Folding and Stability of the N-Terminus of Human Carbonic Anhydrase II†

Göran Aronsson,‡ Lars-Göran Mårtensson,‡,|| Uno Carlsson,§ and Bengt-Harald Jonsson*,‡

Department of Biochemistry, Umeå University, S-901 87 Umeå, Sweden, and IFM-Department of Chemistry, Linköping University, S-581 83 Linköping, Sweden

Received July 7, 1994; Revised Manuscript Received October 18, 1994®

ABSTRACT: Truncations and mutations in the N-terminus of human carbonic anhydrase II were constructed in order to establish what role this part of the protein plays in the folding and stability of the protein. When incubated in various concentrations of guanidine hydrochloride (GuHCl), HCAII unfolds in two transitions, with an intermediate state at about 1.3 M GuHCl. N-Terminal truncations of 5, 17, or 24 amino acid residues destabilize the native state by 4−5 kcal/mol, relative to the intermediate state, but these amino acid residues have virtually no effect on the stability of the intermediate state relative to the unfolded state. These truncated variants of HCAII still have a high enzymatic activity. Deletion of 28 or more amino acid residues, however, results in inactive enzyme variants. The rates at which the active site is formed are practically unaffected by the removal of the 24-amino acid segment, *i.e.*, the active site forms independently of the N-terminus. By using the tryptophans in positions 5 and 16 as intrinsic probes, we conclude that the structure of the N-terminal region is formed very late in folding. The results strongly indicate that this process is dependent on the prior formation of an enzymatically active native-like structure of the rest of the protein.

Human carbonic anhydrase II (HCAII) is a structurally and functionally well-defined protein; it is a zinc-containing enzyme that catalyzes the interconversion of carbon dioxide and water into bicarbonate and protons (Lindskog & Liljas, 1993; Liljas *et al.*, 1994). It is a monomeric, single-domain protein with a molecular mass of 29 300 Da, and its X-ray structure has been resolved to 1.54 Å (Håkansson *et al.*, 1992). HCAII has some helical structure and a dominating β-sheet that extends throughout the entire molecule (Figure 1).

Several parts of HCAII have been studied before, from a folding point of view: the role of the *cis*-prolines has been analyzed (Fransson *et al.*, 1992); the folding of the C-terminus (Freskgård *et al.*, 1991) and its influence on stability (Carlsson *et al.*, 1974) have been studied; several amino acid residues in the central β-sheet have been subjected to mutagenesis and used as probes in folding and stability studies (Mårtensson *et al.*, 1993); and the individual contribution of the seven tryptophans to the fluorescence spectrum (Mårtensson *et al.*, 1995) and the circular dichroism spectrum (Freskgård *et al.*, 1994) have been described. The rather detailed knowledge about the folding pathway of HCAII makes it an excellent model for further studies of different aspects of protein folding. We report here on the role of the N-terminus in the folding and stability of HCAII.

Analysis of the protein structure reveals that the N-terminus of HCAII is highly exposed to solvent. This is similar to the N-terminus of the highly homologous isoenzyme HCAI, which, according to a structural domain analysis (Janin & Wodak, 1983), has few contacts with the rest of the protein. Two 3_{10}-helices are located in the N-terminus and comprise amino acid residues 13−18 and 21−24 (Figure 1). One of the two hydrophobic clusters of HCAII is located in the N-terminus, and it involves four aromatic residues, *i.e.*, Trp5, Tyr7, Trp16, and Phe20 (Eriksson *et al.*, 1988). Considering this, we asked: How does the N-terminus contribute to the total stability of the protein? Is it necessary for folding and enzymatic activity? When, during refolding, are the structures in the N-terminus formed?

We used recombinant DNA techniques to construct five variants of HCAII that lack the first 5, 17, 24, 28, and 31 amino acid residues in the N-terminus (Table 1). These variants were named Trunc5, Trunc17, and so forth. Three mutants of HCAII, *i.e.*, W5F, W16F, and the double mutant W5F/W16F, were also included in this investigation. All of these variants and mutants were characterized using fluorescence spectroscopy, circular dichroism spectroscopy, and measurements of enzymatic activity.

MATERIALS AND METHODS

Chemicals. Guanidine hydrochloride (GuHCl, sequanal grade) was purchased from Pierce (Rockford, IL), and its concentration was determined refractometrically (Nozaki, 1972). Tris(hydroxymethyl)aminomethane (reagent grade) was obtained from Sigma (St. Louis, MO). The restriction endonuclease *Nco*I and T4 DNA ligase were both purchased from Boehringer Mannheim GmbH (Mannheim, Germany).

Construction of N-Terminally Truncated HCAII. We used plasmid pACA (Nair *et al.*, 1991) in a modified form, previously described by us (Mårtensson *et al.*, 1993), which codes for serine rather than cysteine at position 206. Henceforth, this plasmid will be referred to as pCApwt. We used the mutant that this plasmid produces, HCAII C206S, as our reference in this work and will refer to it as $HCAII_{pwt}$.

† This work was supported by grants from the Swedish National Board for Industrial and Technical Development (U.C. and B.-H.J.), the Swedish Natural Science Research Council (B.-H.J.), "Magn. Bergwalls stiftelse" (B.-H.J.), "Carl Tryggers Stiftelse för vetenskaplig forskning" (U.C.), "Sven and Lilly Lawskis fond" (G.A. and L.-G.M.), and "J C Kempes Minnes Stipendiefond" (G.A. and L.-G.M.).

* Author to whom correspondence should be addressed.

‡ Umeå University.

§ Linköping University.

|| Present address: Howard Hughes Medical Institute, Institute of Molecular Biology, University of Oregon, Eugene, OR 97403.

® Abstract published in *Advance ACS Abstracts,* January 15, 1995.

0006-2960/95/0434-2153$09.00/0

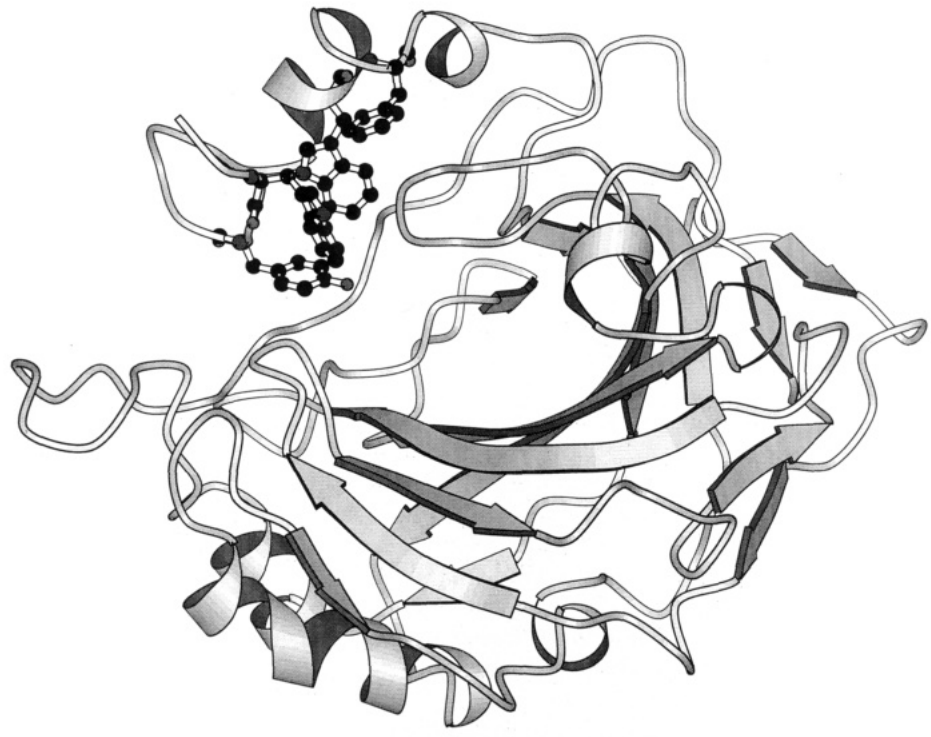

FIGURE 1: Schematic representation of the polypeptide backbone of human carbonic anhydrase II. The four amino acid residues involved in the N-terminal aromatic cluster are displayed as balls and sticks. The figure was produced using the program Molscript (Kraulis, 1991), and the coordinates were kindly provided by Dr. Kjell Håkansson.

