinate sets but the occupancies of side chain atoms are set to zero.

The structures of the five CAII variants reported herein share several common features. First, the hydrogen bond between zinc-bound hydroxide and the side chain of T199 is maintained in each variant. Second, despite the alteration of hydrogen bond networks with histidine zinc ligands, zinc−protein ligand coordination distances do not differ significantly from those of the wild-type enzyme. Additionally, regardless of the substitution at the indirect ligand site, a hydrogen bond to the direct zinc ligand is always maintained, sometimes by the unexpected recruitment of new hydrogen bond partners from the protein or from solvent. These structural consequences help to rationalize the measured properties of each variant.[16] In the remainder of this section we describe the structure of each CAII variant and the effects of each amino acid substitution on protein−zinc structure and function.

Q92A CAII. The difference electron density map in Figure 2 reveals that a well-ordered solvent molecule accepts a hydrogen bond from H94; the carboxamide side chain of Q92 in the wild-type enzyme ordinarily performs this role. This solvent molecule, #335, forms a hydrogen bond with excellent stereochemistry to the N_δ atom of H94: the $C_\gamma-N_\delta\cdots O^{335}$ angle is 130° and the $N_\delta-O^{335}$ separation is 2.7 Å. Moreover, the side chain of N67 moves from its wild-type position to donate a hydrogen bond to water #335 ($N_\delta-O^{335}$ separation = 3.1 Å). In addition, water #495 occupies the location formerly occupied by the side chain carboxamide nitrogen of Q92, and the $O^{495}-O^{336}$ separation is 2.8 Å. The least-squares superposition of wild-type and Q92A CAIIs in Figure 3 allows for a detailed comparison of the structural changes in the enzyme active site which accompanies the point mutation. Clearly, the enzyme structure and its solvation shell are sufficiently plastic to satisfy the hydrogen bond requirements of H94. Otherwise, the structure of Q92A CAII is essentially unchanged relative to the wild-type enzyme as indicated by the rms deviation of C_α atoms of 0.18 Å.

Since electron density corresponding to the non-protein zinc ligand is ellipsoidal in shape, this density is currently interpreted as two hydrogen bonded solvent molecules, #527 and #340

(39) Sack, J. S. *J. Mol. Graphics* **1988**, *6*, 224−225.
(40) Arnez, J. G. *J. Appl. Crystallogr.* **1994**, *27*, 649−653.
(41) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.
(42) 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.
(43) Luzzati, P. V. *Acta Crystallogr.* **1952**, *5*, 802−810.

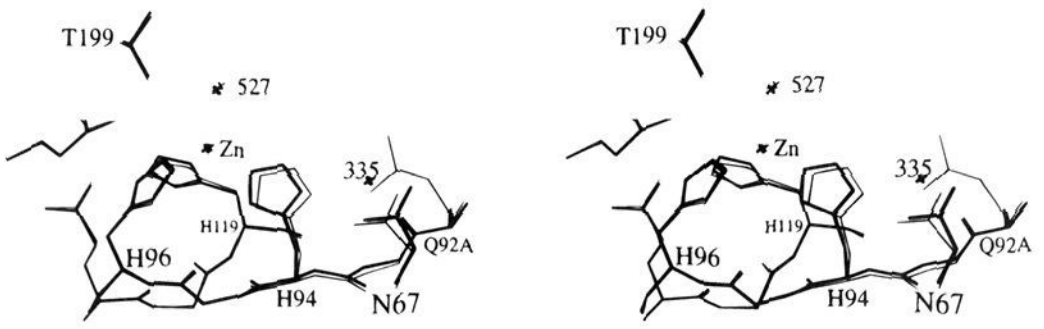

Figure 3. Least-squares superposition of Q92A CAII C_α atoms (thick lines) with those of wild-type CAII (thin lines).

