Volume 289, number 1, 117–122 FEBS 10113 September 1991
© 1991 Federation of European Biochemical Societies 00145793/91/$3.50
ADONIS 001457939100829H

Folding around the C-terminus of human carbonic anhydrase II

Kinetic characterization by use of a chemically reactive SH-group introduced by protein engineering

Per-Ola Freskgård[1], Uno Carlsson[1], Lars-Göran Mårtensson[2] and Bengt-Harald Jonsson[2]

[1]*IFM-Department of Chemistry, Linköping University, S-581 83 Linköping, Sweden* and [2]*Department of Biochemistry, Umeå University, S-901 87 Umeå, Sweden*

Received 26 June 1991

We are characterizing the process of refolding of the enzyme human carbonic anhydrase II from the denatured state in guanidine hydrochloride. To describe the folding in defined parts of the protein we use protein engineering to introduce cysteine residues as unique chemically reactive probes. The accessibility of the cysteine SH-group to the alkylating reagent iodoacetate, at different stages during refolding, is used to give a kinetic description of the folding process. The structuration of the C-terminal part of the polypeptide chain, which is involved in a unique 'knot' topology, was investigated. Our results show that the structure around the C-terminal, composed of the outermost β-strands in a dominating β-structure that extends through the entire protein, is formed relatively late during refolding. In contrast, it was found that β-strands located in the interior of the protein were structured very rapidly. The final native structure is formed in a process that is slower than those observed for formation of β-structure.

Protein folding; Site directed mutagenesis; Side chain accessibility; Carbonic anhydrase; Protein engineering

1. INTRODUCTION

Detailed descriptions of the order of events during folding as well as characterization of folding intermediates are crucial for the understanding of protein folding processes. This study is the first in a series where we seek such information for the protein human carbonic anhydrase II, using a novel approach. The idea is to use site-directed mutagenesis to introduce cysteine residues as local, chemically reactive probes. They are introduced one at a time at numerous positions all over the protein structure giving a repertoire of protein variants each with a single residue. Based on the known three-dimensional structure we have chosen positions for mutagenesis that are found in different types of secondary structure, in aromatic clusters and in the C- and N-terminal parts of the protein. The positions were selected so the cysteine side chain should be buried in the properly folded protein. The accessibility of the SH-group to the alkylating reagent ICH_2COO^- during refolding of the protein from a denatured state is used to give a kinetic description of the folding process viewed from that particular substructure of the protein.

The specific reactivity of SH-groups towards alkylating reagents has earlier been used by us in studies of the refolding of human carbonic anhydrase I, in this case the SH-group of a naturally occurring cysteine residue [1,2].

The aim of this study is to describe the folding of the C-terminal portion of the polypeptide chain. Sequence position 256 was chosen primarily because this amino acid residue is located in a β-strand in a peripheral part of the 10-stranded β-sheet which is the dominating secondary structure of the protein (Fig. 1). Another reason is that position 256 is close to the point in the structure where the polypeptide chain is crossing itself in a way that would give a 'knotted' topology if the polypeptide chain were grasped at each end and stretched out; a feature that, to our knowledge, is unique to carbonic anhydrase [3,4].

It is important that the cysteine residue is introduced in a position which neither drastically alters the stability of the protein nor affects the folding pathway. The analysis of the experiments is simplified if the mutagenesis introduces a cavity in the folded protein that is large enough to accommodate the modified side chain after reaction with ICH_2COO^-.

Analysis of the protein structure on a computer graphics system using FRODO [5] in order to find a position fulfilling the described criteria led to the decision to mutagenize position 256 from isoleucine to cysteine. Initially, the naturally occurring cysteine residue at position 206 was replaced by serine (C206S), so that the introduction of a new cysteine residue at position 256 gives a protein with only one SH-group (I256C/C206S).

Abbreviation: HCAII, human carbonic anhydrase II.

Correspondence address: B.-H. Jonsson, Department of Biochemistry, Umeå University, S-901 87 Umeå, Sweden. Fax: (46) (90) 167661.

