Spinach Carbonic Anhydrase: Investigation of the Zinc-Binding Ligands by Site-Directed Mutagenesis, Elemental Analysis, and EXAFS†

Michael H. Bracey,‡ Jason Christiansen,§ Pilar Tovar,‡ Stephen P. Cramer,§ and Sue G. Bartlett*,‡

Department of Biochemistry, Louisiana State University, Baton Rouge, Louisiana 70803, and Department of Applied Science, University of California, Davis, California 95616

Received May 3, 1994; Revised Manuscript Received August 19, 1994®

ABSTRACT: The enzyme carbonic anhydrase has been well characterized in mammalian systems, but the structural properties of the plant isozymes remain elusive. To investigate the nature of the zinc-binding site in spinach carbonic anhydrase, we targeted potential zinc ligands for mutagenesis and examined the resulting enzymes for catalytic activity and stoichiometric zinc binding. In addition, we examined the wild-type protein using extended X-ray absorption fine structure analysis. Our results suggest that spinach carbonic anhydrase utilizes a Cys–His–Cys–H_2O ligand scheme to bind the zinc ion at the active site.

Carbonic anhydrase (CA;[1] carbonate dehydratase, EC 4.2.1.1) is a ubiquitous zinc metalloenzyme which catalyzes the reversible hydration of carbon dioxide. In mammals, CA plays important roles in facilitating carbon dioxide exchange in capillary beds and alveoli, maintaining the buffering capacity of blood, and reabsorbing bicarbonate across renal tubules (Tashian, 1989). Crystal structures for the human isozymes I and II and the bovine isozymes II and III have been solved, and they show that these enzymes coordinate an active-site zinc through three conserved histidine residues [for example, see Håkansson *et al.* (1992)].

In higher plants which carry out C_3 photosynthesis, the majority of CA activity can be localized to the chloroplast stroma where the enzyme's role is unclear, although it may serve to concentrate carbon dioxide at the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase (Graham *et al.*, 1984). CA from higher plants is quite different from the major mammalian isozymes in both primary sequence and multimeric assembly. CA from C_3 dicotyledonous plants is a hexamer with one zinc atom per monomer (Graham *et al.*, 1984; Kisel & Graf, 1972) while the major mammalian isozymes are monomeric (Deutsch, 1987). Furthermore, sequence analysis reveals no homology between the plant and animal CA's, and this suggests that the animal and plant isozymes do not share a common evolutionary origin (Fukuzawa *et al.*, 1992) and therefore may not share common physical properties.

Until a crystal structure becomes available for a plant-type CA, biochemical investigations may provide hints of structural characteristics of this enzyme. To this end, we have employed techniques of molecular biology and biophysics to investigate the nature of the zinc-binding site in spinach carbonic anhydrase. We present here results of experiments utilizing site-directed mutagenesis, elemental analysis, and EXAFS spectroscopy.

EXPERIMENTAL PROCEDURES

Cloning and Mutagenesis. All DNA manipulations including plasmid isolation, restriction digestions, ligations, and transformations were performed using standard methods (Sambrook *et al.*, 1989). Site-specific mutations were introduced into CA using the Altered Sites Mutagenesis System (Promega). Mutagenesis changing histidine to glutamine, cysteines to alanines, glutamates to glutamines, and aspartate to asparagine resulted in the mutants H210Q, C150A, C213A, E194Q, E266Q, and D152N. Mutations were confirmed by sequencing according to a modified dideoxy chain termination method (Fawcett & Bartlett, 1990).

Mutant constructs were cloned into a derivative of the expression vector pKK233-2 (Pharmacia LKB Biotechnology, Inc.) and expressed in *Escherichia coli* DH5α for activity measurements. To facilitate purification, these CA's were further subcloned into a derivative of the expression vector RIT2T (Pharmacia LKB Biotechnology, Inc.) to yield constructs which upon translation would result in carbonic anhydrase fused to the carboxyl terminus of *Staphylococcus* protein A.

Protein Isolation and Assays. Protein was isolated from the pKK expression system for activity measurements as follows. *E. coli* DH5α harboring the appropriate CA construct or the vector alone was grown at 37 °C in LB media to an OD_{600} of 1.8. Cells were pelleted and resuspended in buffer A (50 mM Tris-HCl, pH 8, 100 mM NaCl, and 1 mM EDTA). After lysing by sonication, the resulting homogenate was clarified and brought to 56% saturation with solid ammonium sulfate. Protein was allowed to precipitate for at least 1 h on ice and then pelleted at 32000*g* for 15 min. The resulting pellets were resuspended and dialyzed overnight at 4 °C against buffer A.

Each dialyzed sample was quantitated with the Coomassie protein assay (Pierce) using BSA as a standard. The presence of recombinant CA in each of the samples was verified by Western blotting.