Table 1: N-Terminal Amino Acid Sequence of $HACII_{pwt}$ and the Determined Sequences of Three Truncated Variants

$HCAII_{pwt}$	AHHWGYGKHNGPEHWHKNFPIAKGERQSPVNIDTHTAKY
Trunc5[a]	GYGKH**********************************
Trunc17[a]	MNFPI******************
Trunc24[a]	GERQS*****************

pCApwt contains one unique sequence for the restriction enzyme *Nco*I at the start of the HCAII-coding region. The construction of an N-terminally truncated HCAII variant started with the introduction of yet another site for *Nco*I, farther into the HCAII−cDNA region, using site-directed mutagenesis. The mutant plasmids thereafter were cleaved by *Nco*I, and this was followed by dilution and addition of T4 DNA ligase. The dilution served the purpose of promoting the religation of plasmids without reinsertion of the small fragments, thereby removing the DNA that coded for the N-terminal amino acid segments of interest. The new plasmids were transfected into *Escherichia coli* BL21/DE3 (Studier & Moffat, 1986); in this host cell, T7 RNA polymerase is supplied from a chromosomal copy of the polymerase gene, which is under the control of the isopropyl β-D-thiogalactopyranoside (IPTG)-inducible lacUV5 promotor. In our plasmid, the coding region is under the regulatory control of the T7 RNA polymerase promotor. The plasmid also contains a gene for ampicillin resistance and an F1 origin of replication.

Mutagenesis. Site-directed mutagenesis was performed according to the method of Kunkel (1985). Single-stranded DNA (ssDNA) was obtained from pCApwt, harbored in *E. coli* CJ236 (ung^-, dut^-), after infection with helper phage M13K07; this procedure gives ssDNA containing uracil, since these host cells are deficient in the enzymes dUTPase and uracil *N*-glycosylase. After *in vitro* second-strand synthesis from oligonucleotides that contained mismatches, the plasmids were transfected into an ung^+,dut^+ strain of *E. coli*. All plasmids were prepared using the Magic MiniPreps kit from Promega (Madison, WI), and mutant plasmids were identified by sequencing the entire HCAII−cDNA according to a slightly modified version of the dideoxy chain-termination method (Sanger *et al.*, 1977).

Expression and Purification. Each strain was grown in 2× Luria−Bertani medium enriched with 0.5 mM $ZnSO_4$ and 50 mg/L ampicillin; culturing was done at 23 °C to prevent the formation of inclusion bodies, and when the cultures had reached $A_{660} = 0.5$, protein synthesis was induced by adding IPTG to a final concentration of 0.5 mM. The cultures were harvested after 16 h, and the cells were lysed by French pressing. The crude cell extracts were loaded onto an affinity chromatography column and purified in a single step (Khalifah *et al.*, 1977). The size and purity of the proteins were verified by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE). Since chromophores had been removed from the truncated and mutated variants, new extinction coefficients had to be determined. This was done by measuring the absorbances at 280 nm of one sample containing protein and 6 M GuHCl and of another sample with identical protein concentration but 0 M GuHCl. The coefficients were calculated according to Gill and von Hippel (1989): Trunc5, 47 100 M^{-1} cm^{-1}; Trunc17, 44 100 M^{-1} cm^{-1}; Trunc24, 41 700 M^{-1} cm^{-1}; W5F, 50 500 M^{-1} cm^{-1}; W16F, 49 000 M^{-1} cm^{-1}; and W5F/W16F, 43 800 M^{-1} cm^{-1}. $HCAII_{pwt}$ was assumed to have the same extinction coefficient as the wild type, *i.e.*, 54 800 M^{-1} cm^{-1} (Nyman & Lindskog, 1964).

Stability Measurements. The stabilities of the enzyme forms (0.025 mg/mL) were determined by incubation for 24 h at 23 °C in various concentrations of GuHCl buffered with 0.1 M Tris−H_2SO_4 (pH 7.5). The transitions were followed by assaying the CO_2 hydration activity and by monitoring the change in intrinsic fluorescence emission properties.

Kinetic Measurements. To study the refolding kinetics, the enzymes (0.625 mg/mL) were first denatured by incubation in 5 M GuHCl and 0.1 M Tris−H_2SO_4 (pH 7.5) at 23 °C for 1 h. Renaturation was achieved by diluting the denatured enzymes 25-fold into buffer, giving a GuHCl concentration of 0.2 M and a protein concentration of 0.025 mg/mL. The renaturation reaction was monitored by measuring the recovery of CO_2 hydration activity and the change in intrinsic fluorescence emission properties.

In another set of experiments, the diluting buffer contained 10 μM 5-(dimethylamino)naphthalene-1-sulfonamide (DNSA). This molecule binds specifically to the active site of HCAII, and upon doing so, its fluorescence properties change dramatically (Chen & Kernohan, 1967; Henkens *et al.*, 1982). Thus, this method reveals the rate at which the active-site cleft becomes fully organized.

CO_2 Hydration Activity Assay. Aliquots were withdrawn from the stability measurement samples and from samples of proteins at refolding conditions and then transferred to a CO_2 hydration assay solution containing 2.0 mL of 25 mM veronal−H_2SO_4 (pH 8.2), 16 mM bromothymol blue, and 0.5 mM ethylenediaminetetraacetic acid (EDTA). Water was added to give a volume of 3.0 mL, and the reaction was initiated by adding 2.0 mL of CO_2-saturated water to a final volume of 5.0 mL. The temperature of the assay solution was held at 0 °C. Under these conditions, no reactivation of the enzyme occurs in the assay medium; hence, a control sample that was 50% denatured showed no change in activity after incubation in the assay solution for up to 20 min prior to the addition of CO_2. This made it possible to collect samples for assay from the early phase of the denaturation and reactivation processes. The CO_2 hydration assay and the calculations were performed according to Rickli *et al.* (1964).

