Biochemistry 1993, 32, 10944–10949

Structure and Energetics of a Non-Proline *cis*-Peptidyl Linkage in a Proline-202 → Alanine Carbonic Anhydrase II Variant[†,‡]

Neil B. Tweedy,[§] Satish K. Nair,[∥] Steven A. Paterno,[§] Carol A. Fierke,[*,§] and David W. Christianson[*,∥]

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

Received June 14, 1993; Revised Manuscript Received July 27, 1993[⊗]

ABSTRACT: The crystal structure of a human carbonic anhydrase II (CAII) variant, *cis*-proline-202 → alanine (P202A), has been determined at 1.7-Å resolution, indicating that the wild-type geometry, including the *cis*-peptidyl linkage, is retained upon substitution of proline by alanine. The CO_2 hydrase activity and affinity for sulfonamide inhibitors of P202A CAII are virtually identical to those of wild type. However, the substitution of *cis*-alanine for *cis*-proline decreases the stability of the folded state by ≈5 kcal mol^{-1} relative to both the unfolded state and an equilibrium intermediate in guanidine hydrochloride-induced denaturation. This destabilization can be attributed mainly to the less favorable *cis/trans* equilibrium of Xaa–alanine bonds compared to Xaa–proline bonds in the denatured state although other factors, including increased conformational entropy of the denatured state and decreased packing interactions in the native state, also contribute to the observed destabilization. The high catalytic activity of P202A CAII illustrates that unfavorable local conformations are nonetheless endured to satisfy the precise structural requirements of catalysis and ligand binding in the CAII active site.

The union of site-directed mutagenesis with high-resolution X-ray crystallographic methods allows for the detailed correlation of protein stability with three-dimensional protein structure. For example, recent studies dissect the contributions of specific hydrogen bonds (Alber et al., 1987; Fersht et al., 1985; Grutter et al., 1987), hydrophobic interactions (Matsumura et al., 1988; Karpusas et al., 1989), helix–dipole interactions (Mitchinson & Baldwin, 1986; Nicholson et al., 1988), and disulfide bridges (Perry & Wetzel, 1984; Sauer et al., 1986; Wells & Powers, 1986) to the stability of folded proteins. Additional studies illuminate entropic factors involving the protein backbone which affect the stability of the folded protein (Matthews, B. W., 1987). Amino acids which have greater backbone conformational entropy in the unfolded state (e.g., glycine) incur a greater energetic cost for conformational restriction during the folding process than those with bulkier side chains. Thus, incorporation of amino acids such as glycine within a polypeptide leads to thermodynamic destabilization. Consequently, amino acid replacements of the form Gly → Xaa[1] increase the thermal stability of T4 lysozyme (Matthews et al., 1987) and λ repressor (Hecht et al., 1986). Similarly, certain amino acid substitutions of the form Xaa → Pro increase protein stability, presumably by decreasing the conformational entropy of unfolding due to the incorporation of the conformationally rigid proline unit. For example, the Ala-82 → Pro substitution in T4 lysozyme results in a protein structure of increased stability (Matthews et al., 1987).

In the current study we focus on the structure–stability relationships of *cis*-peptidyl linkages in human carbonic anhydrase II (CAII). In native proteins, planar peptide bonds occur predominantly in the *trans* conformation; *cis* peptide bonds are rare (0.05%) except for *cis*-Xaa–proline linkages (6.5%) (MacArthur & Thornton, 1991; Stewart et al., 1990), which may reflect the decreased stability of *cis* isomers in non-proline versus Xaa–Pro peptide bonds (Radzicka et al., 1988; Drakenberg et al., 1972; Grathwohl & Wüthrich, 1976; Adler & Scheraga, 1990). Replacement of a *cis*-proline residue might be expected to produce either a *cis*-non-proline peptide bond or isomerization to a *trans* conformer with accompanying changes in the local backbone geometry. For instance, isomerization to a *trans* peptide has been observed in staphylococcal nuclease variants (Hodel et al., 1993; Evans et al., 1989) while the structure of the peptide linkage in RNase A (Schultz & Baldwin, 1992) and thioredoxin (Kelley & Richards, 1987) variants is uncertain. In RNase T1, folding kinetics are interpreted as consistent with isomerization to a *trans* peptide bond in a *cis*-Pro-35 → Asn variant (Kiefhaber et al., 1990) and retention of the *cis* peptide bond in a *cis*-Pro-39 → Ala variant (Mayr et al., 1993). Alternative folded conformations in which a proline residue occurs as either the *cis* or *trans* isomer have been observed in both staphylococcal nuclease (Evans et al., 1987, 1989) and calbindin D_{9K} wild-type proteins (Svensson et al., 1992). However, in this paper we describe the first example in which a *cis* peptide bond is definitively shown to be retained in a single amino acid variant, *cis*-proline-202 → alanine (P202A) carbonic anhydrase II, using X-ray crystallography.

The biological function of human carbonic anhydrase II (CAII) is to catalyze the interconversion of carbon dioxide and bicarbonate (Silverman & Lindskog, 1988). The crystal structure of this zinc metalloenzyme, refined at 1.54-Å resolution (Håkansson et al., 1992), shows two *cis*-proline residues, Pro-30 and Pro-202, which have been implicated in the slow-folding steps of this enzyme (Fransson et al., 1992).

[†] Supported by National Institutes of Health Grants GM40602 (C.A.F.) and GM45614 (D.W.C.). C.A.F. received an American Heart Association Established Investigator Award and a David and Lucile Packard Foundation Fellowship in Science and Engineering. D.W.C. is the recipient of an Alfred P. Sloan Research Fellowship and a Camille and Henry Dreyfus Teacher–Scholar Award.