Activity measurements of the crude CA preparations were determined according to the spectrophotometric method of Khalifah (1971). Briefly, 0.5 mL of an appropriate buffer–indicator pair was placed in a cuvette, followed by the addition

† The National Synchrotron Light Source is supported by the U.S. Department of Energy, Division of Materials Sciences and Division of Chemical Sciences. S.G.B. acknowledges support from LEQSF.

* To whom correspondence should be addressed.
‡ Louisiana State University.
§ University of California.

® Abstract published in *Advance ACS Abstracts,* October 1, 1994.

[1] Abbreviations: CA, carbonic anhydrase; ICP-AES, inductively coupled plasma atomic emission spectroscopy; EXAFS, extended X-ray absorption fine structure; BSA, bovine serum albumin; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; kDa, kilodalton.

0006-2960/94/0433-13126$04.50/0

```
           149                168                 188                 208
             ● o                                           o
Spinach     ACSDSRVCPSHVLDFQPGEAFMVRNIANMVPVFDKDKYAGVGAAIEYAVLHLKVENIVVI
Pea         ACSDSRVCPSHVLDFQPGEAFVVRNVANLVPPYDQAKYAGTGAAIEYAVLHLKVSNIVVI
Tobacco     ACSDSRVCPSHVLNFQPGEAFVVRNIANMVPAYDKTRYSGVGAAIEYAVLHLKVENIVVI
A. thaliana ACSDSRVCPSHVLDFQPGDAFVVRNIANMVPPFDKVKYGGVGAAIEYAVLHLKVENIVVI
E. coli     SCSDSRLVPELVTQREPGDLFVIRNAGNIVPSYGPEP-GGVSASVEYAVAALRVSDIVIC
Synecho.    TCSDSRIDPNLITQSGMGELFVIRNAGNLIPPFGAAN-GGEGASIEYAIAALNIEHVVVC

           209                228                 248                 268
             ●  ●                                                    o
Spinach     GHSACGGIKGLMSFPDAGPTTTDFIEDWVKICLPAKHKVLAEHGNATFAEQCTHCEKEAV
Pea         GHSACGGIKGLLSFPFDGTYSTDFIEEWVKIGLPAKAKVKAQHGDAPFAELCTHCEKEAV
Tobacco     GHSACGGIKGLMSLPADGSESTAFIEDWVKIGLPAKAKVQGEHVDKCFADQCTACEKEAV
A. thaliana GHSACGGIKGLMSFPLDGNNSTDFIEDWVKICLPAKSKVISELGDSAFEDQCGRCEREAV
E. coli     GHSNCGAMTAIAS-C-QCMDHMPAVSHWLRYADSA-RVVNEARPHSDLPSKAAAMVRENV
Synecho.    GHSHCGAMKGLLKLN-QLQEDMPLVYDWLQHAQATRRLVLDNYSGYETDDLVEILVAENV
```

FIGURE 1: Homology among the plant and bacterial carbonic anhydrases. Alignment is shown of the portion of the plant-type CA's spanning the six conserved potential ligands from spinach, pea, tobacco, *Arabidopsis, E. coli*, and *Synechococcus*. Amino acid numbering corresponds to the spinach sequence as deposited in GenBank. Circles indicate potential ligands to zinc. Filled circles indicate potential zinc ligands that, when mutated, resulted in CA that bound zinc poorly. Alignment generated by PILEUP of the UWGCG programs.

of CA extract. The reaction was initiated by the addition of CO_2-saturated water, and the change in absorbance of the indicator was recorded as a function of time. We assayed in 25 mM EPPS, pH 8, with 0.113 mM phenol red monitored at 555 nm and in 25 mM imidazole, pH 8, with 0.072 mM *p*-nitrophenol monitored at 400 nm. These buffer–indicator pairs were chosen on the basis of their similar pK_a values.

For each buffer system, an acid calibration curve was used to calculate the buffer factor (Q) which was then used to convert dA/dt to millimoles of CO_2 per second per milligram. Initial rates of CO_2 hydration were determined by extrapolating back to time zero from a linear sampling time. The hydration rate measured using protein derived from cells harboring the empty pKK233–2 vector was considered spontaneous and subtracted as uncatalyzed background.

Ellman's assays were performed essentially as described (Riddles *et al.*, 1983).

Inductively Coupled Plasma Atomic Emission Spectroscopy. For the metal analysis, fusion proteins were expressed in *E. coli* and purified by affinity chromatography using nonspecific IgG linked to Reacti-Gel 6x (Pierce). Purified protein A/CA fusions were freed of loosely bound zinc using Chelex 100 (Bio-Rad). As negative controls, buffer alone or an amount of BSA comparable to the amount of CA fusion assayed was treated similarly. Each sample was analyzed for zinc concentration using an ARL 34000 inductively coupled plasma atomic emission spectrometer and electronics and software from Labco, Inc. For the final determination of zinc binding, readings for the zinc-free BSA standard were treated as negligible background due to density and viscosity differences within the plasma torch as compared to the buffer control and were subtracted from readings for the CA fusions.