Fluorescence Measurements. All fluorescence data were collected on a Shimadzu RF-5000 spectrofluorophotometer at 23 °C in a 10 × 4 mm cuvette. All samples were excited at 295 nm when measuring intrinsic fluorescence; tyrosines and phenylalanines have negligible absorbances at this wavelength, and the emitted light therefore will originate from the tryptophans only. When the stability was measured, the bandwidths for excitation and emission were set to 3 and 10 nm, respectively, and emission spectra were collected between 310 and 370 nm. All spectra were corrected by subtracting the spectrum of the blank; the blank lacked enzyme but otherwise was identical to the sample. The emission maximum shifted from ~332 to ~352 nm upon denaturation, which indicates a shift of the tryptophans to a more polar environment. Measurement of the change in fluorescence intensity gives noisy denaturation curves due to differences in protein concentrations between samples, caused by pipeting errors and adsorption of protein to the walls of the cuvette. Although these unwanted variations in fluorescence intensity were very small compared to the total intensity, they still gave rise to unsmooth curves. The denaturation was described by plotting the ratio of the

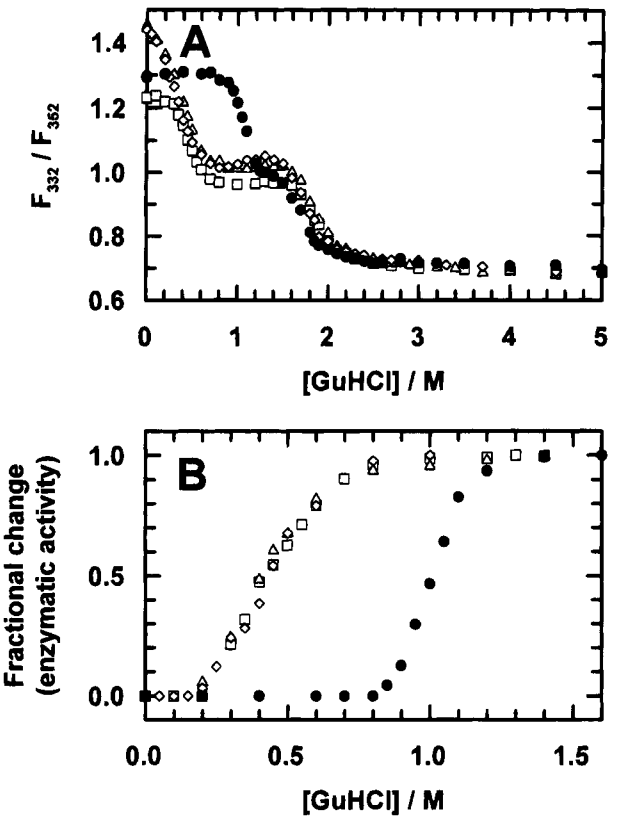

FIGURE 2: GuHCl denaturation of $HCAII_{pwt}$ and the truncated variants. The protein concentration was 0.025 mg/mL in 0.1 M Tris−H_2SO_4 (pH 7.5), and the incubation time was 24 h and the temperature 23 °C. The ratio of the fluorescence intensity at two wavelengths, F_{332}/F_{352} (A), and the fractional change in enzymatic activity (B) are plotted as a function of GuHCl concentration. Symbols: (●) $HCAII_{pwt}$; (□) Trunc5; (△) Trunc17; (◇) Trunc24.

intensities at two wavelengths, F_{332}/F_{352}, versus the GuHCl concentration (Banik *et al.*, 1992). This procedure takes care of the problem with varying concentration. However, the change in fluorescence intensity at 352 nm is, for some mutants, up to 40% of the intensity change at 332 nm, indicating that the ratio of these two wavelengths is not necessarily linearly correlated to the change in concentration. On the other hand, a comparison of denaturation curves produced by this method and curves obtained by measuring differences in absorbance at 292 nm, the most common method used, shows similar denaturation curves for a large set of HCAII mutants (data not shown). Moreover, a comparison of the first transition measured by the F_{332}/F_{352} method with the transition obtained by measuring enzymatic activity shows very similar patterns (Figure 2) and C_m values (Tables 2 and 3). The rationale for using fluorescence, as opposed to absorbance, is that fluorescence gives much better signal-to-noise ratios and that 10-fold lower protein concentrations can be used, minimizing interference from protein−protein interactions.

When intrinsic fluorescence was used to measure refolding kinetics, the samples were excited for only 2 s/min, and the emission was measured at 332 nm. The bandwidths for excitation and emission were set to 1.5 and 30 nm, respectively. The parameters used were chosen in order to minimize the bleaching of the sample: under these conditions, a native sample lost less than 1% of its fluorescence intensity over 3 h.

Table 2: Stability Parameters for GuHCl-Induced Unfolding Measured as the Change in Fluorescence[a]

protein	C_{mNI}[b]	C_{mIU}[b]	$\Delta G_{NI}^{H_2O}$	$\Delta G_{IU}^{H_2O}$	m_{NI}	m_{IU}	$\Delta\Delta G_{NI}$[c]	$\Delta\Delta G_{IU}$[d]
$HCAII_{pwt}$	1.1	1.7	9.3	7.7	8.7	4.5		
Trunc5	0.45	1.9	3.6	7.8	7.8	4.1	5.1	−0.8
Trunc17	0.37	1.9	1.8	7.7	4.9	4.0	4.7	−0.8
Trunc24	0.31	1.8	1.7	8.5	5.5	4.7	5.2	−0.4
W5F	0.77	1.7	8.1	7.7	10.5	4.4	2.5	−0.1
W16F	0.40	1.8	3.9	8.1	9.8	4.5	6.3	−0.3
W5F/W16F	0.32	1.8	1.6	8.6	5.1	4.7	5.0	−0.5

[a] Units are as follows: ΔG, kcal mol^{-1}; m, kcal mol^{-1} M^{-1}. [b] C_m represents the transition midpoint concentration of GuHCl. [c] At 0.75 M GuHCl. [d] At 1.8 M GuHCl.

Table 3: Stability Parameters for GuHCl-Induced Unfolding Measured as Enzymatic Activity[a]

protein	C_{mNI}[b]	$\Delta G_{NI}^{H_2O}$	m_{NI}	$\Delta\Delta G_{NI}$[c]
$HCAII_{pwt}$	1.0	9.9	9.8	
Trunc5	0.42	2.1	5.1	4.2
Trunc17	0.40	2.2	5.6	4.5
Trunc24	0.43	2.4	5.6	4.3
W5F	0.71	6.5	9.1	2.8
W16F	0.38	2.0	5.4	4.5
W5F/W16F	0.39	2.0	5.2	4.4

[a] Units are as follows: ΔG, kcal mol^{-1}; m, kcal mol^{-1} K^{-1}. [b] C_m represents the transition midpoint concentration of GuHCl. [c] At 0.75 M GuHCl.

When the refolding was monitored by measuring the binding of DNSA to the active site, the samples were excited at 320 nm and the emission was measured at 445 nm. The bandwidths were set to 10 and 5 nm, respectively.

Data Analysis. All data were fit with nonlinear least-squares analysis, using the computer program GraFit (Erithacus Software Ltd., Staines, UK). The stability data obtained from the fluorescence spectroscopy measurements indicated a three-state process and were fit to the equation

$$Y = ((N + IA)/(1 + A) + (UB))/(1 + B) \quad (1)$$

$$A = \exp\{(m_{NI}[\text{GuHCl}] - \Delta G_{NI}^{H_2O})/RT\}$$

$$B = \exp\{(m_{IU}[\text{GuHCl}] - \Delta G_{IU}^{H_2O})/RT\}$$

where Y is the observed ratio of fluorescence intensities, F_{332}/F_{352}, and N, I, and U are, respectively, the ratios for the native, intermediate and, unfolded forms (*i.e.*, $(F_{332}/F_{352})_N$, $(F_{332}/F_{352})_I$, and $(F_{332}/F_{352})_U$, respectively). Equation 1 was derived from the following equations:

$$\Delta G_{NI} = -RT[\ln(N - Y)/(Y - I)] \quad (2)$$

$$\Delta G_{IU} = -RT[\ln(I - Y)/(Y - U)] \quad (3)$$

$$\Delta G_{NI} = \Delta G_{NI}^{H_2O} - m_{NI}[\text{GuHCl}] \quad (4)$$

$$\Delta G_{IU} = \Delta G_{IU}^{H_2O} - m_{IU}[\text{GuHCl}] \quad (5)$$

When CO_2 hydration activity was used to measure stability, only one transition appeared. The data were fitted to the equation

$$Y = (N + IA)/(1 + A) \quad (6)$$

$$A = \exp\{(m_{NI}[\text{GuHCl}] - \Delta G_{NI}^{H_2O})/RT\}$$

This equation was derived from eqs 2 and 4, but in this case, Y represents the observed activity. N and I are the activities of the native and the intermediate states, respectively. The intermediate states always exhibited negligible activity.