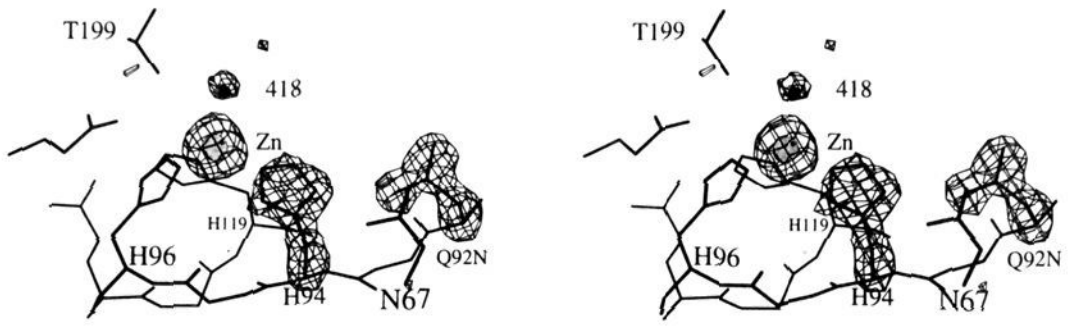

Figure 4. Q92N CAII: Difference Fourier ($|F_o| - |F_c|$) electron density map contoured at 3.5σ in which the C_α atoms and side chains of N92 and H94, Zn^2, and zinc-bound solvent (#418) were omitted from the structure factor calculation.

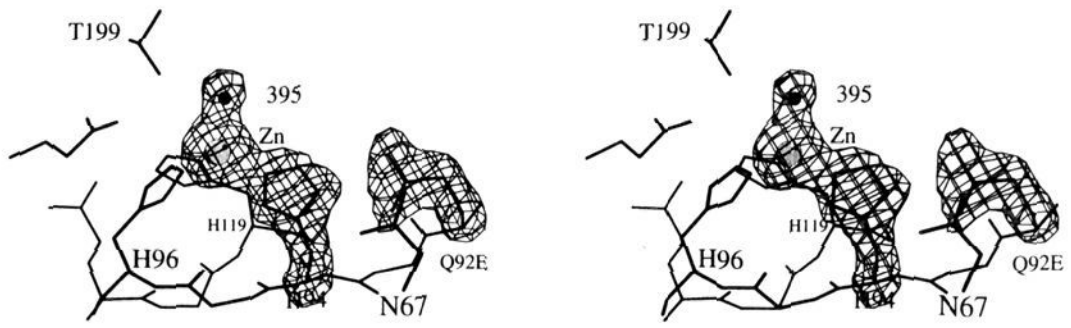

Figure 5. Q92E CAII: Difference Fourier ($|F_o| - |F_c|$) electron density map contoured at 3.5σ in which the C_α atoms and side chains of E92 and H94, Zn^{2+}, and zinc-bound solvent (#395) were omitted from the structure factor calculation.

(O^{527}–O^{340} separation = 2.9 Å). Solvent #527, zinc-bound hydroxide, is coordinated to zinc with a refined Zn^{2+}–O^{527} separation of 2.3 Å; this solvent molecule also hydrogen bonds to the hydroxyl side chain of T199 with an O^{527}–O_γ separation of 2.8 Å. The hydrogen bonded solvent network in the active site extends nearly 8 Å away from zinc-bound hydroxide. Importantly, the interactions of zinc-bound hydroxide, the catalytic nucleophile,[24] are identical to those observed in the wild-type enzyme.[37] Additionally H64, the catalytic proton shuttle, resides in the "in" conformation (i.e., directed toward zinc). This is the predominant rotamer found in the wild-type structure.[37,44,45]

Q92N CAII. Although the side chain of residue 92 is shortened by one methylene group in Q92N CAII, the carboxamide group nonetheless forms a hydrogen bond with H94 as revealed in the electron density map of Figure 4. However, this hydrogen bond exhibits poorer stereochemistry than that observed between H94 and Q92 in the wild-type enzyme. This interaction is achieved primarily by a conformational change of the N92 side chain relative to the wild-type side chain.

Ellipsoidal density characterizes the non-protein metal ligand of Q92N CAII, comparable to that observed in the Q92A variant. This density is interpreted as two solvent molecules, #418 and #486, which exhibit good hydrogen bond geometry at a separation of 2.8 Å. Zinc-bound hydroxide, solvent #418, has a Zn^{2+}–O^{418} separation of 2.2 Å and forms a hydrogen bond with the hydroxyl side chain of T199 (O_γ–O^{418} separation = 2.7 Å). Just as with other variants, crystallographically-observable solvent molecules in the active site extend 8 Å away from zinc-bound hydroxide.