EXAFS. Wild-type protein A/CA fusion was purified as above from 50 g of cell paste. Protein was concentrated using Minicon-B15 clinical sample concentrators (Amicon) to a final concentration of ~2 mM and supplemented with 30% ethylene glycol; 80 μl samples were loaded into lucite cells with a 1 mm path length and sealed with 0.001 in. Kapton tape. X-ray absorption spectra were recorded at the National Synchrotron Light Source, Brookhaven National Laboratory, on beamline X10-C. The beamline was run in the focused mode with an Si(111) double-crystal monochromator configuration. Higher order harmonics were rejected using a mirror position feedback system (Sansone *et al.*, 1991). Frozen samples were loaded into an Oxford Instruments liquid helium flow cryostat maintained at ~10 K. XAFS data were collected in the fluorescence mode using a Canberra Industries 13-element Ge solid-state array detector (Cramer *et al.*, 1988) while incident beam intensity was monitored with a nitrogen-filled ion chamber. Photon energy was calibrated by simultaneously collecting a transmission spectrum of a zinc metal foil and setting the first inflection point energy at 9659.0 eV.

EXAFS oscillations were extracted from the raw data by routine methods (Cramer *et al.*, 1978) and were then quantitatively analyzed using a Levenberg–Marquardt nonlinear least-squares calculated curve-fitting procedure to minimize differences between the data and observed EXAFS. Simulations were derived from the curved-wave functional form:

$$\chi(k) = \sum_i \frac{N_i \gamma_i f_i(k,R_i)}{kR_i^2} e^{-2\sigma_i^2 k^2} \sin[2kR_i + \phi_i(k,R_i)]$$

(McKale *et al.*, 1986). For this analysis, theoretical values for both phase and amplitude were used, and the value γ was fixed at 0.9 for all fits. During fitting, the total Zn coordination was set to four or five, and small changes in the threshold energy (ΔE_0) were fixed at −4.2 eV (Hubbard *et al.*, 1981), while the interatomic distance (R) and the mean square deviation of R (σ^2) were allowed to vary.

RESULTS

Sequences for the *E. coli* CA homologue (Sung & Fuchs, 1988), the *Synechococcus* CA (Fukuzawa *et al.*, 1992), and CA's from *Arabidopsis* (Raines *et al.*, 1992), pea (Roeske & Ogren, 1990), spinach (Fawcett *et al.*, 1989), tobacco (Majeau & Coleman, 1992), and barley[2] were analyzed for conservation of the amino acids whose side chains are known to serve as zinc ligands at enzyme active sites (Vallee & Auld, 1990). Homology alignments showed that one histidine, two cysteines, two glutamates, and one aspartate are conserved among the plant-type CA's (Figure 1). Mutagenesis changing these six residues resulted in mutant forms which displayed altered catalytic activity (Table 1). Western blotting of lysates from *E. coli* expressing each of the mutant proteins revealed that all were expressed in the soluble form.[3]

The six spinach CA mutants and the wild-type enzyme were also analyzed for their ability to bind zinc. In order to facilitate purification of the proteins, the mutant forms as

[2] M.H.B. and S.G.B., in preparation.

[3] Data not shown.

Table 1: Activity of and Zinc Binding to Mutants of Spinach Carbonic Anhydrase

mutant	activity[a] EPPS	activity[a] imidazole	mol of Zn/mol of CA monomer
WT	100	100	1.09
C150A	ND[b]	9	0.03
D152N	1	9	0.81
E194Q	22	5	0.80
H210Q	10	6	ND
C213A	6	5	0.15
E266Q	81	100	1.25

[a] Activities are expressed as percentages of the wild-type enzyme in each buffer system. The blank-corrected activity of the wild-type was 6.8 mmol of CO_2 s^{-1} (mg of *E. coli* protein)$^{-1}$ in EPPS buffer and 76 mmol of CO_2 s^{-1} (mg of *E. coli* protein)$^{-1}$ in imidazole. Reported activities represent an average of four to five independent assays from two preparations. The assay was limited by an uncatalyzed rate of hydration of 7.9 mmol of CO_2 s^{-1} (mg of *E. coli* protein)$^{-1}$ in EPPS buffer and 32 mmol of CO_2 s^{-1} (mg of *E. coli* protein)$^{-1}$ in imidazole. [b] Not detectable.