[‡] Crystallographic coordinates have been deposited in the Brookhaven Protein Data Bank with reference code 1MUA.

[*] Authors to whom correspondence should be addressed.

[§] Duke University Medical Center.

[∥] University of Pennsylvania.

[⊗] Abstract published in *Advance ACS Abstracts*, October 1, 1993.

[1] Abbreviations: CAII, human carbonic anhydrase II; P202A, proline-202 → alanine; Tris, tris(hydroxymethyl)aminomethane; MES, 2-(*N*-morpholino)ethanesulfonic acid; TAPS, 3-[[tris(hydroxymethyl)methyl]-amino]propanesulfonic acid; Xaa, unspecified amino acid; Gdn-HCl, guanidine hydrochloride.

0006-2960/93/0432-10944$04.00/0 © 1993 American Chemical Society

The Pro-201–*cis*-Pro-202 dipeptide is contained in a type VI turn [a four-residue motif containing a *cis*-linked proline at the third position; see Richardson (1981)] which stabilizes the Ser-197–Cys-206 loop—this loop contains catalytically important enzyme residues (Krebs & Fierke, 1993). Here we describe the structure and energetics of the P202A variant, which forms a novel Pro-201–*cis*-Ala-202 peptide linkage. The formation of this non-proline *cis* peptide bond has no effect on the catalytic activity of CAII but significantly decreases the overall stability of the folded protein. Similar Pro–*cis*-Xaa linkages in refined protein structures (MacArthur & Thornton, 1991) include Pro-210–*cis*-Thr-211 in the eglin–subtilisin Carlsberg complex (Bode et al., 1987) and Pro-205–*cis*-Tyr-206 in the zinc protease carboxypeptidase A (Rees et al., 1983). Hence, the Pro-201–*cis*-Pro-202 linkage in wild-type CAII and its engineered variant Pro-201–*cis*-Ala-202 comprise unique structural motifs for structure–stability and structure–activity analyses.

MATERIALS AND METHODS

Enzyme Induction and Purification. BL21(DE3) cells containing a plasmid (pCAM-a1) encoding P202A CAII were grown, and CAII was induced as described (Krebs & Fierke, 1993). P202A CAII was purified using sulfonamide affinity chromatography (Osborne & Tashian, 1975) followed by an S-Sepharose fast-flow column chromatography (2.5 cm × 20 cm) where the enzyme is eluted with a linear ammonium sulfate gradient (0–0.5 M) in 20 mM MES, pH 7. The concentration of CAII was calculated either from absorbance (Tu & Silverman, 1982) or by stoichiometric titration with acetazolamide (Fierke et al., 1991).

Catalysis and Inhibition. Initial rates of CO_2 hydration and HCO_3^- dehydration were measured in a KinTek stopped-flow apparatus at 25 °C by the changing pH-indicator method (Khalifah, 1971) using the following buffer/indicator pairs: TAPS/*m*-cresol purple (pH 9, 578 nm) and MES/chlorophenol red (pH 6.1, 574 nm). The assay contained 50 mM buffer, 25 μM dye, and 0.1 mM (ethylenedinitrilo)tetraacetate with the ionic strength maintained at 0.1 M with sodium sulfate. The kinetic parameters and standard errors were determined using the software SYSTAT (Systat, Inc.) by fitting the data to the equation: rate = $k_{cat}[E][S]/(K_M + [S])$.

Dansylamide dissociation constants were determined by measuring an increase in fluorescence (excitation = 280 nm, emission = 470 nm) upon binding dansylamide to CAII (Fierke et al., 1991). Dissociation constants for acetazolamide were determined by measuring a decrease in fluorescence intensity at fixed concentrations of enzyme and dansylamide and varying concentrations of acetazolamide. Dissociation constants (Fierke et al., 1991) and error estimates were determined using the SYSTAT curve-fitting program.

Stability Measurements. CAII was incubated for 24 h at room temperature (23 °C) with varying concentrations of guanidine hydrochloride (Gdn-HCl) in 100 mM Tris–SO_4, pH 7.5, and 1 mM dithiothreitol, purged with nitrogen. Gdn-HCl-induced unfolding of CAII was followed by determining the CO_2 hydrase activity (Mårtensson et al., 1993) using a pH indicator assay (Brion et al., 1988) maintaining 40 mM Gdn-HCl in the assay. The fraction activity was calculated by subtracting the background rate and then dividing by the observed activity in the absence of denaturant. Denaturation was also monitored by measuring the ratio A_{292}/A_{260} as a function of Gdn-HCl concentration using a reference cell containing the same concentration of denaturant and buffer. A_{260} is a suitable internal standard since it is unaffected by complete denaturation (Edsall et al., 1966).

The transition curves for the solvent-induced inactivation of CAII were analyzed by assuming a two-state transition, N → I. The fraction activity (a_{obs}) was fit as a function of denaturant concentration [D] directly using eq 1 (Pantoliano et al., 1991), where $K_{NI} = \exp\{-(\Delta G°_{NI} - m_{NI}[D])/RT\}$. $a°_N$ and b_N and $a°_I$ and b_I are the *y*-intercept and slope of the pre- and posttransitional baselines, respectively, which describe the linear effect of solvent on the chemical potential of the N and I forms of these proteins. b_N and b_I are calculated from the wild-type and P202A CAII baselines, respectively. (However, fitting all the parameters directly does not significantly affect their value or error.) m_{NI} is the slope of the linear denaturation plot, $-d(\Delta G_{obs})/d[D]$, and $\Delta G°_{NI}$ is the free energy of unfolding at [D] = 0, or the *y*-intercept, and is equivalent to $\Delta G°_{NI}$ commonly measured by the linear extrapolation method (Cupo & Pace, 1983). This equation allows a nonlinear least-squares computer fit (SYSTAT, Systat, Inc.) to all the data and calculation of realistic errors (Santoro & Bolen, 1988). The transition curves for denaturation monitored by absorbance were analyzed on the assumption of a three-state model, N → I → U, using eq 2 (Hurle et al., 1986), where $A = A_{292}/A_{260}$ and $Z = (A_I - A_N)/(A_U - A_N)$. The free energies of unfolding are calculated both in the absence of denaturant ($\Delta G°$) and at the middle of the overlapping data in the transitions at 0.75 and 2.3 M Gdn-HCl ($\Delta G^{0.75}$ and $\Delta G^{2.3}$) (Matthews, C. R., 1987).