The overall structure of this variant enzyme is essentially unchanged relative to wild-type CAII: the rms deviation of C_α atoms is 0.18 Å. Finally, we note that the catalytic proton shuttle (H64) predominantly occupies the "in" conformation as in the structure of the wild-type enzyme.[37,44,45]

Q92E CAII. This variant has a negatively-charged, isosteric carboxylate side chain substituted for the wild-type carboxamide. The refined three-dimensional structure indicates that the carboxylate at position 92 forms a more favorable *syn*-oriented[46] hydrogen bond with H94 than the carboxamide of the wild-type enzyme. An electron density map of the active site is shown in Figure 5. The side chain of E92 undergoes a conformational change from its wild type position to form a bifurcated hydrogen bond with H94 (Figure 6): the H94 N_δ atom is approximately equidistant from both carboxylate oxygens of E92 (N–$O_{\epsilon 1}$ separation = 3.1 Å, N–$O_{\epsilon 2}$ separation = 2.9 Å). Zinc-bound solvent, #395, is 2.1 Å from the zinc cation and 2.7 Å from the O_γ atom of T199. This virtually isosteric substitution does not perturb the hydrogen bond networks in the active site with the exception of H64 interactions as described below. The rms deviation of C_α atoms in this structure compared to the wild-type enzyme is 0.22 Å, indicative of no major structural differences.

Surprisingly, H64 resides primarily in the "out" position; that is, the imidazole side chain of this residue rotates about torsion angle χ_1 away from zinc. In the wild-type enzyme, H64 predominantly occupies the "in" conformation.[37,44,45] This effect may be a result of some subtle differences in active site solvent structure and/or the alteration of the electrostatic environment of the active site brought about by the Q92E substitution.

Q92L CAII. The electron density map of Q92L CAII

(44) Krebs, J. F.; Fierke, C. A.; Alexander, R. S.; Christianson, D. W. *Biochemistry* **1991**, *30*, 9153–9160.
(45) Nair, S. K.; Christianson, D. W. *J. Am. Chem. Soc.* **1991**, *113*, 9455–9458.
(46) Gandour, R. D. *Bioorg. Chem.* **1981**, *10*, 169–76.

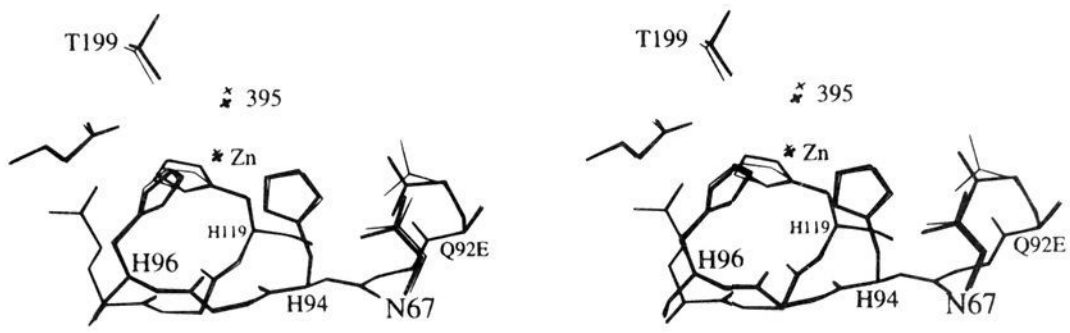

Figure 6. Least-squares superposition of Q92E CAII C_α atoms (thick lines) with those of wild-type CAII (thin lines).