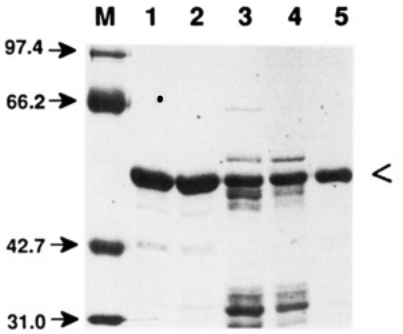

FIGURE 2: Proteolytic susceptibility of selected mutant CA's. SDS–PAGE of the purified CA fusions assayed for zinc content showing the instability of two mutant proteins that bind zinc poorly. Lane 1, WT; lane 2, H199Q; lane 3, H210Q; lane 4, C150A; lane 5, C213A. Molecular weight markers are indicated. The arrowhead at the right indicates the protein A/CA fusion bands.

well as wild-type CA's were expressed in *E. coli* as C-terminal fusions to protein A. The wild-type CA fusion protein was catalytically active, indicating that the protein A moiety does not significantly interfere with proper folding or function of the CA portion of the fusion. After purification, the fusion proteins were depleted of adventitious zinc and analyzed by inductively coupled plasma atomic emission spectroscopy. Wild-type CA bound zinc in a ratio of one atom per subunit. However, the mutants C150A, H210Q, and C213A had greatly diminished capacities to bind zinc (Table 1).

Two of the latter mutant CA's were more susceptible to proteolysis than the wild-type protein when expressed in *E. coli*, and during affinity chromatography various breakdown products copurify. These breakdown products are evident in Figure 2 and are shown with the wild-type enzyme and two stable mutants for comparison. These products likely represent various C-terminal truncations that are still able to interact with the IgG column through the protein A domain at the amino terminus. Regardless, even if intact fusion represented only 35% of the total protein assayed in the case of C150A and only 10% in the case of H210Q, ratios of zinc to CA monomer would still be less than 0.3 based on back-calculations from the ICP-AES data.

The instability of the cysteine mutant C150A is not due to the disruption of a disulfide bond involving this residue. Ellman's assays of the wild-type fusion protein revealed the presence of two reduced cysteines per monomer in the native state and all six cysteines per monomer in the presence of 6 M guanidinium chloride.[4] This indicates that no disulfides exist in the bacterially expressed spinach CA, though four of the six cysteines present per monomer are inaccessible to solvent in the native state.

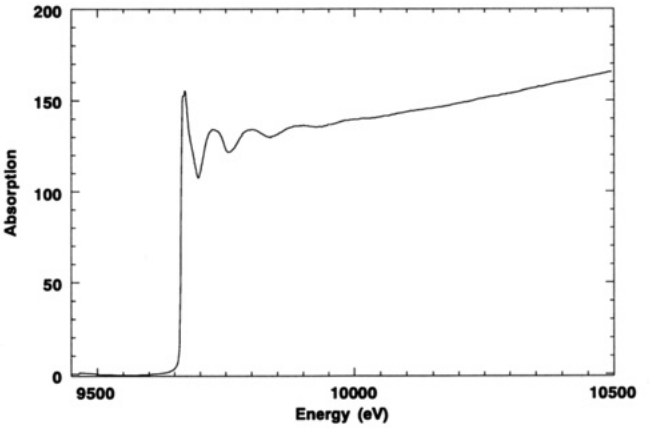

FIGURE 3: Zn *K*-edge XAS spectrum. Pre-edge-subtracted absorption spectrum of protein A/CA fusion.

To further investigate the nature of the ligands in the wild-type CA fusion protein, we examined its Zn X-ray absorption spectrum. The fluorescence-detected Zn *K*-edge spectrum is shown in Figure 3. The CA Zn EXAFS is relatively strong and without a clear beat pattern (Figure 4A). The Fourier transform is dominated by a single peak centered at ~2.3 Å with a few minor peaks above the noise between 3 and 4 Å (Figure 4B). The pattern is similar to the transforms observed in EXAFS studies of plastocyanin and other blue Cu proteins (Scott *et al.*, 1982).

The dominant feature could be simulated by a Zn–S interaction at ~2.3 Å, but additional Zn–N/O interactions near 2 Å were necessary to get a good fit. The small features present beyond the central 2.3 Å peak are most probably due to multiple scattering interactions from an imidazole group and some contributions from the carbons of the cysteines. Unfortunately, with little knowledge of the symmetry and geometry of the site, attempts to fit these features with multiple scattering have not been successful. The quality of fit for all possible combinations of sulfur and nitrogen/oxygen interactions was judged by comparing the fit index between the resulting fit and the raw extracted EXAFS, and these results are summarized in Table 2.

DISCUSSION

The plant-type carbonic anhydrases have long been overshadowed by the animal isozymes, but the former are emerging as rather intriguing enzymes in their own right. Though it is unlikely that the animal and plant-type CA's are evolutionarily related (Fukuzawa *et al.*, 1992), these two groups share many common features. Both are zinc metalloenzymes which catalyze the hydration of carbon dioxide with remarkable efficiency and are susceptible to a number of common inhibitors. Among these, the sulfonamides inhibit the animal isoforms by binding to the catalytic zinc (Vidgren *et al.*, 1990). By analogy, since sulfonamides also inhibit plant-type CA's (Pocker & Ng, 1974), it is presumed that the zinc bound to the plant enzyme is also catalytic in nature. However, in contrast to the animal enzyme, plant CA's do not possess three conserved histidines to account for zinc coordination, consistent with the animal model. Therefore, the two isoforms may have evolved to coordinate the active-site metal by two very different strategies.