In principle, it is preferable to compare the stabilities of the mutants and the reference in the absence of denaturant. In practice, however, data can be accurately measured only in the transition zones, and this leads to long extrapolations that might introduce large errors (Pace & Vanderburg, 1979; Mayr & Schmid, 1993). An alternative approach, which is strictly empirical, is to compare the free energies where they can be most accurately measured, *i.e.*, in or close to the transition regions (Matthews, 1987; Cupo & Pace, 1983; Kellis *et al.*, 1989). Therefore, we have chosen to calculate the differences in stability ($\Delta\Delta G$) at 0.75 and 1.8 M GuHCl for the first and second transitions, respectively. These concentrations were chosen from the middle of the overlapping data in the transition zones. The destabilization of the different mutants (MUT), as compared to that of $HCAII_{pwt}$ (REF), was determined by using the following equations:

at 0.75 M GuHCl for the first transition

$$\Delta\Delta G_{NI} = \Delta G_{NI}^{REF} - \Delta G_{NI}^{MUT} \quad (7)$$

at 1.8 M GuHCl for the second transition

$$\Delta\Delta G_{IU} = \Delta G_{IU}^{REF} - \Delta G_{IU}^{MUT} \quad (8)$$

Circular Dichroism Measurements. Circular dichroism spectra were recorded on a CD6 spectrodichrograph (Jobin-Yvon Instruments SA, Longjumeau, France). Constant N_2 flushing was employed, and the instrument was calibrated with an aqueous solution of *d*-10-(+)-camphorsulfonic acid at 290 nm. The CD spectra represent the averages of three scans obtained by collecting data at 0.5 nm intervals with an integration time of 2 s. Molar ellipticity is reported as mean residue molar ellipticity ($[\theta]$, deg·cm^2 dmol^{-1}) and was calculated by using the following equation:

$$[\theta] = [\theta]_{obs}(\text{mrw})/10cl \quad (9)$$

where $[\theta]_{obs}$ is the ellipticity measured in degrees, mrw is the mean residue molecular weight, c is the protein concentration in grams per milliliter, and l is the optical path length of the cell in centimeters. Far-UV CD spectra were recorded in a 0.5 mm quartz cell, and near-UV CD spectra were recorded in a 5 mm quartz cell. All proteins were buffered in 1 mM sodium phosphate (pH 7.5) at 23 °C, and the protein concentrations were 17 μM in both near- and far-UV measurements.

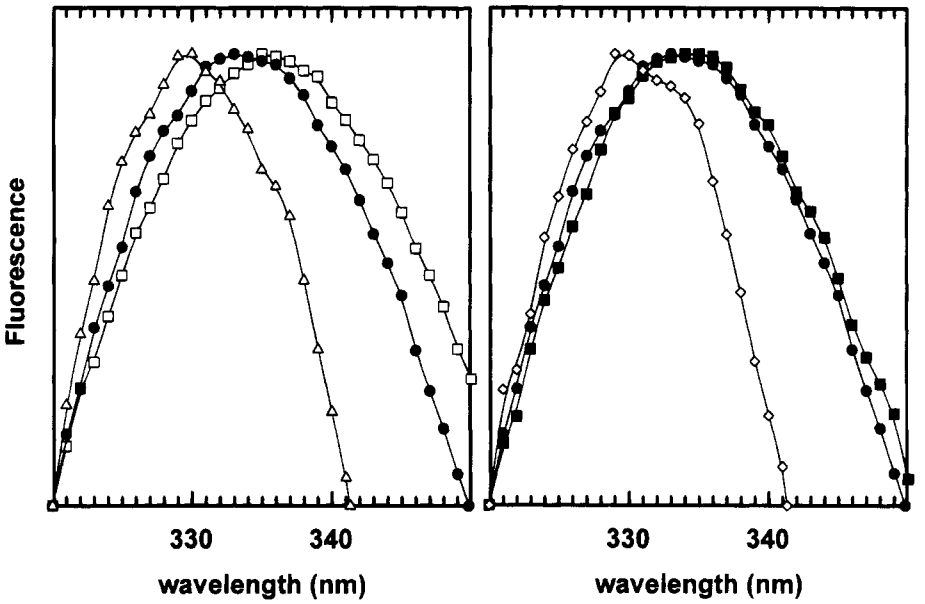

FIGURE 3: Native fluorescence spectra of $HCAII_{pwt}$, W5F, and the truncated variants (divided into two panels for clarity). The intensities of the spectra were set equal for an easier comparison of the λ_{max}. Symbols: (●) $HCAII_{pwt}$; (□) Trunc5; (△) Trunc17; (◇) Trunc24; (■) W5F.

Amino Acid Sequencing. N-Terminal amino acid sequences were determined in an Applied Biosystems Model 477A sequencing system.

Structure Analysis. Analysis of the protein structure was done on a computer graphics system using the program Insight II (Biosym Technologies, San Diego, CA). The coordinates, from X-ray crystallography data, were kindly provided by Håkansson *et al.* (1992).

RESULTS

Initial Characterization of the Truncated Variants. As a precautionary measure, we had the N-terminal parts of the purified, truncated HCAII variants determined by amino acid sequencing (Table 1). In all variants, except Trunc17, the N-terminal methionine was removed. As in these cases, methionyl aminopeptidase usually excises the methionine when the penultimate amino acid residue is small (Hirel *et al.*, 1989). Otherwise, the N-terminal sequences of the truncated variants conformed with the known primary structure of the enzyme (Henderson *et al.*, 1976).

The enzymatic activity of $HCAII_{pwt}$ is identical to the activity of the wild type, that is, with a k_{cat} of 10^6 s^{-1} (Lindskog & Liljas, 1993). The enzymatic activities of the truncated variants are not as different from that of $HCAII_{pwt}$ as one might have expected considering the large structural changes introduced. Trunc5 has roughly 30% activity compared to that of $HCAII_{pwt}$, whereas 15% of the activity of $HCAII_{pwt}$ remains in Trunc17 and Trunc24. This corresponds to a k_{cat} of at least 10^5 s^{-1}, which indicates that the truncated variants still are very efficient enzymes. It should be noted that these relatively small decreases in the enzymatic activities of Trunc5, Trunc17, and Trunc24 correspond to an increase in the free energy of activation of approximately 1 kcal/mol, which is an energy of the same magnitude as that caused by, for example, a hydrogen bond.

Trunc28 and Trunc31 were expressed in large quantities, as evidenced by SDS−PAGE, but showed no detectable activity and did not bind to the affinity gel. For these reasons, Trunc28 and Trunc31 were not purified or studied further.