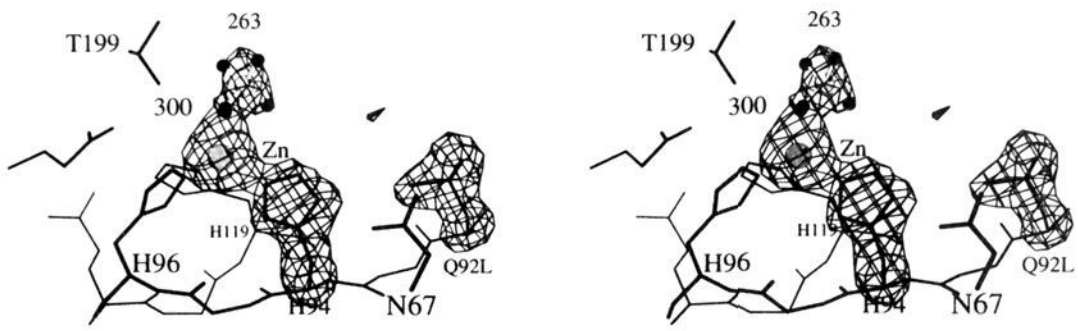

Figure 7. Q92L CAII: Difference Fourier ($|F_o| - |F_c|$) electron density map contoured at 3.5σ in which the C_α atoms and side chains of L92, H94, and N67, Zn^{2+}, zinc-bound solvent (#300), and zinc-bound sulfate (#263) were omitted from the structure factor calculation.

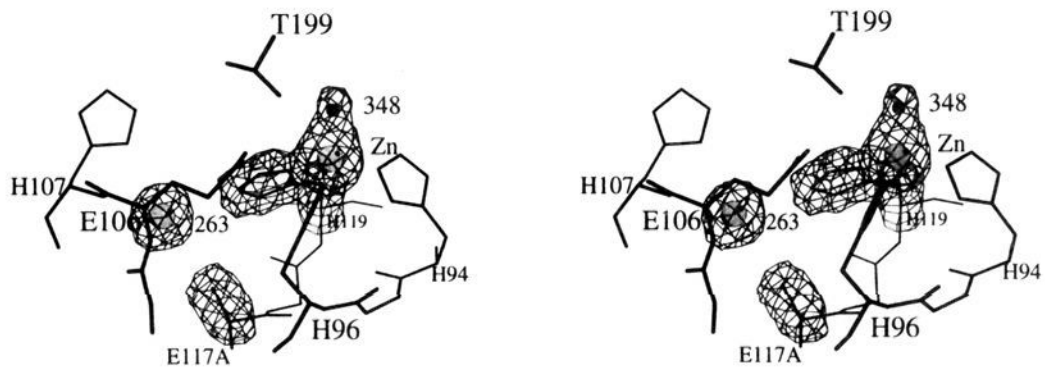

Figure 8. E117A: Difference Fourier ($|F_o| - |F_c|$) electron density map contoured at 3.5σ in which the C_α atoms and side chains of residues A117 and H119, Zn^{2+}, zinc-bound solvent (#348), and chloride ion (#263) were omitted from the structure factor calculation. Note that the backbone amide nitrogens of E106 and H107 are oriented toward the recruited chloride ion.

(Figure 7) reveals that a solvent molecule cannot be recruited to maintain the hydrogen bond to H94. However, the side chain of N67 (beyond hydrogen bonding range to H94 in the wild-type enzyme) rotates by 30° about χ_2 toward H94, and the imidazole ring of H94 pivots 15° about χ_2 toward N67, so that a hydrogen bond forms between $O_\delta{}^{67}$ and $N_\delta{}^{94}$ (O−N separation = 3.1 Å, $C_\gamma{}^{67}-O_\delta{}^{67}\cdots N_\delta{}^{94}$ angle = 146°). However, since the hydrogen bond acceptor ($O_\delta{}^{67}$) lies 1.1 Å out of the plane defined by the imidazole ring of H94, this hydrogen bond must be somewhat weaker than that between Q92 and H94 in the wild-type enzyme.

Electron density corresponding to the non-protein zinc ligand appears as a broad, 10σ peak centered approximately 3 Å from zinc. This peak refines satisfactorily as a partially-occupied (35%) sulfate ion, with one of its oxygen atoms coordinated to zinc ($O-Zn^{2+}$ separation = 1.9 Å), and a partially-occupied (65%) solvent molecule.

Just as in Q92E CAII, H64 undergoes a conformational change to the "out" position in Q92L CAII.[37,44,45] This may be a consequence of the small structural changes which facilitate the H67−H94 hydrogen bond, but an unambiguous molecular cause cannot be inferred from the structural data. Finally, the rms deviation of C_α atoms from their wild-type locations is 0.24 Å in this variant.