[4] Data not shown.

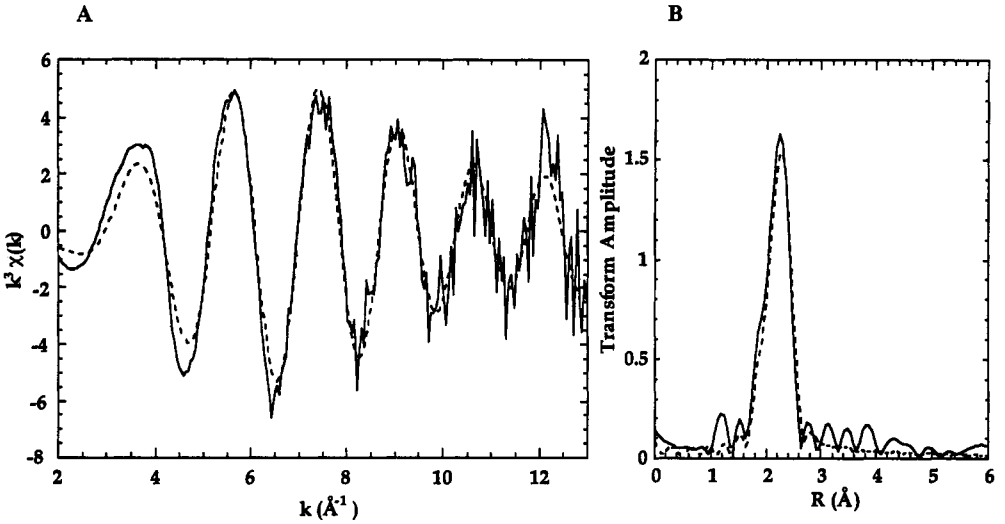

FIGURE 4: Results of EXAFS curve-fitting analysis. (A) Experimental k^3 EXAFS (solid line) and final-fit results for 2 Zn-S—2 Zn-N/O coordination to Zn (dashed line). (B) Fourier transform of EXAFS (solid line) with final-fit results (dashed line).

Table 2: Results of EXAFS Fitting Analysis

model	N	R (Å)	σ^2 (Å^2 × 10^{-3})	fit index (F)[a]
tetracoordinate				
Zn–S	4	2.31	6.29	314
Zn–S	3	2.32	4.08	
Zn–N/O	1	2.04	1.35	212
Zn–S	2	2.32	2.03	
Zn–N/O	2	2.06	3.02	171
Zn–S	1	2.32	1.00	
Zn–N/O	3	2.08	5.09	255
Zn–N/O	4	2.11	4.90	777
pentacoordinate				
Zn–S	2	2.32	2.37	
Zn–N/O	3	2.06	5.63	143

[a] $F = [\Sigma(\chi_e - \chi_s)^2 k^6]/n$ where the difference is between each data point of the experimental (χ_e) and simulated (χ_s) EXAFS, and n is the number of data points in the fitting page. All fits were performed on the range k = 2–13 Å^{-1}. During fitting, the total Zn coordination was set to four or five, and small changes in the threshold energy (ΔE_0) were fixed at –4.2 eV (Hubbard *et al.*, 1981), while the interatomic distance (R) and the mean square of R (σ^2) were allowed to vary.

To investigate how spinach CA might coordinate zinc, we analyzed available sequences of plant-type CA's for the conservation of all amino acids whose side chains are known to serve as ligands to active-site zincs, namely, histidine, aspartate, glutamate, and cysteine (Vallee & Auld, 1990). Six such conservations were identified, and each residue was targeted for mutagenesis (Figure 1). Among these, the aspartate, histidine, and cysteine mutants each exhibited less than 10% of the activity of the wild-type enzyme when assayed in EPPS buffer (Table 1).

Mutations in the carbonic anhydrase from pea at locations equivalent to those described here result in similar enzyme performance (Provart *et al.*, 1993). Pea mutants corresponding to the spinach H210Q, C150A, and C213A (H220N, C160S, and C223S, respectively) had no measurable activity, reinforcing our results with the spinach enzyme. Additionally, Coleman's group found the pea E204A to exhibit no activity and E276A to be rather compromised. However, we find the equivalent, more conservative spinach mutant E266Q to display near-wild-type activity, while E194Q is 22% as active as the wild-type enzyme. We attribute these differences of results to Coleman's more severe charged to aliphatic mutations which may serve to disrupt the pea enzyme more than the conservative spinach mutations. Since Coleman's group did not construct a mutation comparable to D152N, we are unable to compare our results regarding this variant. Nonetheless, a duplication of enzymatic trends among various mutants from two species provides further evidence that H210, C150, and C213 are critical residues in spinach CA.