$$a_{obs} = \frac{\{(a°_I + b_I[D])(K_{NI}) + (a°_N + b_N[D])\}}{K_{NI} + 1} \quad (1)$$

$$F_{app} = (A_{obs} - A_N)/(A_U - A_N) = K_{NI}(Z + K_{IU})/(1 + K_{NI} + K_{NI}K_{IU}) \quad (2)$$

Crystallography. Crystals of P202A CAII were grown in 1 week at 4 °C by the sitting drop method using Charles Supper crystal growth chambers. A 10-μL drop containing 0.3 mM enzyme, 50 mM Tris-HCl (pH 8.5 at 4 °C), and 3 mM NaN_3 was added to a 10-μL drop of precipitant buffer containing 50 mM Tris-HCl (pH 8.0 at room temperature), 3 mM NaN_3, and 1.75–2.5 M ammonium sulfate, and the resulting 20-μL sitting drop was equilibrated against 1 mL of precipitant buffer in the crystallization reservoir. Both solutions were saturated with methylmercuric acetate (Tilander et al., 1965). Crystals were then back-soaked with 2.5 M ammonium sulfate and 50 mM Tris-HCl (pH 8.5) in order to dialyze out azide. P202A CAII crystallized isomorphously with the wild-type enzyme (Alexander et al., 1991) in space group $P2_1$ and exhibited unit cell parameters similar to those of the wild-type enzyme (a = 42.7 Å, b = 41.7 Å, c = 73.0 Å and β = 104.6°).

Intensity data to 1.7-Å resolution were collected from a single crystal using a Seimens X-100A multiwire area detector mounted on a Rigaku RU-200 X-ray generator. Data collection at room temperature was performed by the oscillation method with the crystal-to-detector distance fixed at 10 cm. Data frames of 0.08° oscillation about ω were collected with exposure times of 60 s/frame. Raw data frames were processed using the package of Durbin et al. (1986), and intensities were corrected for Lorentz and polarization effects. All reflections with $I < 2\sigma(I)$ were discarded. Scaling and reduction of replicate data and calculation of electron density maps were done with the PROTEIN suite of programs (Steigemann, 1974). Phases calculated from the refined structure of wild type CAII (Alexander et al., 1991), less the atoms of residue 202 and active site solvent molecules, were applied to structure factors obtained from the corrected

Table I: Data Collection and Refinement Statistics for P202A CAII

no. of crystals	1
no. of measured reflections	34796
no. of unique reflections	19327
max resolution (Å)	1.70
R_{merge}[a]	0.10
no. of water molecules in final cycle of refinement	93
no. of reflections used in refinement (6.5–1.70 Å)	18232
R factor[b]	0.188
completeness of data (%)	72
rms deviation from ideal bond lengths (Å)	0.009
rms deviation from ideal bond angles (deg)	1.3
rms deviation from ideal planarity (Å)	0.012
rms deviation from ideal chirality (Å^3)	0.117

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

Table II: Catalytic Activity and Inhibitor Binding

		wild type	Pro-202 → Ala
CO_2 hydration[a]	k_{cat}/K_M (μM^{-1} s^{-1})	110 ± 10	112 ± 4
	k_{cat} (μs^{-1})	1.0 ± 0.1	1.3 ± 0.1
	$K_M^{CO_2}$ (mM)	11 ± 1	10 ± 2
HCO_3^- dehydration[b]	k_{cat}/K_M (μM^{-1} s^{-1})	13 ± 1	8 ± 1
	k_{cat} (μs^{-1})	0.56 ± 0.04	0.36 ± 0.05
	$K_M^{HCO_3^-}$ (mM)	45 ± 9	45 ± 12
inhibitor binding[c]	K_{DNSA} (μM)	1.2 ± 0.1[d]	1.1 ± 0.1
	K_{ACET} (nM)	8.2 ± 0.3[d]	19 ± 1

[a] Activity measured as a function of [CO_2] in 50 mM TAPS, pH 9.0, I = 0.1, with Na_2SO_4, 25 °C. [b] Activity measured as a function of [HCO_3^-] in 50 mM MES, pH 6.2, I = 0.1, with Na_2SO_4, 25 °C. [c] K_{DNSA} and K_{ACET} are the dissociation constants for dansylamide and acetazolamide, respectively, determined in 20 mM potassium phosphate, pH 7.4, 25 °C. [d] Taken from Krebs et al. (1991).

intensity data in order to generate difference electron density maps using Fourier coefficients $2|F_o| - |F_c|$ and $|F_o| - |F_c|$. All model building was done with the software FRODO (Jones, 1985). Initial inspection of difference electron density maps unambiguously revealed that the *cis*-peptidyl linkage was maintained in P202A CAII. The reciprocal space least-squares algorithm PROLSQ (Hendrickson, 1985) was used to refine the model. Active site water molecules and the methyl group of Ala-202 were added when the crystallographic R factor dropped below 0.20. Solvent molecules with thermal B factors greater than 50 Å^2 were deleted from the model. Pertinent refinement statistics are recorded in Table I. Atomic coordinates have been deposited in the Brookhaven Data Bank (reference code 1MUA) (Bernstein et al., 1977).