E117A CAII. A difference Fourier electron density map (Figure 8) reveals a 22σ peak in the former location of the $O_{\epsilon 2}$ atom of E117. This peak is interpreted and satisfactorily refined as a chloride ion, recruited from solvent to occupy the void resulting from the E117A substitution. In addition to the hydrogen bond with H119 ($N_\epsilon-Cl^-$ separation = 3.2 Å), this chloride ion accepts potential hydrogen bonds from the backbone NH groups of residues E106 and H107 ($N-Cl^-$separations of 3.3 and 3.2 Å, respectively), the N_δ of H107 ($N_\delta-Cl^-$separation = 3.1 Å), and an ordered solvent molecule, #376 ($O^{376}-Cl^-$-separation = 2.6 Å). Interestingly, when crystals of E117A are soaked in 10 mM KBr for 2 days, this electron density peak is 40% higher, consistent with the binding of an electron-rich bromide anion. Thus, a halide binding site is created in E117A CAII.

Electron density maps show clear density for a non-protein zinc ligand as well as other ordered solvent molecules in and around the active site. The peak corresponding to zinc-bound solvent, #348, lies 1.9 Å from zinc and exhibits good hydrogen bond geometry to the O_γ of T199 ($O^{348}-O_\gamma$ separation = 2.9 Å). Zinc-bound solvent in this model, just as in all other refined structures, is part of an extensive hydrogen bond network of ordered waters extending almost 10 Å from zinc.

No significant structural changes are observed elsewhere in the structure of E117A CAII. The rms deviation of C_α atoms from their wild-type positions is unremarkable at 0.21 Å, and the predominant rotamer of H64 is in the "in" conformation, unchanged from its wild-type position.[37,44,45] A superposition of the active site atoms of wild-type and E117A CAIIs is found in Figure 9.

Discussion

Modulating the pK_a of Zinc-Bound Water. In the absence of any drastic structural changes in a CAII variant, the pK_a of zinc-bound water provides a readily-measurable indication of the electrostatic effect of zinc ligands on the ability of zinc to polarize its bound water molecule. For example, in H94C and

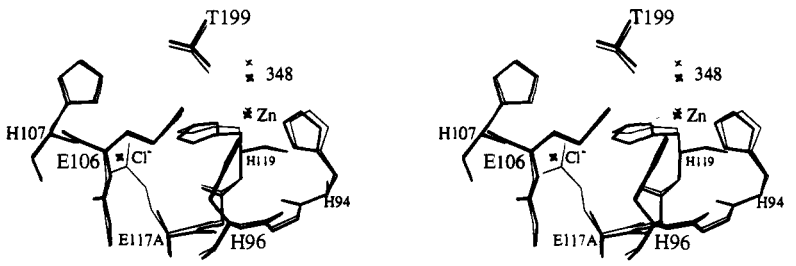

Figure 9. Least-squares superposition of E117A CAII C_α atoms (thick lines) with those of wild-type CAII (thin lines).

H94D CAIIs a negatively-charged group replaces a neutral histidine ligand. When the positive charge of zinc is offset by the introduction of negatively-charged ligands, zinc cannot polarize its bound water as effectively. As a result, the pK_a of zinc-bound water increases in H94C and H94D CAIIs to >9.5 and >9.6, respectively, from the wild-type pK_a of 6.8.[9,11]

Likewise, the pK_a of zinc-bound water affords a convenient means of quantifying the electrostatic contribution of hydrogen bonds to the metal ligands with respect to the ability of zinc to polarize its bound water molecule. Any hydrogen bond acceptor to H94 comparable to the neutral carboxamide side chain of Q92 should maintain the pK_a of zinc-bound water at the wild-type level. Similarly, any hydrogen bond acceptor to H119 comparable to the negatively-charged carboxylate side chain of E117 should likewise maintain the pK_a of zinc-bound water. However, if the hydrogen bond to a histidine zinc ligand were completely removed, the pK_a of zinc-bound water would decrease: the positive charge of the metal ion would not be offset by the (partial) negative charge of the deleted hydrogen bond acceptor.