In previous studies on other zinc proteins, site-directed mutagenesis of potential zinc ligands followed by the loss of function of the metalloprotein in question has been provided as evidence implicating residues in metal coordination. This has been demonstrated for leukotriene A_4 hydrolase (Medina *et al.*, 1991), neutral endopeptidase (LeMoual *et al.*, 1993), and a potential zinc finger in the glucocorticoid receptor (Severne *et al.*, 1988). We have shown here the same correlation in spinach CA; of six conserved potential zinc ligands, four can be mutated to cause a 90% loss of function (Table 1). As the zinc of CA is presumably catalytic in nature, the loss of the enzyme's ability to coordinate this metal would necessarily result in a concomitant loss of activity. Thus, based on our activity assays, the conserved histidine, cysteines, and aspartate are all candidates for zinc ligands in spinach carbonic anhydrase.

Additionally, we have definitively ascertained which of these mutations affect the enzyme's ability to bind zinc rather than affect catalysis through some secondary disruption. Such a disruption could conceivably occur by various means, for example, by abolishing a hydrogen bond network within the active site as suggested for the pea mutant E276A (Provart *et al.*, 1993). Through elemental analysis, we have established that the histidine and cysteine mutants uniquely exhibit severely diminished capacities to bind zinc (Table 1). This correlation between a loss of activity and loss of zinc binding has also been shown for the zinc ligand mutant H94D in human CA II (Kiefer *et al.*, 1993) and mutants of leukotriene A_4 hydrolase (Medina *et al.*, 1991). We interpret this as additional proof indicating that these three residues are involved in the coordination of the zinc ion, while the aspartate may play an essential role unrelated to metal coordination.

While the ICP-AES data do not implicate the glutamate and aspartate residues as potential zinc ligands, these variants do demonstrate significant catalytic disruptions. To investigate if these residues might be involved in a proton transfer step within the active site as suggested for various amino acids in the animal isozymes, we also assayed in an imidazole buffer system. As a relatively small buffering molecule, imidazole

is thought to complement the catalytic deficiency of proton shuttle mutants by entering the active site and essentially replacing the missing functional group. This property has been exploited in the animal systems to identify proton shuttling amino acids such as R67 in human CA III (Tu *et al.*, 1990). We find that imidazole does indeed restore the activity of E266Q to wild-type levels, implicating this residue as a possible proton shuttle in the active site. This conclusion has also been reached for the equivalent pea CA mutation (Provart *et al.*, 1993). The remaining mutants do not exhibit a stimulatory imidazole effect. It is interesting that the activity of E194Q actually decreases relative to the wild-type enzyme in the imidazole system (Table 1). The reasons for this decline are not immediately obvious.

An increased susceptibility to endogenous *E. coli* proteases during the purification of C150A and H210Q is demonstrated in Figure 2, and this may represent an instability of the mutant proteins. Similar instability has been documented in other systems where zinc metalloenzymes have been depleted of zinc and assayed for resistance to trypsin proteolysis. In the case of the DNA-binding domain of the GAL4 transcription factor, the zinc-free apoenzyme was reduced to peptides upon trypsinization while the zinc-bound form was cleaved to a 13 kDa core particle (Pan & Coleman, 1989). Similar results were also obtained with the gene 32 protein from bacteriophage T4 (Giedroc *et al.*, 1987). Furthermore, differential scanning calorimetry investigations of the folding processes of the gene 32 protein as well as alkaline phosphatase from *E. coli* showed metal-dependent changes in the denaturation profiles of these zinc metalloproteins (Keating *et al.*, 1988; Chlebowski & Mabrey, 1977). These results imply that the binding of the metal imparts significant stabilization to these proteins at the level of tertiary structure, or, as suggested for aspartate carbamoyltransferase, at the level of quaternary structure (Monaco *et al.*, 1978).

These characteristics may also hold true for spinach carbonic anhydrase. Possibly, the formation of the active site is necessary for the proper folding of CA monomers; the inability to bind the metal could disturb elements of the protein's secondary or tertiary structure, exposing protease-sensitive domains. Alternatively, the active sites of spinach CA may be located at subunit interfaces, and zinc could stabilize the quaternary structure of the holoenzyme. Though the zinc of CA is presumably catalytic in nature, this established correlation of protein stability and metal binding is a possible explanation for the susceptibility to proteolysis in our CA mutant proteins.

