Mapping the Folding Intermediate of Human Carbonic Anhydrase II. Probing Substructure by Chemical Reactivity and Spin and Fluorescence Labeling of Engineered Cysteine Residues†

Magdalena Svensson,‡ Per Jonasson,§ Per-Ola Freskgård,‡ Bengt-Harald Jonsson,§ Mikael Lindgren,|| Lars-Göran Mårtensson,§,⊥ Massimiliano Gentile,‡ Kristina Borén,‡ and Uno Carlsson*,‡

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

Received January 10, 1995; Revised Manuscript Received May 1, 1995⊗

ABSTRACT: Several conformation-sensitive parameters have shown that human carbonic anhydrase II exists as a stable and compact equilibrium folding intermediate of molten globule type. In this study we have continued a previously initiated mapping of the intermediate structure. Cys residues were engineered, one at a time, into various regions of the protein structure, so as to obtain chemically reactive probes and handles for spectroscopic probes. These probes were used to specifically report on conformational changes accompanying the folding process. Thus, the accessibility of the introduced Cys residues to specific chemical labeling by radioactive iodoacetate was used to monitor the stability and compactness of the substructure surrounding each Cys residue. In addition, a spin-label (nitroxide radical) and a fluorescent probe (IAEDANS) were attached to the inserted SH-groups to give complementary information. The mobility of the spin-label was used to indicate local changes in structure, and the fluorophore was used to probe local changes in polarity at various stages of unfolding. Much of the predominant β-structure, consisting of 10 β-strands extending throughout the entire molecule, appears to be compact and largely intact in the intermediate. Thus, β-strands 3−7, probed at positions 68, 97, 118, 123, 206, and 245, seem to have a native-like structure in the folding intermediate. In contrast, a more flexible structure is found around positions 56, 176, and 256 in the peripheral β-strands 1, 2, and 9, showing that the stability of the secondary structure in the intermediate state is less in the outer parts of the protein. A hydrophobic region, containing β-strands 3−5, seems to be remarkably stable and is not ruptured until strong denaturing conditions (5 M GuHCl) are applied. The stability of this hydrophobic β-core appears to increase toward the center. This stable region is contained in the middle of a sequentially continuous antiparallel structure that spans β-strands 2−6, suggesting that this part might represent a site where folding is initiated.

A significant goal in biophysical chemistry is to understand how a specific amino acid sequence encodes all the information required for a protein to fold spontaneously into a unique, three-dimensional and biologically active conformation (Anfinsen & Scheraga, 1975). Experiments indicate that the native state of a protein is reached via formation of folding intermediates (Kuwajima, 1989; Kim & Baldwin, 1990). Under certain experimental conditions one can enrich different equilibrium intermediates, making it feasible to study the structural properties of the intermediate states. Investigations have shown that most of these intermediates have structural features in common, and this state is most often referred to as the molten globule state (Ohgushi & Wada, 1983; Ptitsyn, 1992). Moreover, the molten globule state has been suggested to represent a universal intermediate in the folding pathway (Ptitsyn et al., 1990). Interestingly, some kinetic molten globule intermediates have by pulsed hydrogen exchange methods been found to be in close similarity to the equilibrium molten globule state (Ptitsyn, 1995; Baldwin, 1993; Jennings & Wright, 1993). It has also recently been discovered that there is residual structure in unfolded proteins even if unfolding was induced by exposure to very strong denaturing conditions (Neri et al., 1992; Shortle, 1993; Mårtensson et al., 1993; Oliveberg et al., 1994).

Previously, it has been demonstrated that bovine carbonic anhydrase II can form a stable folding equilibrium intermediate (Henkens et al., 1982; Dolgikh et al., 1984). In a previous study, we showed that human carbonic anhydrase II (HCAII)[1] also forms a stable and compact intermediate at a moderate concentration of GuHCl (Mårtensson et al., 1993). Hence,

† This work was supported by grants from the Swedish National Board for Industrial and Technical Development (88-04439P, U.C.; 88-04392P, B.-H.J.), the Swedish Natural Science Research Council (K-Ku 4241-301, U.C.; K-Ku 9426-300, B.-H.J.; K 7375-309, M.L.), Stiftelsen Bengt Lundqvists Minne (M.S.), Sven och Lily Lawskis Fond (P.-O.F. and L.-G.M.), Carl Tryggers Stiftelse för Vetenskaplig Forskning (M.L. and U.C.), and Knut och Alice Wallenbergs Stiftelse (M.L.).

* To whom correspondence should be addressed.

‡ IFM-Department of Chemistry, Linköping University.

§ Umeå University.

|| IFM-Department of Chemical Physics, Linköping University.

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

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

[1] Abbreviations: EPR, electron paramagnetic resonance; GuHCl, guanidine hydrochloride; HCAII, human carbonic anhydrase II; $HCAII_{pwt}$, pseudo-wild-type of HCAII; IAEDANS, 5-((((2-iodoacetyl)-amino)ethyl)amino)naphthalene-1-sulfonic acid; IPTG, isopropyl β-thiogalactopyranoside; NBD-Cl, 7-chloro-4-nitrobenzofurazan.