In view of the seemingly drastic structural nature of the Q92A and E117A substitutions—that is, the removal of the entire side chain functionality which normally hydrogen bonds to the histidine metal ligand—it is surprising at first glance that the pK_a values of zinc-bound solvent in these variants are identical within experimental error: the pK_a values of zinc-bound water in Q92A, E117A, and wild-type CAIIs are 6.8, 6.9, and 6.8, respectively.[16] However, the high-resolution crystal structures of these variants reveal why a greater effect is not observed: each of these amino acid substitutions creates a void in the active site sufficiently large for a water molecule or a chloride ion to engage in a compensatory hydrogen bond with the histidine zinc ligand. The dipole–dipole Q92–H94 hydrogen bond is replaced by a dipole–dipole H_2O–H94 hydrogen bond in Q92A CAII. Likewise, the charge–dipole E117–H119 hydrogen bond is replaced by a charge–dipole Cl^-–H119 hydrogen bond in E117A CAII.[47] Each of these substitutions yields no net change in the electrostatic environment of zinc, so it is logical that no net change is observed in the pK_a of its bound water molecule (pK_a = 6.9).[16]

Interestingly, the hydrogen bond to H94 is *not* deleted by shortening the side chain of the acceptor at position 92: the Q92–H94 hydrogen bond is replaced by a somewhat longer, but well-oriented, N92–H94 hydrogen bond in Q92N CAII. Here, too, there is little net change in the electrostatic environment of zinc, so no significant change is observed in the pK_a of zinc-bound solvent.

The most striking effects on the pK_a of zinc-bound water are observed in Q92L and Q92E CAIIs. In Q92L CAII, the engineered aliphatic side chain is essentially isosteric with that of the wild-type enzyme but it cannot accept a hydrogen bond from H94. Therefore, the Q92L substitution blocks a water molecule from satisfying the hydrogen bond potential of H94, although N67 moves in order to form a poorly oriented hydrogen bond with H94. Thus, the effective positive charge on zinc is increased in Q92L CAII relative to the wild-type enzyme, which allows zinc to further polarize its bound solvent molecule. This accounts for the decreased pK_a of zinc-bound water by 0.4 pK_a unit to 6.4.[16] In contrast, the opposite effect is observed in Q92E CAII: the substitution of a neutral hydrogen bond acceptor by a negatively-charged acceptor diminishes the effective positive charge of the zinc ion and thereby raises the pK_a of zinc-bound water by 0.9 pK_a unit to 7.7.[16] We conclude that hydrogen bonds with histidine zinc ligands in metalloenzymes play an important role in fine-tuning the pK_a and reactivity of zinc-bound solvent.

Modulating the K_D of Zinc Binding. The structure of Q92N CAII is very similar to that of the wild-type enzyme, particularly with regard to the hydrogen bond network in the active site. For the Q92N variant then, no factor—entropic or enthalpic—should alter the zinc-binding behavior of the enzyme, which is in accord with the measurements of Kiefer and co-workers.[16] However, in Q92A and E117A CAIIs, protein–zinc affinity is decreased by nearly an order of magnitude. This may result from the increased entropic cost of fixing the histidine metal ligand and its hydrogen bonded solvent molecule for zinc binding. In fact, a Q92A, E117A double mutant isolated and characterized by Kiefer and colleagues[16] indicates that these entropic effects are additive.

Other variants in which the zinc ligands are significantly more weakly hydrogen bonded should likewise exhibit diminished protein–zinc affinity: the entropic cost of fixing the liganding histidine in proper orientation for metal binding is ordinarily "paid for" in the folding of the apoprotein (i.e., when the H94–Q92 and H119–E117 hydrogen bonds are initially formed). Therefore, the compromise of these hydrogen bonds will affect protein–metal affinity by incurring an entropic cost to fix not only the zinc ion but also its ligands during zinc binding.

In the engineering of hydrogen bonds to metal ligands to increase metal binding affinity, it is clear that these interactions should involve only protein residues and not solvent molecules. Fixing the zinc ion in a catalytically-optimal position necessarily requires well-oriented direct ligands. Indirect ligands thus serve to fix the direct ligands, and ordinarily dynamic solvent molecules are insufficient for such a structural task (unless such solvent molecules are firmly hydrogen bonded by other protein residues). Q92L CAII presents an intermediate scenario: although the H94–X92 hydrogen bond is lost, a compensatory (but weak) hydrogen bond is formed between H94 and N67. So, prior to zinc binding, H94 may be fixed in sub-optimal orientation for metal binding, and a possible requirement for its conformation change manifests itself in a 7-fold decrease in protein–zinc affinity.