To further investigate the zinc-binding site in wild-type spinach carbonic anhydrase, we analyzed this enzyme's extended X-ray absorption fine structure. EXAFS has the advantage of providing a direct examination of a metal's environment in metalloproteins. Analysis of the spectrum of spinach CA yields evidence that the nearest neighbors of the zinc are sulfur and nitrogen/oxygen species. Of the five possible combinations of two distinct Zn–S and Zn–N/O interactions in a four-ligand system, 2 Zn–S—2 Zn–N/O gives the best fit, in agreement with the results of the mutagenesis. This model represents a 20% improvement in fit index over the next lowest value (Table 2). It is clear that a model with four N/O ligands coordinated to Zn, as found in mammalian CA, is not consistent with the EXAFS since the corresponding fit is of much worse quality than the others performed on this system. Attempts to fit a longer N/O shell at ~2.3 Å were also unsuccessful, yielding fit indexes $\sim 10^4$ or causing the second Zn–N/O distance to contract back to ~2.1 Å, giving the same result as the 4 N/O model. Furthermore, the presence of only two conserved cysteines per polypeptide in the plant-type CA's is consistent with our 2 Zn–S—2 Zn–N/O EXAFS model and renders the 3 Zn–S—1 Zn–N/O and 4 Zn–S—0 Zn–N/O models improbable.

Since four separate CA mutants yielded enzymes with activities of 10% or less, we also examined the EXAFS data for the possibility of a pentacoordinate zinc as implicated for adenosine deaminase (Wilson & Quiocho, 1993; Bhaumik *et al.*, 1993). The fit index resulting from the addition of a third N/O ligand at ~2.0 Å for a 2 Zn–S—3 Zn–N/O system does yield a 16% improvement over the 2 Zn–S—2 Zn–N/O model. However, several lines of evidence discredit this five-ligand model. First, the fit indexes for the best tetracoordinate model and the pentacoordinate model lie within the usual error range placed on values of coordination derived using the EXAFS technique. Therefore, it is objectively impossible to favor one model or the other based purely on EXAFS results. Second, when a tetracoordinate zinc is converted to a pentacoordinate one in a synthetic system, the bond length of the leaving group is longer than the four remaining ligand bond lengths (Auf der Heyde & Nassimbeni, 1984). We can find no such asymmetry in the CA pentacoordinate model; all N/O bond lengths remain the same. Third, only three of our potential ligand mutants exhibit a diminished capacity for zinc binding. If we define a *ligand* as a residue that contributes to the binding of a metal, then the EXAFS data may indicate that D152 is a near neighbor of the metal, while the ICP-AES data indicate that this residue is not a ligand *per se*. Overall, we find the tetracoordinate system to be a more credible model.

Monozinc enzymes may be classified into families based on the spacing of zinc ligands and the conservation of amino acids adjacent to them (Vallee & Auld, 1990). A typical catalytic zinc-binding site in this nomenclature is made up of two closely spaced ligands, L_1 and L_2, which comprise a zinc-binding nucleus separated by an amino acid spacer, X, plus a third, distant L_3 separated by a spacer, Y. In the case of spinach CA, H210 and C213 may represent L_1 and L_2, respectively, while C150 may represent L_3, with $X = 2$ and $Y = 59$. A preliminary search of the SwissProt database for homology with the amino acid motif encompassing the L_1–L_2 nucleus revealed no significant homology with any known zinc metalloenzymes. Thus, the plant-type carbonic anhydrases may represent a novel family of zinc metalloproteins.

In summary, site-directed mutagenesis of potential zinc ligands in spinach carbonic anhydrase revealed that mutant proteins in which cysteine 150 was converted to alanine, aspartate 152 was converted to asparagine, histidine 210 was converted to glutamine, or cysteine 213 was converted to alanine exhibited severely diminished catalytic activity. Furthermore, the cysteine and histidine mutants had greatly diminished capacities to bind zinc. In addition, EXAFS data are consistent with a proposed metal coordination by two sulfurs, while the bond lengths suggest tetracoordination, presumably by additional nitrogen or oxygen donor ligands. The residues corresponding to cysteine 150, histidine 210, and cysteine 213 are conserved in all plant-type CA's sequenced to date, and so our results suggest that the side chains of these amino acids are the protein ligands to the active-site Zn, while a reactive water molecule may be inferred to complete the coordination sphere. With this ligand scheme, spinach CA joins alcohol dehydrogenase and cytidine deaminase as the third zinc metalloenzyme in which a catalytic zinc is coordinated by sulfur ligands (Vallee & Auld, 1990; Betts *et al.*, 1994). Though a definitive resolution of the role of the conserved aspartate will await the elucidation of a CA crystal

structure or a detailed enzymological study of D152 mutants, the plant-type CA's seem to differ from the mammalian isozymes not only in primary sequence and quaternary structure but also in the nature of the side chains responsible for binding the catalytic zinc.

ACKNOWLEDGMENT

We thank Robyn Migues and Joni Haigler for performing the atomic emission spectroscopy, Brian Zganjar for assistance with the mutagenesis, Melissa M. Grush for assistance in EXAFS data collection, and Dr. Graham George for EXAFS curve-fitting software.