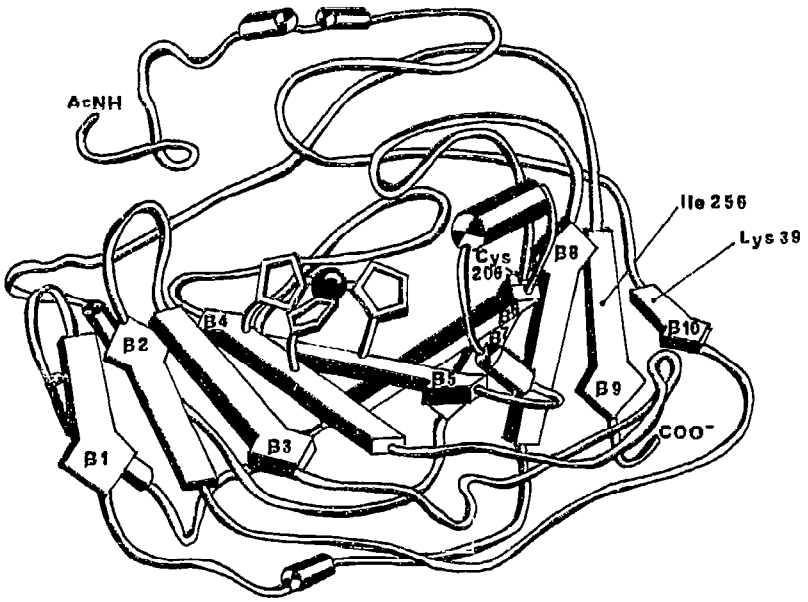

Fig. 1. Schematic drawing of the polypeptide backbone of carbonic anhydrase II with the positions of some amino acid residues indicated. (By courtesy of Prof. A. Liljas)

2. MATERIALS AND METHODS

2.1. *Chemicals*

Sequential grade guanidine hydrochloride (GuHCl) was obtained from Pierce. Metal ion impurities from GuHCl were removed by extraction with dithizone (7 mg·l^{-1}) in carbon tetrachloride. Concentrations of GuHCl were determined from refractive index measurements using an equation given by Nozaki [6]. Iodo[2-^{14}C]acetic acid (54 mCi/mmol) was purchased from Amersham. All other chemicals were reagent grade.

2.2. *Protein isolation and purification*

Human erythrocyte carbonic anhydrase II (HCAII) was initially purified by affinity chromatography according to Khalifah [7] and was further chromatographed on a DEAE-cellulose column using the conditions described by Henderson and Henriksson [8]. For in vitro site-directed mutagenesis we used the method based on host cell deficiency in the enzymes dUTPase (*dut*$^-$) and uracil *N*-glycosylase (*ung*$^-$) described by Kunkel [9]. For production of the enzyme variants we used the expression plasmid pHCAII described earlier [10]. The enzyme mutants were purified by affinity chromatography [7]. The purity of the enzymes used in these studies was verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis.

2.3. *Stability measurements*

Solutions of enzyme (8.5 mM) in 0.1 M Tris-H_2SO_4 (pH 7.5) with various concentrations of GUHCl were allowed to equilibrate for 24 h at 23°C. The transition was followed by assaying the CO_2 hydration activity.

2.4. *Kinetic measurements*

The enzyme (14.2 mM) was denatured in 5 M GuHCl, buffered with 0.1 M Tris-H_2SO_4 (pH 7.5) for 5 min at 23°C. Renaturation of the enzyme was achieved by rapid dilution of the denaturing agent to 0.3 M with 0.1 M Tris-H_2SO_4, pH 8.0. The final concentration of enzyme was 0.85 mM. Aliquots were withdrawn during the renaturation experiments and transferred to the enzyme activity assay solution which contained 2.0 ml of 25 mM Veronal-SO_4 buffer pH 8.2., Bromothymol blue (16 mM) and 0.5 mM EDTA. The enzymic reaction was initiated by the addition of 2 ml of saturated CO_2 solution. The final volume of the assay solution was 5 ml and the temperature was 0°C [11]. Under these conditions no reactivation or inactivation of the enzyme occurs in the assay medium. Thus, a control sample which was denaturated to 50% showed a stable activity after incubation in the assay buffer for 20 min prior to the addition of CO_2. This made it possible to collect samples for assay from the early phase of the renaturation reaction at intervals shorter than the duration of an assay of enzyme activity.