Finally, it is interesting to note the differential effects of alanine substitutions at Q92 and E117. Alanine substitutions

(47) Although E117A CAII is overexpressed in *E. coli* in the presence of chloride, it is neither purified nor assayed in chloride-containing buffers; hence, it is not clear what species binds in the chloride pocket of E117A CAII under assay conditions.

have a 2-fold greater effect on weakening the protein−zinc K_D at position 117 than at position 92.[16] Likewise, the removal of a methylene group from the side chain in the Q92N and E117D variants has a 2-fold greater effect on zinc affinity at position 117 relative to position 92. The implication is that the indirect ligand at position 117 anchors the zinc ion more tightly than the indirect ligand at position 92.[16] In other words, carboxylate−histidine−zinc triads are stronger than carboxamide−histidine−zinc triads.[2,3]

Conclusions

We conclude this dissection of the CAII zinc binding site with the following general points relevant to the structure-assisted design of protein−metal binding sites, which must include consideration of the architecture of *both* direct and indirect metal ligands:

(1) Amino acid residues that hydrogen bond with histidine zinc ligands play a subtle yet important role in fine-tuning the pK_a and reactivity of zinc-bound solvent.[2,3] Alteration of these hydrogen bonding groups by site-directed mutagenesis yields pK_a adjustments on the order of 0.1−0.2 pK_a unit for structurally conservative mutations and 0.5−1.0 pK_a unit for mutations which alter the net charge of the hydrogen bonding residues.[16] Histidine zinc ligands in the CAII active site have an extraordinarily high hydrogen bond potential, and the three-dimensional structures reported herein reveal modes by which compensatory hydrogen bonding groups are recruited from solvent (as in Q92A and E117A CAIIs) or from the protein scaffolding itself (as in Q92N and Q92L CAIIs).

(2) Amino acid residues that hydrogen bond with histidine zinc ligands play a modest yet important role in modulating protein−zinc affinity.[2,3] Due to their structural and electrostatic contributions, these hydrogen bond networks may contribute up to a factor of 10^4 to affinity in a tetracoordinate metal site.[16] For example, the substitution of the Q92−H94 hydrogen bond by H_2O−H94 diminishes protein−zinc affinity 5-fold, probably due to an increased entropic cost to fix H94 in the proper conformation for zinc complexation. We have been only marginally successful in engineering an enzyme with increased zinc affinity by strengthening the hydrogen bond to H94: Q92E CAII does not differ significantly in its zinc affinity, whereas the expected result was a modest increase due to the replacement of the Q92−H94 dipole−dipole interaction with a stronger E92−H94 charge−dipole interaction.[16]

(3) Just as hydrogen bond networks with metal ligands are important for protein−zinc affinity in CAII, we advance that such networks are likewise important for the binding of certain proteins, such as the C_2H_2 class of zinc-finger transcription factors, to nucleic acids.[2] For instance, the three-dimensional structures of Zif268, GLI, and Tramtrack complexed with DNA reveal phosphate−histidine−zinc triads between the DNA fragment and the protein.[48−50] These triads are analogous in structure and electrostatics to the E117−H119-Zn^{2+} triad of CAII, although inverted in function: zinc-finger proteins bind DNA phosphate groups—the indirect zinc ligand—because zinc preorients the histidine ligand; in CAII, indirect ligand E117 preorients the histidine ligand for zinc complexation. By analogy with the measurements of Kiefer and colleagues,[16] it is possible that the phosphate−histidine hydrogen bond in the protein−DNA complex may contribute 5−10-fold to affinity. It is likely that hydrogen bonds between phosphate and histidine metal ligands in C_2H_2 (and possibly C_3H) zinc-finger protein−DNA complexes play an important role in protein−nucleic acid recognition, as previously advanced by Christianson and Alexander.[2]

Acknowledgment. We thank the Office of Naval Research for support of this work; C.A.L. is supported in part by NIH Cell and Molecular Biology Training Grant No. GM07229. We also are grateful to Drs. Carol Fierke and Laura Kiefer for providing protein and to Drs. Joseph Ippolito and Satish Nair for many helpful discussions.

JA9502381

(48) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809−817.
(49) Pavletich, N. P.; Pabo, C. O. *Science* **1993**, *261*, 1701−1707.
(50) Fairall, L.; Schwabe, J. W. R.; Chapman, L.; Finch, J. T.; Rhodes, D. *Nature* **1993**, *366*, 483−487.