RESULTS AND DISCUSSION

Functional screens of single amino acid variants at the highly conserved *cis*-Pro-202 residue in CAII (Hewett-Emmett & Tashian, 1991) suggest that the major effect of these substitutions is to increase thermal lability (Krebs & Fierke, 1993). To further explore the functional properties of one of these variants, P202A CAII, we determined the kinetic parameters for CO_2 hydration and HCO_3^- dehydration as well as the dissociation constants for two sulfonamide inhibitors, dansylamide and acetazolamide (Table II). These data indicate that the substitution of alanine at position 201 has little or no effect on either the catalytic efficiency or the binding of inhibitors.

To investigate the stability of the *cis*-P202A CAII, we characterized the Gdn-HCl-induced denaturation of CAII. In both wild type and variant, two transitions are observed spectroscopically (Figure 1B), indicating a stable intermediate

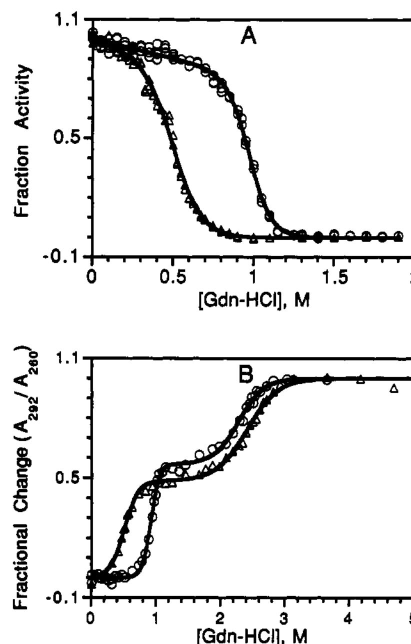

FIGURE 1: Gdn-HCl denaturation of wild-type (○) and P202A (△) CAIIs assayed by either (A) CO_2 hydrase activity or (B) the ratio of absorbance at 292 nm to 260 nm after CAII [0.01 (A) or 0.2 mg/mL (B)] was incubated for 24 h at 23 °C in 0.1 M Tris–SO_4, pH 7.5, 1 mM dithiothreitol, and 0–5 M Gdn-HCl. The solid lines represent the nonlinear least-squares fit of the data using eq 1 and 2 derived for either a two- (A) or three-state (B) model, respectively, and the parameters in Table II.

state (Martensson et al., 1993; Henkens et al., 1982). Unfolding was analyzed (see Materials and Methods) as a two-step (three-state) process made up of an N ↔ I transition observable by both activity and spectroscopic measurements (panels A and B of Figure 1) and an I ↔ U transition which can be characterized spectroscopically (Figure 1B). The parameters for these fits, including the transition midpoints and free energy for unfolding, are listed in Table III. Substitution of alanine for Pro-202 has little or no effect on the relative stability of the I and U states; the concentrations of Gdn-HCl at the transition midpoints are similar, and $\Delta\Delta G°_{IU}$ ($\Delta G°_{IU\text{-}WT} - \Delta G°_{IU\text{-}P202A}$) and $\Delta\Delta G^{2.3}_{IU}$ are 1 ± 1 and 0.2 ± 0.1 kcal mol^{-1}, respectively. [This latter comparison removes the long extrapolation to 0 M Gdn-HCl (Cupo & Pace, 1983; Matthews, C. R., 1987).] However, the midpoint concentration of Gdn-HCl for the first transition (N ↔ I), as assayed by both activity and absorbance, is decreased significantly (Table III) for P202A CAII. Therefore, this substitution destabilizes the native state relative to the intermediate and unfolded states compared to wild type by $\Delta\Delta G°_{NI}$ = 5.1 ± 0.5 or $\Delta\Delta G^{0.75}_{NI}$ = 3.5 ± 0.2 kcal mol^{-1}. This is consistent with the ≈10^3 increase in thermal lability observed for variants at position 202 (Krebs & Fierke, 1993).

A difference electron density map, calculated with phases from the final model less all atoms of residues Pro-201, Ala-202, and Leu-203, illustrates that the *cis*-peptidyl linkage is maintained between Pro-201 and Ala-202 (Figure 2). The overall structure of P202A CAII is very similar to that of the wild-type enzyme: the rms deviation in atomic coordinates is 0.14 Å for Cα atoms (wild-type enzyme coordinates are superimposed in Figure 2). The conformations of residues flanking residue 202 change very slightly between wild-type and P202A CAIIs, and there are only minor local changes in the backbone conformation of Ala-202 relative to the wild-type enzyme (Table I). The P202A substitution does not disrupt the Ser-197–Cys-206 loop, so we conclude that proline is not obligatory for the third residue of a type VI turn.

Table III: Stability Parameters for Gdn-HCl-Induced Unfolding[a]

variant	$\Delta G°_{NI}$	m_{NI}	$C^m{}_{NI}$[b]	$\Delta G^{0.75}{}_{NI}$[c]	$\Delta G°_{IU}$	m_{IU}	$C^m{}_{IU}$	$\Delta G^{2.3}{}_{IU}$[d]
wild type	8.4 ± 0.4[e]	8.6 ± 0.4[e]	0.97 ± 0.01[e]	1.9 ± 0.1[e]				
	8.3 ± 0.5[f]	9.0 ± 0.6[f]	0.92 ± 0.01[f]	1.5 ± 0.1[f]	7.1 ± 0.6[f]	3.1 ± 0.2[f]	2.29 ± 0.02[f]	−0.05 ± 0.05[f]
P202A	3.3 ± 0.1[e]	6.6 ± 0.3[e]	0.50 ± 0.01[e]	−1.6 ± 0.1[e]				
	2.8 ± 0.3[f]	5.5 ± 0.5[f]	0.51 ± 0.01[f]	−1.3 ± 0.1[f]	6.0 ± 0.4[f]	2.5 ± 0.2[f]	2.41 ± 0.02[f]	0.28 ± 0.05[f]

[a] Units are as follows: ΔG, kcal mol^{-1}; m, kcal mol^{-1} M^{-1}; C^m, M. [b] C^m represents the concentration of Gdn-HCl at the transition midpoint. [c] At 0.75 M Gdn-HCl. [d] At 2.3 M Gdn-HCl. [e] Measured by activity changes. [f] Measured by absorbance changes.