When the intrinsic fluorescence of $HCAII_{pwt}$ is measured in various concentrations of GuHCl, the protein unfolds in two separate transitions (Figure 2A), and the enzymatic activity vanishes completely during the first transition (Figure 2B). The intermediate state, which is highly populated at a concentration of about 1.3 M GuHCl, has been characterized in some detail (Mårtensson *et al.*, 1993, 1994).

Deletion of 5, 17, or 24 amino acid residues from the N-terminus of $HCAII_{pwt}$ results in enzyme variants that are destabilized by 4−5 kcal/mol relative to the intermediate state (Figure 2, Tables 2 and 3). As is evident from the second transition at 1.7 M GuHCl, the stability of the intermediate state, relative to the unfolded state, is not changed significantly by these deletions.

The emission maxima of the native fluorescence spectra of Trunc5, Trunc17, Trunc24, and W5F are more or less shifted relative to $HCAII_{pwt}$ (Figure 3). Trunc5 is red-shifted by 2 nm relative to $HCAII_{pwt}$, while the other truncated variants are blue-shifted by 3 nm. These shifts are significant; as a comparison, the variations in several spectra of $HCAII_{pwt}$ never exceeded ±0.5 nm. The spectrum of W5F does not differ significantly from that of $HCAII_{pwt}$.

Circular Dichroism Spectra. CD spectra in the far-UV region (Figure 4A) allow an analysis of the secondary-structure content. Calculations performed according to the matrix-multiplication method (Compton & Johnson, 1986) give helical contents of 9%, 5%, 4%, and 4% for $HCAII_{pwt}$, Trunc5, Trunc17, and Trunc24, respectively. For the mutants W5F, W16F, and W5F/W16F, the result is 7% helix in all three cases (Freskgård *et al.*, 1994). The relative fluctuations in the content of the other secondary-structure elements, for all these variants, were much smaller than for those for helix.

The CD spectrum of Trunc5 in the near-UV region (Figure 4B) strongly resembles the spectrum of W5F (Freskgård *et al.*, 1994), and the spectra of Trunc17 and Trunc24 resemble that of W5F/W16F (G. Aronsson, unpublished results).

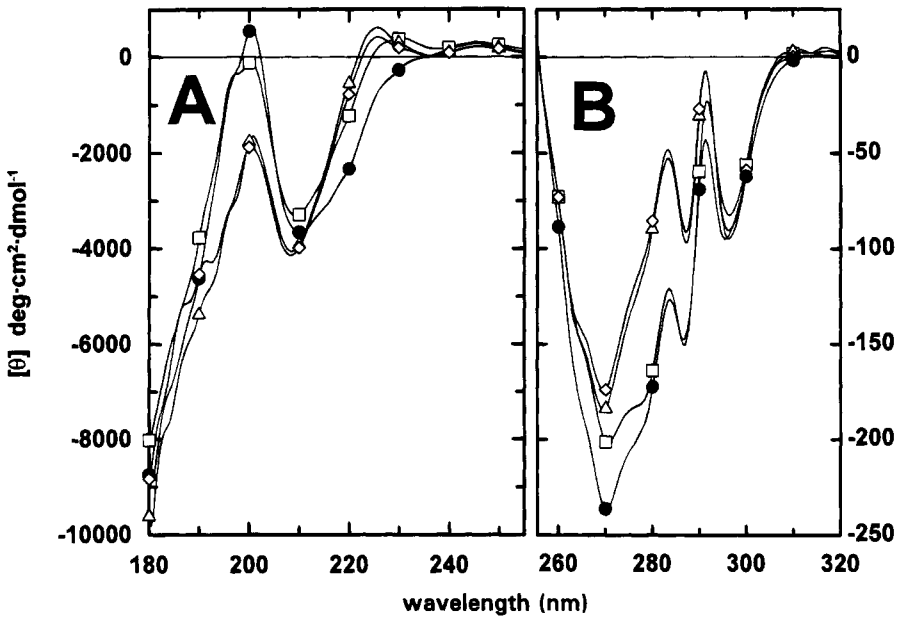

FIGURE 4: Far-UV (A) and near-UV (B) CD spectra of $HCAII_{pwt}$ and the truncated variants. The protein concentration was 17 μM in 1 mM sodium phosphate (pH 7.5), and the far- and near-UV spectra were recorded in 0.5 and 5 mm quartz cells, respectively. The symbols represent a few randomly chosen data points that are included so that the curves can be distinguished. Symbols: (●) $HCAII_{pwt}$; (□) Trunc5; (△) Trunc17; (◇) Trunc24.

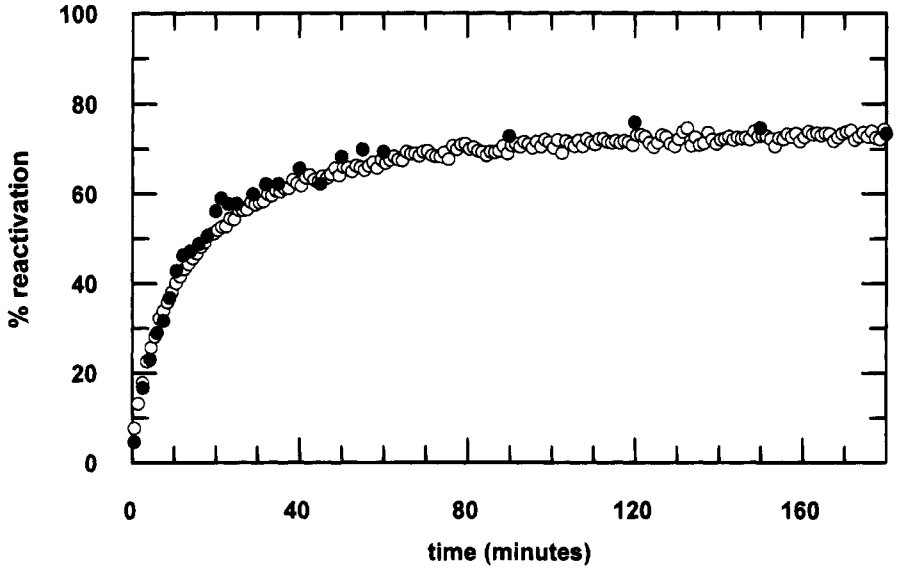

FIGURE 5: Reactivation of $HCAII_{pwt}$ monitored by two different methods: regain in enzymatic activity (●) and the binding of DNSA to the active site upon renaturation (○). $HCAII_{pwt}$ was incubated in 5 M GuHCl and 0.1 M Tris−H_2SO_4 (pH 7.5) at 23 °C for 1 h. Renaturation was achieved by rapid dilution to 0.2 M GuHCl and a protein concentration of 0.025 mg/mL.

Refolding Monitored by Various Methods. Refolding of $HCAII_{pwt}$ measured either by the regain in enzymatic activity or by the binding of DNSA to the active site gives identical results (Figure 5). The DNSA-refolding trace is biphasic with respective rate constants of 0.22 and 0.034 min^{-1} and amplitudes of 32% and 41%, giving a total yield of 73%; the remaining 27% never reactivates. The DNSA-refolding traces of the three active truncated variants and the three tryptophan mutants are very similar to that of $HCAII_{pwt}$ (Table 4). All of these traces fit well to double-exponential curves with rate constants that fall within the range 0.22−0.33 min^{-1} for the first phase and 0.034−0.044 min^{-1} for the second. The amplitudes are also similar to those of $HCAII_{pwt}$. The accuracy of the enzymatic activity assay is not as good as for the DNSA-binding assay, but these data are also best described by two-exponential phases. The reactivation kinetics produce essentially the same kinetic parameters as do the DNSA-refolding experiments, and all reactivation yields fall within the range 67−73% (data not shown).