2.5. *Alkylation with iodoacetate to monitor renaturation*

Carboxymethylation was performed with a 145-fold molar excess of iodo[2-^{14}C]acetate (19.5 mCi) over enzyme (0.85 mM, 2.5 nmol in 3 ml). To follow the course of renaturation the iodoacetate was added to the renaturation mixture at different times (t_n; time interval 0–60 min) after the onset of renaturation. For measurement at time zero (t_0) the iodoacetate was premixed with the dilution solution. Following the addition of ICH_2COO^- at time t_n in the renaturation was allowed to proceed for 90 min. The reaction was stopped by applying the reaction mixture to an affinity chromatography column (CM-Bio-Gel A, Bio-Rad Laboratories, with *p*-aminomethylbenzenesulfonamide attached; gel bed dimensions: 5 × 0.5 cm). By this procedure the carboxymethylated enzyme samples were separated from excess reagent and inactive material (25%). The fractions containing enzyme activity were pooled and concentrated 10- to 12-fold using Centriprep-10 (Amicon). Final enzyme concentrations were determined spectrophotometrically at 280 nm assuming an $A_{280}^{1\%}$ of 18.7 cm^{-1}, and a molecular weight of 29 300. The incorporated radioactivity during the 90 min period from time t_n is x_{in}. To determine the accessibility of the SH-group at different times during refolding the difference in incorporated radioactivity between two successive time points was calculated. The difference $(x_{in}-x_{in+1})$ was taken as a measure of the accessibility at time $(t_n + t_{n+1})/2$. The quenching of accessibility at the SH-group was calculated using the formula $100\times (1-(x_{in}-x_{in+1})/(x_{i0}-x_{i1}))$. For this approach to be valid the renaturation must be complete within 90 min and the incorporation of radioactivity into renatured (native) protein must be negligibly small. Both these conditions are fulfilled. For example, the incorporation per time unit of radioactivity into renatured (native) protein is less than 1% of the incorporation into denatured protein for the 1256C/C206S mutant.

2.6. *Measurement of radioactivity*

The radioactivity of the enzyme samples (dissolved in 15 ml scintillation liquid, Liquid Scintillator Supersolve X, Zinsser Analytic) was determined in a liquid scintillation counter (Beckman LS 1801). The

radioactivity of a reference sample, the filtrate from the concentration step dissolved in the same volume of scintillation liquid was subtracted from the values obtained for the enzyme samples. These reference values were 100–110 cpm, while the background values were 90–100 cpm.

3. RESULTS AND DISCUSSION

3.1. *Accessibility of the SH-group*

We performed two experiments to quantify the accessibility of the SH-group of cysteine 256 to ICH_2COO^- in the initial denatured state and in the final native state (in 0.3 M GuHCl). Using a 145-fold molar excess of iodo[2-^{14}C]acetate over enzyme as in the kinetic experiments, we obtained a labeling of 240 cpm/nmol after 90 min incubation with refolded I256C/C206S mutant in 0.3 M GuHCl. The corresponding labeling of the denatured enzyme in 5 M GuHCl was 19800 cpm/nmol. Incubation of cloned, unmodified HCAII with iodoacetate in the same manner led to an incorporation of radioactivity at position 206 of 370 cpm/nmol in 0.3 M GuHCl and 9100 cpm/nmol in 5 M GuHCl. These differences in accessibility are large enough to make the kinetics studies feasible.