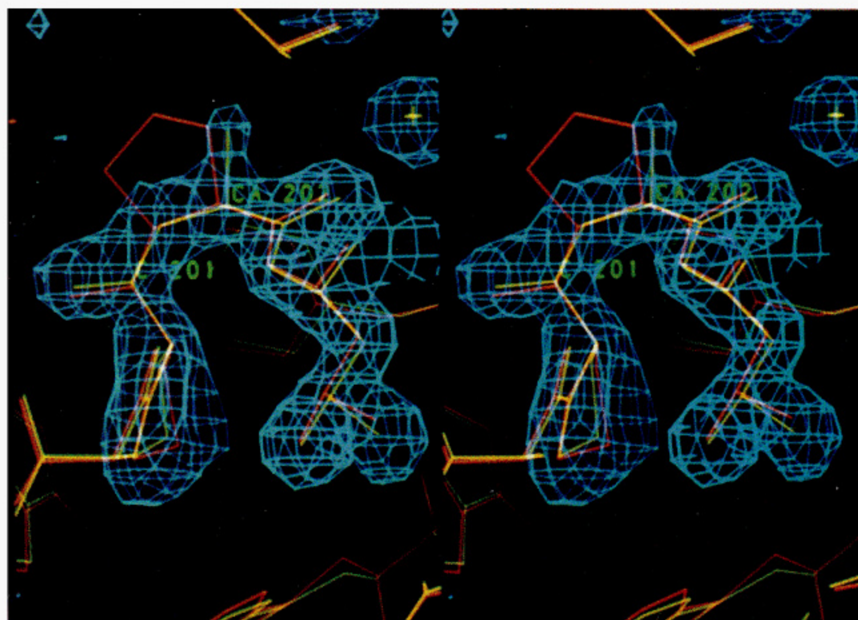

FIGURE 2: Difference electron density map of Pro-202 → Ala CAII calculated with Fourier coefficients ($|F_o| - |F_c|$) and phases calculated from the final model less the atoms of Pro-201, Ala-202, and Leu-203 and water molecule 354. Refined atomic coordinates (yellow) are superimposed on the map (contoured at 2.6σ); residues 201 and 202 are indicated, and water 354 appears as a star. Note that the *cis*-peptidyl linkage is maintained between Pro-201 and Ala-202. For reference, the coordinates of the wild-type enzyme are superimposed (red).

Interestingly, the Ser-197–Cys-206 loop exhibits notable plasticity in certain protein engineering experiments involving residues contained in this loop (Alexander et al., 1993; Ippolito & Christianson, 1993) as well as residues adjacent to this loop (Fierke et al., 1991; Alexander et al., 1991); indeed, the compensatory plastic response of this loop to single-point mutations in the CAII amino acid sequence can trigger structural changes of greater than 1 Å in certain loop residues. However, this loop is not sufficiently plastic to allow the Pro-201–Ala-202 peptide linkage to isomerize to the *trans* conformation.

Active site solvent structure is quite similar in wild-type and P202A CAIIs, and well-defined electron density characterizes the non-protein zinc ligand and the "deep" water molecule at the mouth of the hydrophobic pocket. Both of these solvent molecules are catalytically important, so their structural invariance accounts for the catalytic similarities between wild-type and P202A CAIIs: the deep water molecule is believed to be displaced by substrate CO_2 prior to its nucleophilic attack by the zinc-bound hydroxide atom (Lindahl et al., 1993). However, the steric void arising from the loss of the pyrrolidine ring in P202A CAII allows for a water molecule to hydrogen bond to the main chain carbonyl of residue 202 with good stereochemistry (water 354 in Figure 2).

In the wild-type enzyme, the backbone torsion angles for Pro-202 are $\phi = -68°$ and $\psi = +6°$, which deviate slightly from the average values of −86° and +1°, respectively, tabulated for *cis*-prolines contained in refined protein structures (MacArthur & Thornton, 1991). The backbone torsion angles of Ala-202 in P202A CAII are $\phi = -74°$ and $\psi = +11°$; these values are outside of sterically favorable parameters but nonetheless allowable under circumstances of steric strain, as estimated from Ramachandran maps of *cis*-polyalanine (Ramachandran & Venkatachalam, 1968). Hence, the folded conformation of active site residue 202 is maintained in a strained conformation despite the substitution of alanine for proline. Similarly unfavorable torsion angle values have been found in 17 other refined protein structures containing non-prolyl *cis*-peptidyl linkages (Stewart et al., 1990). However, only five of these structures have been determined to high resolution (≤2.0 Å with crystallographic R values < 0.2): the eglin–subtilisin complex (Bode et al., 1987), neurotoxin β (Tsernglou et al., 1978), carboxypeptidase A (Rees et al., 1983), dihydrofolate reductase (Bolin et al., 1982), and *Staphylococcus aureus* PCI β-lactamase (Herzberg, 1991). Herzberg and Moult (1991) point out that such steric strain typically characterizes regions of protein structure implicated in function. As illustrated by the engineering of Pro-202, torsional strain is tolerated to satisfy the precise structural requirements of ligand binding and catalysis at protein functional sites.