0006-2960/95/0434-8606$09.00/0

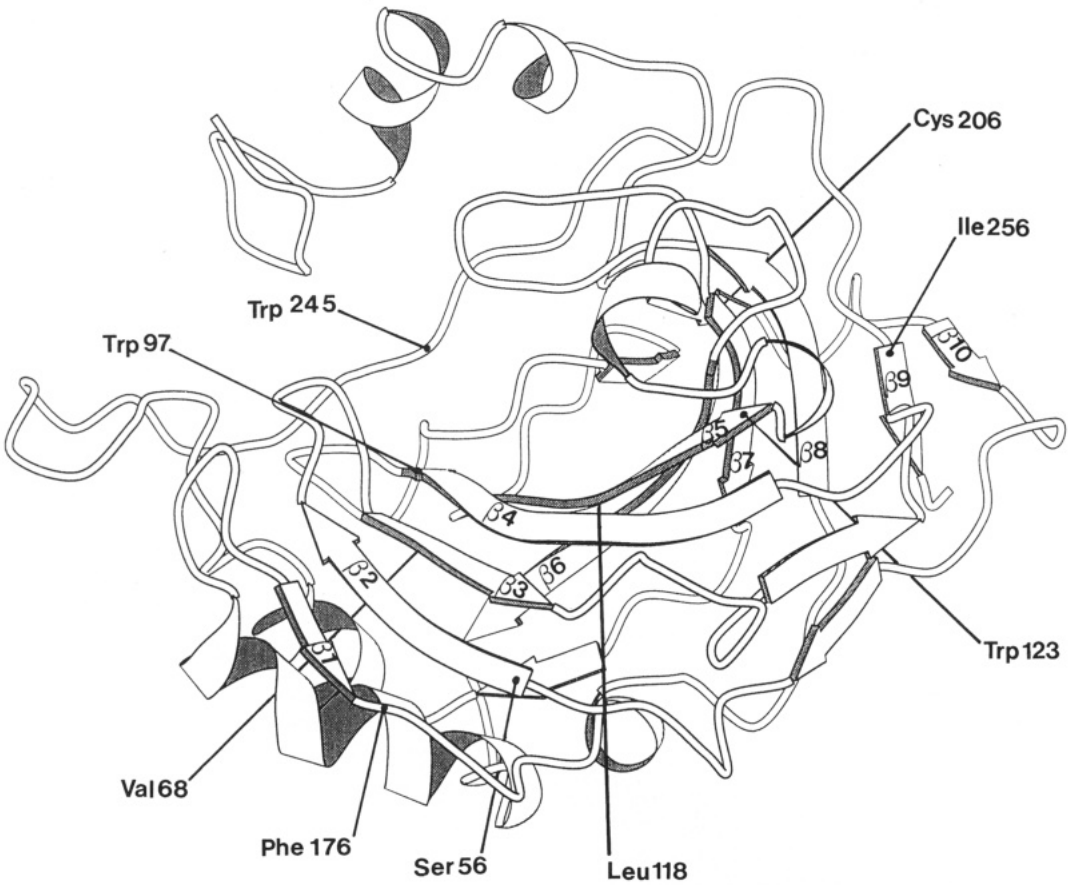

FIGURE 1: Schematic view of the polypeptide backbone of human carbonic anhydrase II. Positions for substituted amino acid residues are indicated. This figure was drawn using the program Molscript (Kraulis, 1991).

in the present study we used HCAII as a model protein for structural characterization of folding intermediates.

HCAII has been very well characterized, both structurally and functionally. This protein has a molecular weight of 29 300; its primary structure is known (Henderson et al., 1976), and its X-ray structure has been determined to high resolution (Eriksson et al., 1988; Håkansson et al., 1992). The enzyme consists of a single domain, with partial helix structure and a dominating β-sheet structure that divides the molecule into two halves. The upper half as illustrated in Figure 1 consists mainly of the N-terminal region and the active site, and the lower half contains an extensive hydrophobic core.

We have previously probed various substructures of the folding intermediate of HCAII by chemical labeling of SH-groups introduced by site-directed mutagenesis, and by doing so, we were able to partially map the intermediate (Mårtensson et al., 1993). The results of that study indicate that the folding intermediate has an ordered native-like secondary structure in the central part of the β-sheet core, whereas the parts of the β-sheet that are situated close to the surface appear to be less ordered. The large hydrophobic cluster seemed to be present in the intermediate and to remain stable even under strong denaturing conditions.