It has earlier been shown that pH 7.5 is optimal for efficient reactivation of HCAII after denaturation in 5 M GuHCl [12]. However, the alkylation of cysteine residues is more rapid at higher pH, since the pK_a values for the SH-groups in cysteine residues are usually close to 9. As a compromise the kinetic renaturation experiments were done at pH 8, while the preceding denaturation was done at pH 7.5 to preclude formation of disulfide bonds. In control experiments at pH 8 it was found that the carboxymethylated and unmodified forms of the I256C/C206S mutant were reactivated to the same degree after 1 h of refolding. HCAII from red blood cells and the I256C/C206S mutant were reactivated to 75% under those experimental conditions. Hence, interactions with the side chain of residue 256 must be of minor importance in the folding process. Alkylation of the SH-group gives a 'side chain' that is somewhat larger than the original side chain of isoleucine and contains a carboxyl group. Thus, the protein structure must be rather flexible in this part since the necessary adjustment apparently does not affect the yield of the folding process.

Since the structures of those protein molecules that fold into inactive (25%) enzyme molecules are unknown we decided to measure the incorporation of radioactivity into only those molecules that gave active enzyme at the end of the renaturation experiment. To accomplish this, the material after each experiment was purified by affinity chromatography using an agarose gel with covalently linked sulfonamide groups, which bind very specifically only to active enzyme molecules. Analysis of the fraction containing inactive enzyme molecules showed that they were labeled to a substantially lower degree than active enzyme wht ICH_2COO^- was added at the onset of renaturation. Only 15% of the totally incorporated radioactivity was found in the fraction containing inactive enzyme. The solution was opalescent and after filtration most of the radioactivity was found on the filter. This might indicate that the main cause of formation of inactive enzyme species is aggregation of partially folded enzyme molecules. Support for this idea comes from a recent study on the folding of bovine carbonic anhydrase II [13]. With the intention to study the folding process using EPR methods the spin-label-*N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)-iodoacetamide was introduced at position 256 in the denatured I256C/C206S mutant. All attempts to renature this modified protein resulted in a totally inactive protein showing that gross alterations at the C-terminal are deleterious for the folding process or for the final stability of the protein. Earlier it has been shown that removal of the last three C-terminal residues reduces the stability of the protein [14].

3.2. *Stability of the I256C/C206S mutant*

The stabilities of HCAII and the different variants thereof were tested by measuring their resistance towards denaturation in GuHCl solutions. To probe the stability we determined the CO_2 hydration activity remaining after 24 h of incubation. Assuming a two state transition for the inactivation process the equilibrium unfolding constant, K_u, is calculated from

$$e^{-\Delta G_u/RT} = K_u\,(a_n - a)/(a - a_u) \quad \text{(Eqn. 1)}$$

where a is the observed enzymic activity, a_n and a_u are the activities of the native and denatured protein forms [15]. A least squares analysis presented in Fig. 2 was used to fit data to the equation

$$\Delta G_u = \Delta G_u^{H_2O} - m[\text{GuHCl}] \quad \text{(Eqn. 2)}$$

The stabilities relative to the stability of C206S are also presented as a difference in free energy of stabilization $\Delta\Delta G_u$, where

$$\Delta\Delta G_u = \Delta G_u(C206S) - \Delta G_u \text{ (protein variant)} \quad \text{(Eqn. 3)}$$

Values of $\Delta G_u^{H_2O}$, $\Delta\Delta G_u$ and m are listed in Table I along with the denaturant concentration at the midpoint of denaturation curves $[\text{GuHCl}]_{\frac{1}{2}}$. To avoid long extrapolations the values of $\Delta\Delta G_u$ were calculated at 0.95 M GuHCl which is the midpoint of denaturation for the reference protein (C206S). All variants of carbonic anhydrase II that are produced in *E. coli* differ from the enzyme isolated from erythrocytes in that their N-termini are not acetylated. A comparison of erythrocyte HCA II with the 'cloned' enzyme produced in *E. coli* shows that the acetylation has a negligible effect on the enzyme stability ($\Delta\Delta G$ = 0.2 kcal/mol at 0.95 M GuHCl).