In aqueous buffer, native wild-type CAII is stabilized by 8.4 kcal mol^{-1} relative to a folding intermediate while the stability of the P202A variant is decreased to 3.3 kcal mol^{-1}. This ≈5 kcal mol^{-1} difference in stabilities may arise from energetic differences in both the folded and unfolded forms of the two variants. To dissect the peptidyl isomerization reaction from unfolding, we can express the unfolding process as a pair of linked reactions using a thermodynamic cycle (see Figure 3). Our spectroscopic techniques (Figure 1) measure the folding characteristics of CAII, irrespective of the

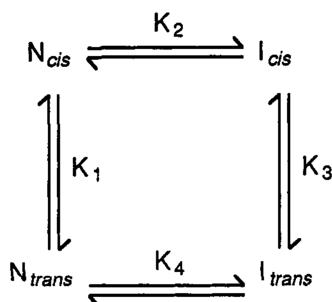

FIGURE 3: Thermodynamic cycle illustrating the linkage between folding and peptidyl isomerization in CAII. N and I are the native and intermediate states, respectively, observed in the equilibrium denaturation of CAII distinguished with respect to the *cis/trans* conformation of the Xaa–peptide at position 202. In this figure, $K_1 = N_{trans}/N_{cis}$, $K_2 = I_{cis}/N_{cis}$, $K_3 = I_{trans}/I_{cis}$, and $K_4 = I_{trans}/N_{trans}$. Each of these constants depends on the amino acid at position 202 and will be distinguished by subscripts in the text, e.g., $K_{1,\mathrm{Pro}}$. In the text these equilibria are discussed in terms of free energy changes using a corresponding notation, e.g., $\Delta G_{1,\mathrm{Pro}}$.

conformation at position 202; therefore, the *observed* equilibrium constant for this reaction can be described by the ratio of $[I_{tot}]/[N_{tot}]$ and expressed in terms of the unfolding (K_2) and isomerization (K_1, K_3) equilibria as shown in eq 3.

$$K_{obs} = (I_{cis} + I_{trans})/(N_{cis} + N_{trans}) = K_2(1 + K_3)/(1 + K_1) \approx K_2(1 + K_3) \quad (3)$$

Since the X-ray crystallographic data indicate that the *cis* conformer is predominant in the native state of both wild type and P202A, we may postulate that $K_1 \ll 1$, which simplifies eq 3. Therefore, the observed free energy change for unfolding can be represented as a combination of the *cis/trans* isomerization free energy in the unfolded intermediate state and the free energy for folding CAII with the native *cis* conformer at position 202. The free energy for isomerization to a *trans* Xaa–Pro peptide bond from a *cis* conformer, $\Delta G_{3,\mathrm{Pro}}$, has been estimated at –0.5 kcal mol^{-1} in both small peptide models and denatured RNase A (Grathwohl & Wuthrich, 1976; Adler & Scheraga, 1990). This allows the estimation of the free energy for unfolding for wild-type CAII, $\Delta G_{2,\mathrm{Pro}} \approx 9$ kcal mol^{-1}, from the observed ΔG for unfolding (using eq 3). However, reliable estimates of the *cis/trans* equilibrium in Xaa–Ala peptide bonds are not available. $\Delta G \approx -2.5$ kcal mol^{-1} was measured for the *cis/trans* isomerization of the peptide analog *N*-methylacetamide (Radzicka et al., 1988; Drakenberg et al., 1972), setting a lower limit for the ΔG for isomerization of non-proline peptide bonds because of the existence of additional unfavorable steric interactions in *cis*-peptidyl linkages (Brandts et al., 1975; Kim & Baldwin, 1982). However, the unfolding energy for P202A, $\Delta G_{2,\mathrm{Ala}}$,can be estimated by adjusting the wild-type value, $\Delta G_{2,\mathrm{Pro}}$, for changes introduced by the substitution of *cis*-alanine for *cis*-proline at position 202. The difference free energy for unfolding ($\Delta\Delta G_2 = \Delta G_{WT} - \Delta G_{variant}$) for substitution of *trans*-proline with alanine, determined for eight staphylococcus nuclease and T4 lysozyme variants, ranges from –0.1 to 1.0 kcal mol^{-1} with an average of ≈0.4 kcal mol^{-1} (Green et al., 1992; Chen et al., 1992). Additionally, in staphylococcus nuclease these changes in stability correlate with the fractional solvent accessibility of the side chain (Green et al., 1992). In the case of Pro-202, the side chain is a surface residue so the fractional solvent accessibility is moderately high (0.55); therefore, the destabilizing effect of substitution of alanine for proline at position 202 should be smaller than 0.4 kcal mol^{-1}. This value reflects entropic differences in the unfolded state (Matthews et al., 1987; Yun et al., 1991) as well as changes in the packing interactions and solvation of the folded state. Assuming that $\Delta\Delta G_2 = 0.4$ kcal mol^{-1} and, therefore, $\Delta G_{2,\mathrm{Ala}} \approx 8.6$ kcal mol^{-1}, the free energy for the *cis/trans* isomerization of the Xaa–Ala bond at position 202 in the folding intermediate, $\Delta G_{1,\mathrm{Ala}}$, is estimated at ≈–5 kcal mol^{-1}. Since there is no energetic difference between the intermediate and denatured forms upon substitution of Ala for Pro-202 (Table III), we suggest that this value also represents the *cis/trans* equilibrium in the denatured state as well. This number may also be a reasonable value for this equilibrium in any polypeptide. This would suggest that 0.02% non-proline *cis* peptide bonds should be observed in proteins if their distribution reflects stability alone [compared to 0.05% tabulated from known structures (MacArthur & Thornton, 1991; Stewart et al., 1990)].