REFERENCES

Auf der Heyde, T. P. E., & Nassimbeni, L. R. (1984) *Acta Crystallogr., Sect. B B40,* 582.
Betts, L., Xiang, S., Short, S. A., Wolfenden, R., & Carter, C. W., Jr. (1994) *J. Mol. Biol. 235,* 635.
Bhaumik, D., Medin, J., Gathy, K., & Coleman, M. S. (1993) *J. Biol. Chem. 268,* 5464.
Chlebowski, J. F., & Mabrey, S. (1977) *J. Biol. Chem. 252,* 7042.
Cramer, S. P., Hodgson, K. O., Steifel, E. I., & Newton, W. O. (1978) *J. Am. Chem. Soc. 100,* 2748.
Deutsch, H. F. (1987) *Int. J. Biochem. 19,* 101.
Fawcett, T. W., & Bartlett, S. G. (1990) *BioTechniques 9,* 46.
Fawcett, T. W., Browse, J. A., Volokita, M., & Bartlett, S. G. (1989) *J. Biol. Chem. 265,* 5414.
Fukuzawa, H., Suzuki, E., Komukai, Y., & Miyachi, S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89,* 4437.
Giedroc, D. P., Keating, K. M., Williams, K. R., & Coleman, J. E. (1987) *Biochemistry 26,* 5251.
Graham, D., Reed, M. L., Patterson, B. D., Hockley, D. G., & Dwyer, M. R. (1984) *Ann. N.Y. Acad. Sci. 429,* 222.
Håkansson, K., Carlsson, M., & Svensson, L. A. (1992) *J. Mol. Biol. 227,* 1192.
Hubbard, S. R., Bishop, W. R., Kirschmeier, P., George, S. J., Cramer, S. P., & Hendrickson, W. A. (1981) *Science 254,* 1776.
Keating, K. M., Ghosaini, L. R., Giedroc, D. P., Williams, K. R., Coleman, J. E., & Sturtevant, J. M. (1988) *Biochemistry 27,* 5240.
Khalifah, R. G. (1971) *J. Biol. Chem. 246,* 2561.
Kiefer, L. L., Ippolito, J. A., Fierke, C. A., & Christianson, D. W. (1993) *J. Am. Chem. Soc. 115,* 12581.
Kisel, W., & Graf, G. (1972) *Phytochemistry 11,* 113.
Le Moual, H., Roques, B. P., Crine, P., & Boileau, G. (1993) *FEBS Lett. 324,* 196.
Majeau, N., & Coleman, J. R. (1992) *Plant Physiol. 100,* 1077.
McKale, A. G., Knapp, G. S., & Chan, S.-K. (1986) *Phys. Rev. B 33,* 841.
Medina, J. F., Wetterholm, A., Rådmark, O., Shapiro, R., Haeggström, J. Z., Vallee, B. L., & Samuelsson, B. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88,* 7620.
Monaco, H. L., Crawford, J. L., & Lipscomb, W. N. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75,* 5276.
Pan, T., & Coleman, J. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86,* 3145.
Pocker, Y., & Ng, J. S. Y. (1974) *Biochemistry 13,* 5116.
Provart, N. J., Majeau, N., & Coleman, J. R. (1993) *Plant Mol. Biol. 22,* 937.
Raines, C. A., Horsnell, P. R., Holder, C., & Lloyd, J. C. (1992) *Plant Mol. Biol. 20,* 1143.
Riddles, P. W., Blakely, R. L., & Zerner, B. (1983) *Methods Enzymol. 9,* 49.
Roeske, C. A., & Ogren, W. L. (1990) *Nucleic Acids Res. 18,* 3413.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Sansone, M., Via, G., George, G., Meitzner, G., Hewitt, R., & Marsch, J. (1991) in *X-ray Absorption Fine Structure* (Hasnain, S. S., Ed.) pp 656–658, Ellis Horwood, Ltd., W. Sussex, England.
Scott, R. A., Hahn, J. E., Doniach, S., Freeman, H. C., & Hodgson, K. O. (1982) *J. Am. Chem. Soc. 104,* 5364.
Severne, Y., Wieland, S., Schaffner, W., & Rusconi, S. (1988) *EMBO J. 7,* 2503.
Sung, Y. C., & Fuchs, J. A. (1988) *J. Biol. Chem. 263,* 14769.
Tashian, R. E. (1989) *Bioessays 10,* 186.
Tu, C., Paranawithana, S., Jewell, D. A., Tanhauser, S. M., LoGrasso, P. V., Wynns, G. C., Laipis, P. J., & Silverman, D. N. (1990) *Biochemistry 29,* 6400.
Vallee, B. L., & Auld, D. S. (1990) *Biochemistry 29,* 5647.
Vidgren, J., Liljas, A., & Walker, N. P. C. (1990) *Int. J. Biol. Macromol.12,*342.
Wilson D. K., & Quiocho, F. A. (1993) *Biochemistry 32,* 1689.