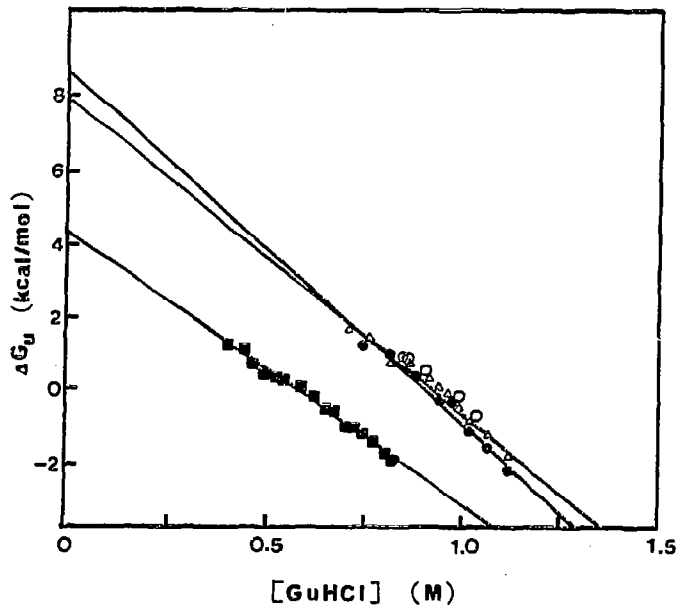

Fig. 2. Free energy of unfolding of erythrocyte and cloned HCA II and mutants as a function of GuHCl concentration. ΔG_u was calculated from Eqn 1. The proteins are designated by the following symbols: (○), C206S; (-△-), cloned HCAII; (-●-), erythrocyte HCAII; (-■-), I256C/C206S. The line for the C206S mutant has been omitted for lack of space.

The mutation C206S, where the natural cysteine in position 206 is replaced by a serine residue, also results in what is probably an insignificant difference ($\Delta\Delta G_u$ = 0.2 kcal/mol) in stability. However, introduction of a second mutation in position 256, where cysteine is a substitute for isoleucine (the I256C/C206S mutant), lowers the protein stability by 2.8 kcal/mol in comparison with C206S mutant. A reduction of the hydrophobic effect may be responsible for the decreased stability in I256C/C206S mutant. Replacing the aliphatic group in isoleucine by the smaller sulfhydryl group in cysteine apparently diminishes the hydrophobic effect. In addition a cavity is created in the protein which leads to less hydrophobic contributions from groups lining the cavity. Recent studies of mutant proteins suggest that the creation of a cavity the size of a $-CH_2-$ group destabilizes the enzyme by 1.1 kcal/mol, and a cavity the size of three such groups by 4.0 kcal/mol [16]. Hence, the destabilization of the I256C/C206S mutant by 2.8 kcal/mol possibly indicates a slight stabilizing effect of the SH-group in comparison with the cavity created by the loss of three CH_2 groups.

3.3. *Kinetics of renaturation*

The kinetics of renaturation of erythrocyte and cloned HCAII and the mutants I256C/C206S and C206S was followed by measuring the reappearance of enzymatic activity and, except for C206S which is devoid of cysteine, by incorporation of radioactive $-CH_2COO-$ into the protein.

3.3.1. *Global folding.* The reappearance of activity was taken as a measure of folding into the final, native state. Reactivation of the enzymes was measured by taking aliquots from the renaturation solution at specified times followed by assay of the CO_2-hydration activity. Dilution into the assay buffer has been found to halt the renaturation process effectively. The time

Table I

Parameters characterizing GuHCl denaturation

protein	$[GuHCl]^a_{1/2}$ (M)	$\Delta G_u^{H_2O}$ (at 0 M GuHCl) (kcal/mol)	m^b [kcal/(mol M)]	$\Delta\Delta G_u$ (at 0.95 M GuHCl) (kcal/mol)
C206S	0.95	7.9	8.2	
Cloned HCA II	0.93	7.9	8.5	0.1
Erythrocyte HCA II	0.91	8.7	9.5	0.4
I256C/C206S	0.57	4.2	7.4	2.8

[a]Midpoint concentration of the GuHCl-induced unfolding transition.
[b]From Eqn 2.
[c]$\Delta\Delta G_u$ is calculated from $\Delta\Delta G_u = \Delta G_u$ (C206S)$-\Delta G_u$ (protein variant).