In P202A CAII the *cis* linkage is preferred over a conformational change to *trans*, suggesting that intramolecular interactions between the protein and the Ser-197–Cys-206 loop play an important role in stabilizing the *cis* conformation. The observed decreases in stability of the native state caused by substitution of alanine for a *cis*-proline in other proteins range from ≈5 kcal mol^{-1} for the P39A variant of RNase T1 (Mayr et al., 1993) to 2–3 kcal mol^{-1} for RNase A variants (Schultz & Baldwin, 1992), reflecting either a backbone conformational change or insertion of a non-proline *cis* peptide bond. Conformational energy calculations for incorporation of *cis*-proline linkages in bovine pancreatic trypsin inhibitor suggest that the stability of the *cis* stereochemistry is highly context dependent (Levitt, 1981).

The stability of the equilibrium folding intermediate in CAII, relative to unfolded protein, is insensitive to the P202A substitution, suggesting a *trans* linkage at position 202 in the intermediate. This is consistent with data on the accessibility of introduced thiol groups (Mårtensson et al., 1993), indicating that the folding intermediate has native-like secondary structure only in the central part of the β-sheet. The detailed, correct geometry necessary for high catalytic activity is achieved only in the native state.

Cis/trans isomerization of proline peptide bonds has been implicated as rate-limiting for slow-folding species ($\tau \approx$ 10–100 s) in a variety of proteins, including CAII (Brandts et al., 1975; Kim & Baldwin, 1982; Fransson et al., 1992). While little is known about the rate of *cis/trans* isomerization of non-proline peptidyl linkages (Radzicka et al., 1988; Drakenberg et al., 1972), a recent kinetic study of Pro-202 → Asn CAII indicates that the rate of refolding ($\tau \approx 800$ s) is unaffected by this substitution (Fransson et al., 1992). Assuming a *cis* peptide in this variant [as suggested by our results (Krebs & Fierke, 1993)], the isomerization must be at least as fast as the observed refolding rate, thus setting a lower limit for the isomerization rate of the Pro-201–Asn-202 peptide bond in CAII. The extremely slow refolding rates ($\tau \approx 4000$ s) for the *cis*-Pro-39 → Ala variant of RNase T1 may reflect the rate constant for *trans* to *cis* isomerization of the Tyr-38–Ala-39 bond, assuming the *cis*-peptidyl bond is retained in the variant (Mayr et al., 1993; Mayr & Schmidt, 1993). Complete refolding kinetics of the P202A variant may provide an opportunity to further investigate *cis/trans* isomerization of non-proline peptide bonds. Finally, since significant conformational changes often accompany substitutions at *cis*-prolines, we emphasize that three-dimensional crystal structures are essential for dissecting protein stabilization and folding pathways using such variants.

ACKNOWLEDGMENT

We thank Drs. D. E. Stewart, T. G. Oas, and J. F. Krebs for helpful discussions.

REFERENCES

Adler, M., & Scheraga, H. A. (1990) *Biochemistry 29*, 8211–8216.

Alber, T., Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., & Matthews, B. W. (1987) *Nature (London) 330*, 41–46.

Alexander, R. S., Nair, S. K., & Christianson, D. W. (1991) *Biochemistry 30*, 11064–11072.

Alexander, R. S., Kiefer, L. L., Fierke, C. A., & Christianson, D. W. (1993) *Biochemistry 32*, 1510–1518.

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. (1977) *J. Mol. Biol. 112*, 535–542.

Bode, W., Papamokos, E., & Musil, D. (1987) *Eur. J. Biochem. 166*, 673–692.

Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C., & Kraut, J. (1982) *J. Biol. Chem. 257*, 13650–13662.

Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953–4963.

Brion, L. P., Schwartz, J. H., Zavilowitz, B. J., & Schwartz, G. J. (1988) *Anal. Biochem. 175*, 289–297.

Chen, B.-L., Baase, W. A., Nicholson, H., & Schellman, J. A. (1992) *Biochemistry 31*, 1464–1476.

Cupo, J. F., & Pace, C. N. (1983) *Biochemistry 22*, 2654–2658.

Drakenberg, T., Dahlqvist, K.-I., & Forsén, S. (1972) *J. Phys. Chem. 76*, 2178–2183.

Durbin, R. M., Burns, R., Moulai, J., Metcalf, P., Freymann, D., Blum, M., Anderson, J. E., Harrison, S. C., & Wiley, D. C. (1986) *Science 232*, 1127–1132.

Edsall, J. T., Mehta, S., Myers, D. V., & Armstrong, J. M. (1966) *Biochem. Z. 345*, 9–36.

Evans, P. A., Dobson, C. M., Kautz, R. A., Hatfull, G., & Fox, R. O. (1987) *Nature (London) 329*, 266–268.

Evans, P. A., Kautz, R. A., Fox, R. O., & Dobson, C. M. (1989) *Biochemistry 28*, 362–370.

Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature (London) 314*, 235–238.

Fierke, C. A., Calderone, T., & Krebs, J. F. (1991) *Biochemistry 30*, 11054–11063.

Fransson, C., Freskgard, P. O., Herbertsson, H., Johansson, A., Jonasson, P., Martensson, L. G., Svensson, M., Jonsson, B. H., & Carlsson, U. (1992) *FEBS Lett. 296*, 90–94.

Grathwohl, C., & Wüthrich, K. (1976) *Biopolymers 15*, 2025–2041.

Green, S. M., Meeker, A. K., & Shortle, D. (1992) *Biochemistry 31*, 5717–5728.