When the renaturation was measured with intrinsic fluorescence, however, a more complex pattern arose (Figure 6, Table 4). W5F has a large increase in total amplitude, and Trunc5 has approximately the same total amplitude as $HCAII_{pwt}$. The total amplitudes of the remaining variants are, more or less, decreased when compared to $HCAII_{pwt}$. Despite these large differences in total amplitude, $HCAII_{pwt}$, W5F, W16F, and Trunc5 all have a biphasic increase in fluorescence intensity: the rate constants fall within the ranges of 0.09−0.16 min^{-1} for the first phase and 0.028−0.043 min^{-1} for the second. For Trunc17 and W5F/W16F the signal-to-noise ratio is not high enough to determine the

Table 4: Kinetic Parameters from Refolding Experiments Employing Two Different Methods[a]

	DNSA binding				intrinsic fluorescence			
protein	k_1[b]	k_2[b]	Amp_1[c]	Amp_2[c]	k_1[b]	k_2[b]	Amp_1[d]	Amp_2[d]
$HCAII_{pwt}$	0.22 ± 0.01	0.034 ± 0.001	32 ± 1	41 ± 1	0.15 ± 0.06	0.043 ± 0.004	32 ± 10	68 ± 10
Trunc5	0.23 ± 0.01	0.034 ± 0.001	30 ± 1	44 ± 1	0.11 ± 0.01	0.028 ± 0.005	43 ± 9	85 ± 9
Trunc17	0.33 ± 0.03	0.044 ± 0.001	28 ± 1	43 ± 1				
Trunc24	0.27 ± 0.02	0.040 ± 0.001	35 ± 1	38 ± 1				
W5F	0.24 ± 0.01	0.034 ± 0.001	30 ± 1	43 ± 1	0.13 ± 0.002	0.039 ± 0.001	64 ± 10	177 ± 10
W16F	0.26 ± 0.02	0.036 ± 0.001	33 ± 1	40 ± 1	0.10 ± 0.03	0.034 ± 0.008	26 ± 21	34 ± 18
W5F/W16F	0.30 ± 0.02	0.042 ± 0.001	26 ± 1	44 ± 1				

[a] The numbers are reported with 95% confidence intervals. [b] The units of all rate constants are minutes^{-1}. [c] These amplitudes are reported as the percent bound DNSA compared to a native reference. [d] These amplitudes are normalized to the total amplitude of $HCAII_{pwt}$.

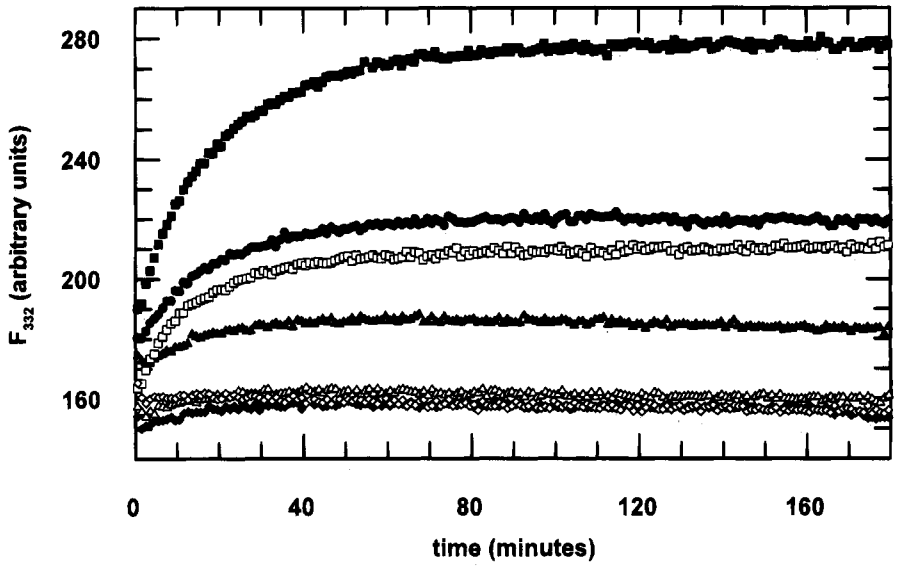

FIGURE 6: Renaturation measured as intrinsic fluorescence. $HCAII_{pwt}$, the truncated variants, and the tryptophan mutants were incubated in 5 M GuHCl and 0.1 M Tris−H_2SO_4 (pH 7.5) at 23 °C for 1 h. Renaturation was achieved by dilution to 0.2 M GuHCl and a protein concentration of 0.025 mg/mL. The fluorescence intensity at 332 nm, F_{332}, is plotted as a function of time. Symbols: (●) $HCAII_{pwt}$; (□) Trunc5; (△) Trunc17; (◇) Trunc24; (■) W5F; (▲) W16F; (♦) W5F/W16F.

kinetics of refolding with any resonable accuracy; however, the total amplitudes of these two variants are roughly 20% of the total amplitude of $HCAII_{pwt}$. Trunc24 does not exhibit any measurable fluorescence change between 30 s and 3 h of refolding.

DISCUSSION

Unfolding Probed by Intrinsic Fluorescence. The fluorescence emission spectrum of HCAII changes as the GuHCl concentration increases: the peak decreases in intensity and undergoes a red shift. This reflects the increasing exposure of aromatic side chains to a polar environment. Excitation of samples at 295 nm reflects environmental changes of tryptophans, since other amino acid residues have negligible absorbances at this wavelength. HCAII has seven tryptophans positioned in various parts of the three-dimensional structure, and the change in fluorescence therefore will monitor global conformational changes. However, all of the tryptophans do not contribute equally to the fluorescence spectrum, at least not in the native state or in native-like species (Mårtensson *et al.*, 1995).

Destabilization Caused by Truncations of the N-Terminus. HCAII has an intermediate state that is highly populated at a concentration of approximately 1.3 M GuHCl (Mårtensson *et al.*, 1993). Deletion of the first five amino acid residues from the N-terminus of $HCAII_{pwt}$ results in only moderate losses of stability, considering the large structural changes introduced (Figure 2, Tables 2 and 3), and deletion of 17 amino acid residues does not cause additional destabilization, relative to the intermediate state. Not even when 24 amino acid residues are cut off, thereby removing an entire hydrophobic cluster and almost one-tenth of the molecule, is there any further destabilization, which indicates that the only amino acid residues in the N-terminal 24-amino acid segment that interact favorably with the rest of the molecule are one or more of the first five amino acid residues. These interactions are not present in the intermediate state of $HCAII_{pwt}$; truncation of the N-terminus only marginally affects the stability of the intermediate states relative to the unfolded states. Hence, in the intermediate state, there are no net stabilizing tertiary contacts between the N-terminus and the rest of the molecule.

Truncation of 28 amino acid residues or more, however, results in a protein without any enzymatic activity, and the native structure is probably ruptured. This is not surprising, since it means perturbation of a highly conserved region which is very sensitive to mutation (Fransson *et al.*, 1992; Mårtensson *et al.*, 1992).