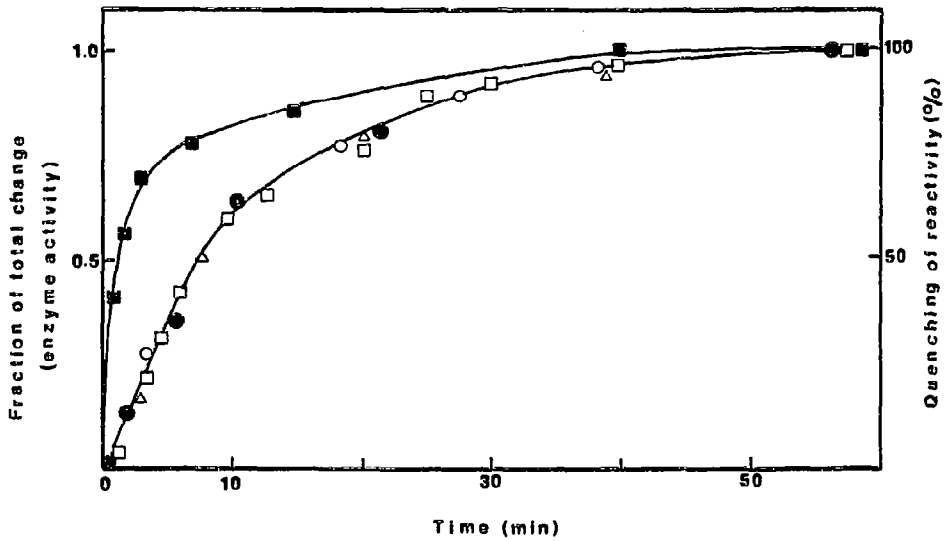

Fig. 3. Recovery of enzyme activity and quenching of chemical reactivity of Cys-256 towards [^{14}C]iodoacetate during renaturation. The quencing of accessibility at Cys-256 is shown by the symbol (-■-). The reactivation of the different forms of the protein are shown by the following symbols: (○), C206S; (△), cloned HCAII; (●), erythrocyte HCAII; (-□-), I256C/C206S.

courses for reactivation of erythrocyte and cloned HCAII and the mutants as well as the carboxymethylated I256C/C206S mutant were found to be virtually identical (Fig. 3). Hence, the incorporation of CH_2COO^- at position 256 does not have any significant impact on the folding process.

3.3.2. *Local folding.* To study the local folding at position 256 close to the C-terminal we added I[2-^{14}C]CH_2COO^- to the renaturation solution at different times after the onset of renaturation. The incorporation of radioactivity into the protein was taken as a measure of solvent accessibility of cysteine in position 256 at the time of addition of ICH_2COO^- (see section 2 for a detailed description of the kinetic measurements). Protection of the SH-group of cysteine 256 from the reagent during the renaturation was described resonably well by a single exponential function giving a half-time of 75 s (Fig. 3). For comparison an identical experiment using HCAII was performed in which the accessibility of the cysteine at position 206 was investigated. The result was that only 410 cpm/nmol radioactivity was incorporated at position 206 even when ICH_2COO^- was present from the onset of renaturation. A similar experiment where native HCAII in 0.3 M GuHCl was incubated with iodoacetate gave an incorporation of 370 cpm/nmol. These results are compatible with results from previous studies where the SH-group at position 206 was modified with a spin label and studied with ESR methods. Those studies indicated that the side chain at this position is buried within less than 0.1 s of renaturation [1].

3.4. *Concluding remarks*

A comparison of the folding processes as described by the kinetics of reactivation and by the kinetics describing burial of the side chains of position 256 and 206 into the protein molecule lead us to the following comments.