Grutter, M. G., Gray, T. M., Weaver, L. H., Alber, T., Wilson, K., & Matthews, B. W. (1987) *J. Mol. Biol. 197*, 315–329.

Håkansson, K., Carlsson, M., Svensson, L. A., & Liljas, A. (1992) *J. Mol. Biol. 227*, 1192–1204.

Hecht, M. H., Sturtevant, J. M., & Sauer, R. T. (1986) *Proteins: Struct., Funct., Genet. 1*, 43–46.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Henkens, R. W., Kitchell, B. B., Lottich, S. C., Stein, P. J., & Williams, T. J. (1982) *Biochemistry 21*, 5918–5923.

Herzberg, O. (1991) *J. Mol. Biol. 217*, 701–719.

Herzberg, O., & Moult, J. (1991) *Proteins 11*, 223–229.

Hewett-Emmett, D., & Tashian, R. E. (1991) in *The Carbonic Anhydrases: Cellular Physiology and Molecular Genetics* (Dodgson, S. J., Tashian, R. E., Gros, G., & Carter, N. D., Eds.), pp 15–32, Plenum Press, New York.

Hodel, A., Kautz, R. A., Jacobs, M. D., & Fox, R. O. (1993) *Protein Sci. 2*, 838–850.

Hurle, M. R., Tweedy, N. B., & Matthews, C. R. (1986) *Biochemistry 25*, 6356–6360.

Ippolito, J. A. & Christianson, D. W. (1993) *Biochemistry* (in press).

Jones, T. A. (1985) *Methods Enzymol. 115*, 157–171.

Karpusas, M., Baase, W. A., Matsumura, M., & Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8237–8241.

Kelley, R. F., & Richards, F. M. (1987) *Biochemistry 26*, 6765–6774.

Khalifah, R. G. (1971) *J. Biol. Chem. 246*, 2561–2573.

Kiefhaber, T., Grunert, H. P., Hahn, U., & Schmid, F. X. (1990) *Biochemistry 29*, 6475–6480.

Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–489.

Krebs, J. F., & Fierke, C. A. (1993) *J. Biol. Chem. 268*, 948–954.

Krebs, J. F., Fierke, C. A., Alexander, R. S., & Christianson, D. W. (1991) *Biochemistry 30*,9153–9160.

Levitt, M. (1981) *J. Mol. Biol. 145*, 251–263.

Lindahl, M., Svensson, A., & Liljas, A. (1993) *Proteins: Struct., Funct., Genet. 15*, 177–182.

MacArthur, M. W., & Thornton, J. M. (1991) *J. Mol. Biol. 218*, 397–412.

Mårtensson, L.-G., Jonsson, B.-H., Freskgård, P.-O., Kihlgren, A., Svensson, M., & Carlsson, U. (1993) *Biochemistry 32*, 224–231.

Matsumura, M., Bechtel, W. J., & Matthews, B. W. (1988) *Nature (London) 334*, 406–410.

Matthews, B. W. (1987) *Biochemistry 26*, 6885–6888.

Matthews, B. W., Nicholson, H., & Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663–6667.

Matthews, C. R. (1987) *Methods Enzymol. 154*, 498–511.

Mayr, L. M., & Schmidt, F. X. (1993) *J. Mol. Biol. 231*, 913–926.

Mayr, L. M., Landt, O., Hahn, U., & Schmidt, F. X. (1993) *J. Mol. Biol. 231*, 897–912.

Mitchinson, C., & Baldwin, R. L. (1986) *Proteins: Struct., Funct., Genet. 1*, 23–33.

Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature (London) 336*, 651–656.

Osborne, W. R. A., & Tashian, R. E. (1975) *Anal. Biochem. 64*, 297–303.

Pantoliano, M. W., Bird, R. E., Johnson, S., Asel, E. D., Dodd, S. W., Wood, J. F., & Hardman, K. D. (1991) *Biochemistry 30*, 10117–10125.

Perry, L. J., & Wetzel, R. (1984) *Science 226*, 555–557.

Radzicka, A., Pedersen, L., & Wolfenden, R. (1988) *Biochemistry 27*, 4538–4541.

Ramachandran, G. N., & Venkatachalam, C. M. (1968) *Biopolymers 6*, 1255–1262.

Rees, D. C., Lewis, M., & Lipscomb, W. N. (1983) *J. Mol. Biol. 168*, 367–387.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063–8068.

Sauer, R. T., Hehir, K., Stearman, R. S., Weiss, M. A., Jeitler-Nilsson, A., Suchanek, E. G., & Pabo, C. O. (1986) *Biochemistry 25*, 5992–5998.

Schultz, D. A., & Baldwin, R. L. (1992) *Protein Sci. 1*, 910–916.

Silverman, D. N., & Lindskog, S. (1988) *Acc. Chem. Res. 21*, 30–36.

Steigemann, W. (1974) Ph.D. Dissertation, Max-Planck-Institut für Biochemie, Germany.

Stewart, D. E., Sarkar, A., & Wampler, J. E. (1990) *J. Mol. Biol. 214*, 253–260.

Svensson, L. A., Thulin, E., & Forsen, S. (1992) *J. Mol. Biol. 223*, 601–606.

Tilander, B., Strandberg, B., & Fridborg, K. (1965) *J. Mol. Biol. 12*, 740–760.

Tsernoglou, D., Petsko, G. A., & Hudson, R. A. (1978) *Mol. Pharmacol. 14*, 710–716.

Tu, C. K., & Silverman, D. N. (1982) *Biochemistry 21*, 6353–6360.

Wells, J. A., & Powers, D. B. (1986) *J. Biol. Chem. 261*, 6564–6570.

Yun, R. H., Anderson, A., & Hermans, J. (1991) *Proteins 10*, 219–228.