Changes in the Secondary and Tertiary Interactions. Secondary-structure predictions based on the CD spectra in the far-UV region (Figure 4A) show that some helical structure has disappeared in the truncated enzymes. The

tryptophans of HCAII interfere with the far-UV CD spectrum (Freskgård *et al.*, 1994), but this interference is not the reason for the observed decrease in helical structure upon truncation: when the same secondary-structure prediction algorithm was applied to the spectra of W5F, W16F, and W5F/W16F, a large decrease in helical structure, as for the truncated variants, was not observed. Thus, these results indicate that the deletion of five amino acid residues is sufficient to induce unfolding of the helical segments comprising amino acid residues 13−18 and 21−24.

The helical content of HCAII deduced from X-ray crystallography data (Eriksson *et al.*, 1988) is 16%. The discrepancy between this value and our calculated value of 9% probably reflects the contribution of tryptophan side chains to the far-UV CD spectrum. Indeed, HCAII has more tryptophans than does the average protein, on which the prediction method used is based. This does not invalidate our predictions of helical content, since we use the result only to observe relative changes.

CD spectra in the near-UV region (Figure 4B) show that the tertiary fold is not perturbed by N-terminal deletions. The differences between these spectra could be accounted for solely by the removal of Trp5 and/or Trp16 (Freskgård *et al.*, 1994). Although the environment of Trp16 in Trunc5 is different from the one it had in $HCAII_{pwt}$, it is still asymmetrical; if this tryptophan were in a symmetrical environment in Trunc5, as in a random coil, the near-UV CD spectrum of Trunc5 should overlap the spectra of Trunc17 and Trunc24.

Exposure of Tryptophans. Trp5 and Trp16 display interdependent fluorescence. These tryptophans are located very close to one another, and energy is transferred from Trp16 to Trp5, which, in turn, is quenched by the neighboring histidines (Mårtensson *et al.*, 1995). Trunc5 and W5F lack Trp5, and therefore, the emission from Trp16 can no longer be transferred to Trp5, which will allow radiative relaxation of the excited Trp16. Trunc5 and W5F have the same tryptophan content, yet their native fluorescence spectra differ (Figure 3): Trunc5 is red-shifted by 2 nm relative to W5F. This shows that Trp16 becomes more exposed to the polar environment in Trunc5 than it is in W5F or $HCAII_{pwt}$. This is in accordance with the CD measurements, which also indicated that N-terminal structures had already unfolded in the Trunc5 variant. The fluorescence spectra of Trunc17 and Trunc24, on the other hand, are blue-shifted by 3 nm relative to the spectrum of $HCAII_{pwt}$, which is what should be expected when removing an exposed tryptophan such as Trp16. The observed emission maxima of each variant (Figure 3) parallel the F_{332}/F_{352} ratio at 0 M GuHCl (Figure 2A), *i.e.*, the F_{332}/F_{352} ratio of Trunc5 is much lower than that of $HCAII_{pwt}$ (red shift) and Trunc17 and Trunc24 have higher ratios (blue shift).

Another observation concerning these two N-terminal tryptophans is the destabilization they cause when replaced by phenylalanines, *i.e.*, the $\Delta\Delta G_{NI}$ values (Tables 2 and 3). The sum of the destabilizations that the two single mutants W5F and W16F give rise to is far greater than the destabilization caused by the double mutant W5F/W16F alone. This synergism shows that there should be positive interactions between these two positions in the N-terminal aromatic cluster. Indeed, these tryptophans are in an edge-to-face position, which has been proposed to be the most favorable aromatic−aromatic interaction (Burley & Petsko, 1985).

Refolding of the Truncated Variants and the Tryptophan Mutants. The extrinsic probe DNSA binds with high affinity to the active-site crevice of HCAII (Chen & Kernohan, 1967), and the reappearance of enzymatic activity that occurs when $HCAII_{pwt}$ is refolded from the denatured state parallels the amount of bound DNSA (Figure 5, Table 4). Removal of more or less of the amino acid segment 1−24 does not have any effect at all on the reactivation kinetics. Normally the denaturation time before renaturation was 1 h, but when $HCAII_{pwt}$ was refolded after a denaturation time of only 10 s, the amplitudes of the phases changed drastically while the rate constants were essentially unchanged (data not shown). Thus, the biphasic reappearance of the activity is probably due to the presence of two or more populations of proline *cis*−*trans* isomers.

When the refolding of $HCAII_{pwt}$ and its variants, from the denatured state, was measured by intrinsic fluorescence, the first data points were recorded after 30 s. At that time point, the fluorescence intensity at 332 nm had already reached a level approximately halfway between the fluorescence levels of the denatured and native states of $HCAII_{pwt}$ (data not shown); the rest of the fluorescence intensity change, however, occurred between 30 s and 3 h (Figure 6). When the N-terminal 24-amino acid segment is removed, the fluorescence intensity change after 30 s is gone, while the fluorescence intensity change before 30 s remains essentially unchanged. This clearly shows that the two tryptophans residing in the N-terminus, Trp5 and Trp16, are the dominating source of fluorescence change after 30 s and that these two tryptophans do not contribute significantly to the fluorescence change before 30 s. The other five tryptophans, *i.e.*, numbers 97, 123, 192, 209, and 245, obviously have acquired their native fluorescence properties within 30 s of renaturation. In the case of Trp123, Trp192, and Trp209, this does not necessarily mean that these positions are not involved in minor structural rearrangements after this time, since these tryptophans have been shown to contribute only a little to the native fluorescence spectrum (Mårtensson *et al.*, 1994). The rest of this section will deal only with the events taking place after 30 s of refolding.

When the tryptophan in position 5 is replaced by a phenylalanine, there is a large increase in fluorescence change upon renaturation, as compared to $HCAII_{pwt}$ (Figure 6, Table 4). Again, this can be explained by energy transfer from Trp16 to Trp5 and by the fact that Trp5 is quenched by neighboring histidines (Mårtensson *et al.*, 1995), *i.e.*, removal of Trp5 will allow the radiation from Trp16 to reach the detector rather than becoming trapped by Trp5. Trunc5 does not exhibit the same large total amplitude as W5F, even though Trp5 is removed from this variant as well. The reason for this must be that Trp16 never assumes the wild-type position in Trunc5 and, therefore, emits less radiation than it does in its normal position in $HCAII_{pwt}$. As we have seen, the CD and stability measurements also point in this direction, *i.e.*, deletion of the first five amino acid residues of the polypeptide chain of $HCAII_{pwt}$ unfolds a larger part of the N-terminus and leaves Trp16 fairly exposed.

When Trp16 is replaced by a phenylalanine, the fluorescence change upon renaturation is reduced (Figure 6); the total amplitude that remains within the experimental time

window is approximately 50% of the total amplitude of $HCAII_{pwt}$ (and approximately 25% of that of W5F). When Trp5 and Trp16 are removed simultaneously, either by mutating them to phenylalanines or by truncating the first 17 amino acid residues in the sequence, the total amplitude is lower yet, *i.e.,* roughly 20% of the total amplitude of $HCAII_{pwt}$. Thus, it seems obvious that the large change in fluorescence upon renaturation of $HCAII_{pwt}$ arises mainly because of the environmental changes that Trp16 undergoes. Part of the observed fluorescence change might also originate from Trp5, since the total amplitude of W16F is larger than that of W5F/W16F.