1. The amino acid residue 206 is situated in β-strand 7 which is contained in the central part of the dominating β-sheet structure. Notably the β-strands 6 and 7 constitute the most hydrophobic parts of the polypeptide according to a mean hyropathy profile calculated as described by Kyte and Doolittle [17]. This observation and the kinetic results showing burial of residue 206 within 0.1 s of renaturation [1] indicate that the structure composed of β-strands 6 and 7 adopts a compact structure very early during folding and might act as an initiation site.
2. A compact conformation is achieved much later in the region around position 256 as evidenced by the half time of 75 s for burial of the side chain into the protein. This may indicate that not all of the β-structure is formed early in the folding process. One reason why a compact structure in this region is formed late may be that a *trans–cis* isomerization at the peptide bond between Ser-29 and Pro-30 affects the folding. The enzyme contains two proline residues (position 30 and position 202) with peptide bonds in cis conformation. Proline 30 is only 9 residues away from residue 39 which directly interacts with the side chain at position 256 and is part of its shield against solvent. *Cis–trans* isomerizations have half-times similar to the half-time of this process. Another reason for the slow structuration of the C-terminal might be that it is part of a structure with 'knotted' topology (Fig. 1). The polypeptide chain 'crosses' itself at residues 256 and 39. Hence, β-strand 9, which contains residue 256 and is shielded by the outermost β-strand 10, has to come close to its final position before strand 10 comes in place. Otherwise the last 4 residues in the polypep-

tide would have to pass through a narrow tunnel to enter their final position. Thus, the observed kinetics for burial of residue 256 into the protein most probably reflects the final formation of the outermost β-strand no 10.

3. Enzymatic activity reappears even later in the process with a half-time around 8 minutes. The rate of this step may also be limited by *trans–cis* isomerization at proline residues. But in this case isomerization at both position 202 and position 30 might contribute to the rate. The proline at position 30 is part of the wall of the active site and situated only 5 Å from the Zn^{2+} ion.

Acknowledgements: This study has been financially supported by grants to U.C. from the Swedish National Board for Technical Development (732-88-04392), and O.E. and Edla Johansson Vetenskapliga Stiftelse, and a grant to B.-H.J. from the Swedish Natural Science Research Council (K-KU 9426-300). A fellowship to L.-G.M. from Sven and Lilly Lawski's Fond is further kindly acknowledged. We thank S. Lindskog for helpful discussions and reading the manuscript.

REFERENCES

[1] Carlsson, U., Aasa, R., Hendersson, L.E., Jonsson, B.-H. and Lindskog, S. (1975) Eur. J. Biochem. 52, 25–36.
[2] Bergenhem, N., Carlsson, U. and Karlsson, J.-Å. (1989) Int. J. Pept. Prot. Res. 33, 140–145.
[3] Eriksson, A.E., Jones, T.A. and Liljas, A. (1988) Proteins: Structure, Function and Genetics 4, 274–282.
[4] Creighton, T.E. (1984) Proteins, Structures and Molecular Properties, Freeman, New York, pp. 226.
[5] Jones, A.T. (1978) J. Appl. Cryst. 11, 268–272.
[6] Nozaki, Y. (1972) Methods enzymol. 26, 43–50.
[7] Khalifah, R.G., Strader, D.J., Bryant, S.H. and Gibson, S.M. (1977) Biochemistry 16, 2241–2247.
[8] Henderson, L.E. and Henriksson, D. (1973) Anal. Biochem. 51, 288–296.
[9] Kunkel, P.A. (1985) Proc. Natl. Acad. Sci. USA 82, 488–492.
[10] Forsman, C., Behravan, G., Osterman, A. and Jonsson, B.-H. (1988) Acta Chem. Scand. B42, 314–318.
[11] Rickli, E.E., Ghazanfar, S.A.S., Gibbons, B.H. and Edsall, J.T. (1964) J. Biol. Chem. 239, 1065–1078.
[12] Carlsson, U., Henderson, L.E. and Lindskog, S. (1973) Biochim. Biophys. Acta 310, 376–387.
[13] Cleland, J.L. and Wang, D.I.C. (1990) Biochemistry 29, 11072–11078.
[14] Carlsson, U., Henderson, L.E., Nyman, P.O. and Samuelsson, T. (1974) FEBS Lett. 48, 167–171.
[15] Pace, C.N. (1975) CRC Cric. Rev. Biochem. 3, 1–43.
[16] Kellis, J.T., Jr, Nyberg, K., Sali, D. and Fersht, A.R. (1988) Nature 333, 784–786.
[17] Kyte, J. and Dolittle, F. (1982) J. Mol. Biol. 157, 105–132.