The small amplitude remaining in W5F/W16F and Trunc17 is completely gone in Trunc24; there is no increase in the fluorescence intensity at all between 30 s and 3 h of refolding. The reason for this difference might be that amino acid segment 18−24, which is absent from Trunc24 but present in the other variants, influences the fluorescence properties of some of the five remaining tryptophans.

The fluorescence reports on the folding of the N-terminus, and the binding of DNSA reports on the formation of the active site. Both events occur in a biphasic manner, and both methods exhibit similar amplitudes (Table 4). The formation of the active site is not affected at all when smaller or larger portions of the N-terminus are removed, *i.e.,* the reactivation kinetics of $HCAII_{pwt}$ and all its variants are identical, showing that the formation of the active site occurs totally independent of the N-terminus. The rates at which the N-terminus forms, on the other hand, are somewhat slower than the regain in enzymatic activity; k_1 is a factor 2 lower when measured by fluorescence and k_2 is identical. This shows that the N-terminal structure is formed only after the formation of an enzymatically active native-like structure of the rest of the molecule (amino acid segment 25−259). That structure possibly acts as a template for the proper formation of the N-terminus.

Concluding Remarks. A rather detailed picture of the folding of HCAII is starting to emerge. The molecule does not fold in a cooperative manner: different parts become organized after different times of refolding. Previous results, together with the results from the present study, give the following picture of the order of events during the folding of human carbonic anhydrase II.

1. The naturally occurring Cys206, situated in β-strand 7, becomes buried early in folding (Carlsson *et al.*, 1975; Freskgård *et al.*, 1991). Notably, β-strands 6 and 7 constitute the most hydrophobic parts of the polypeptide (Bergenhem *et al.*, 1989).

2. Tryptophans 97, 123, 192, 209, and 245 acquire their native fluorescence properties within 30 s of renaturation.

3. A compact conformation is achieved later in the region around position 256, which is located in a peripheral part of the β-sheet in the C-terminus (Freskgård *et al.*, 1991). It has also been shown that the central parts of the β-sheet are buried at relatively high concentrations of denaturant and that positions become progressively more accessible the closer they are to the peripheral parts of the central β-sheet (Mårtensson *et al.*, 1993).

4. There is a biphasic reappearance of the enzymatic activity. The two phases have similar amplitudes, and their respective rate constants are approximately 0.3 and 0.04 min^{-1} (corresponding to half-times of 2 and 17 min), respectively.

5. The structure of the N-terminus is also formed in a biphasic manner, with rates of approximately 0.15 and 0.04 min^{-1} (corresponding to half-times of 4 and 17 min), respectively. The results strongly indicate that this process is dependent on the prior formation of an enzymatically active native-like structure of the rest of the protein. Partaking in this late process are, at a minimum, the amino acid residues in segment 1−17 and one or more amino acid residues in segment 18−24.

ACKNOWLEDGMENT

We thank Dr. Per-Ola Freskgård (Linköping University) for expertly performing the circular dichroism measurements. We also thank Dr. Nils Bergenhem (University of Michigan, Ann Arbor, MI) for critical reading of this manuscript.

REFERENCES

Banik, U., Saha, R., Mandal, N. C., Bhattacharyya, B., & Roy, S. (1992) *Eur. J. Biochem. 206*, 15−21.
Bergenhem, N., Carlsson, U., & Karlsson J.-Å. (1989) *Int. J. Pept. Protein Res. 33*, 140−145.
Burley, S. K., & Petsko, G. A. (1985) *Science 229*, 23−28.
Carlsson, U., Henderson, L. E., Nyman, P. O., & Samuelsson, T. (1974) *FEBS Lett. 48*, 167−171.
Carlsson, U., Aasa, R., Hendersson, L. E., Jonsson, B.-H., & Lindskog, S. (1975) *Eur. J. Biochem. 52*, 25−36.
Chen, R. F., & Kernohan, J. C. (1967) *J. Biol. Chem. 242*, 5813−5823.
Compton, L. A., & Johnson, W. C., Jr. (1986) *Anal. Biochem. 155*, 155−167.
Cupo, J. F., & Pace, C. N. (1983) *Biochemistry 22*, 2654−2658.
Eriksson, A. E., Jones, T. A., & Liljas, A. (1988) *Proteins: Struct., Funct., Genet. 4*, 274−282.
Fransson, C., Freskgård, P.-O., Herbertsson, H., Johansson, A., Jonasson, P., Mårtensson, L.-G., Svensson, M., Jonsson, B.-H., & Carlsson, U. (1992) *FEBS Lett. 296*, 90−94.
Freskgård, P.-O., Carlsson, U., Mårtensson, L.-G., & Jonsson, B.-H. (1991) *FEBS Lett. 289*, 117−122.
Freskgård, P.-O., Carlsson, U., Mårtensson, L.-G., Jonasson, P., & Jonsson, B.-H. (1994) *Biochemistry 33*, 14281−14288.
Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
Håkansson, K., Carlsson, M., Svensson, L. A., & Liljas, A. (1992) *J. Mol. Biol. 227*, 1192−1204.
Henderson, L. E., Henriksson, D., & Nyman, P. O. (1976) *J. Biol. Chem. 251*, 5457−5463.
Henkens, R. W., Kitchell, B. B., Lottich, S. C., Stein, P. J., & Williams, T. J. (1982) *Biochemistry 21*, 5918−5923.
Hirel, P.-H., Schmitter, J.-M., Dessen, P., Fayat, G., & Blanquet, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8247−8251.
Janin, J., & Wodak, S. J. (1983) *Prog. Biophys. Mol. Biol. 42*, 21−78.
Kellis, J. T., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914−4922.
Khalifah, R. G., Strader, D. J., Bryant, S. H., & Gibson, S. M. (1977) *Biochemistry 16*, 2241−2247.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
Liljas, A., Håkansson, K., Jonsson, B.-H., & Xue, Y. (1994) *Eur. J. Biochem. 219*, 1−11.
Lindskog, S., & Liljas, A. (1993) *Curr. Opin. Struct. Biol. 3*, 915−920.
Mårtensson, L.-G., Jonsson, B.-H., Andersson, M., Kihlgren, A., Bergenhem, N., & Carlsson, U. (1992) *Biochim. Biophys. Acta 1118*, 179−186.
Mårtensson, L.-G., Jonsson, B.-H., Freskgård, P.-O., Kihlgren, A., Svensson, M., & Carlsson, U. (1993) *Biochemistry 32*, 224−231.
Mårtensson, L.-G., Jonasson, P., Freskgård, P.-O., Svensson, M., Carlsson, U., & Jonsson, B.-H. (1995) *Biochemistry* (in press).

Matthews, C. R. (1987) *Methods Enzymol. 154*, 498−511.
Mayr, L. M., & Schmid, F. X. (1993) *Biochemistry 32*, 7994−7998.
Nair, S. K., Calderone, T. L., Christiansson, D. W., & Fierke, C. A. (1991) *J. Biol. Chem. 266*, 17320−17325.
Nozaki, Y. (1972) *Methods Enzymol. 26*, 43−50.
Nyman, P. O., & Lindskog, S. (1964) *Biochim. Biophys. Acta 85*, 141−151.
Pace, C. N., & Vanderburg, K. E. (1979) *Biochemistry 18*, 288−292.
Rickli, E. E., Ghazanfar, S. A. S., Gibbons, B. H., & Edsall, J. T. (1964) *J. Biol. Chem. 239*, 1065−1078.
Sanger, S., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113−130.

BI941513J