To further characterize the intermediate, we constructed another set of mutants in which Cys residues were introduced, one in each mutant, in various positions of the structure. The SH-groups of these Cys residues were exploited as handles to which spectroscopic probes, *i.e.*, a spin-label and a fluorescent probe (Figure 2), were attached,

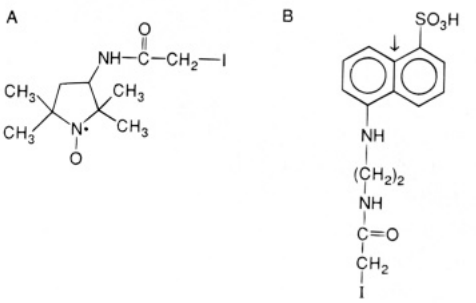

FIGURE 2: Panel A: Spin-label reagent, *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide. The maximal and minimal distances occupied by the spin-label in the protein (9.7−5.5 Å) were calculated from the attachment point (Cys-C_α) and the ring-N in the spin-label. Panel B: Fluorescent reagent, 5-((((2-iodoacetyl)-amino)ethyl)amino)naphthalene-1-sulfonic acid (1, 5-IAEDANS). The maximal and minimal distances occupied by the AEDANS moiety in the protein (13.5−4.4 Å) were calculated from the attachment point (Cys-C_α) to the naphthalene C-9; see arrow). The calculations of the distances above are explained in Material and Methods.

and the inserted Cys residues were furthermore used as chemical reactivity probes, as in the cited study above.

A spin-label that is covalently attached to a Cys residue in various positions of the protein structure will report on its local environment; *e.g.*, information regarding mobility is gathered by noting the EPR line shape during various stages of the folding process. This approach has been successfully applied in our earlier studies on HCAII and mutants thereof (Carlsson et al., 1975; Lindgren et al., 1993) and in work on other proteins (Todd et al., 1989; Calciano et al., 1993; Liu et al., 1994). In the present experiments, the Cys mutants were also labeled specifically with the

widely used fluorescent reagent IAEDANS [*e.g.*, Murry-Brelier and Goldberg (1989)] in order to monitor local changes in polarity that accompany the formation of folding intermediates. The chemical reactivity approach measures accessibility and reveals compactness of various substructures (Ballery et al., 1990; Mårtensson et al., 1993).

MATERIAL AND METHODS

Chemicals. Guanidine-HCl, sequential grade, was purchased from Pierce. Concentrations of GuHCl were confirmed by measuring the index of refraction (Nozaki, 1972). Isopropyl β-thiogalactopyranoside (IPTG) was obtained from Promega. 7-Chloro-4-nitrobenzofurazan (NBD-Cl) was obtained from Fluka, and iodo[2-^{14}C]acetic acid (54 mCi/nmol) was purchased from Amersham. 5-((((2-Iodoacetyl)amino)-ethyl) amino)naphthalene-1-sulfonic acid (1,5-IAEDANS) was obtained from Molecular Probes, and *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide was purchased from Sigma. All other chemicals were of reagent grade.

Spectrophotometers. Light absorbance was measured on either a Perkin-Elmer 320 or a Hitachi U2000 spectrophotometer. Fluorescence was measured on a Shimadzu RF-5000.

Protein Isolation and Purification. The *in vitro* site-directed mutagenesis was performed as described in Freskgård et al. (1994). All enzyme variants were produced and purified as described in Mårtensson et al. (1993, 1995).

Stability Measurements. Monitoring the denaturation of the enzyme variants by measuring UV absorbance (A_{292}/A_{260}) as a function of GuHCl concentration was performed according to Mårtensson et al. (1993). The inactivation was followed by CO_2 hydration activity measurements, as described elsewhere (Rickli et al., 1964; Mårtensson et al., 1992).

Data Analysis. The course of enzyme unfolding is sequential, and it was therefore possible to evaluate the data by analysis of a three-state model including a native state (N), an intermediate state (I), and an unfolded state (U) (Mårtensson et al., 1993). The equilibrium data were converted to plots of the apparent fraction of unfolded protein, F_{app}, by using an equation derived by Garvey and Matthews (1989):

$$F_{app} = [\exp(-\Delta G_{NI}/RT)][Z + \exp(-\Delta G_{IU}/RT)]/ [1 + \exp(-\Delta G_{NI}/RT) + \exp(-\Delta G_{IU}/RT)]$$

where ΔG_{NI} and ΔG_{IU} are the apparent free energy differences between native and intermediate protein species and between intermediate and unfolded species, respectively, and $Z = [(A_{292}/A_{260})_I - (A_{292}/A_{260})_N]/[(A_{292}/A_{260})_U - (A_{292}/A_{260})_N]$. A_{292}/A_{260} is the observed ratio of absorbance, and $(A_{292}/A_{260})_N$, $(A_{292}/A_{260})_I$, and $(A_{292}/A_{260})_U$ are the ratios for the native, intermediate, and unfolded forms, respectively. The fits were done by using a nonlinear least-squares fitting program, Table Curve (Jandel Scientific).

Carboxymethylation and Measurements of Radioactivity. Carboxymethylation of the enzyme variants in various GuHCl concentrations was performed as described elsewhere (Mårtensson et al., 1993).

Labeling with 1,5-IAEDANS and Fluorescence Measurements. All the enzyme variants were incubated in 5 M GuHCl, except for V68C (8 M GuHCl), buffered with 0.1 M Tris-H_2SO_4, pH 7.5, and a 10-fold molar excess of 1,5-IAEDANS was used; incubations were carried out for 1−2 h at 23 °C in the dark. The labeled proteins were allowed to refold by diluting the denaturant to 0.15−0.30 M GuHCl, depending on the stability of the mutant, and the process was monitored by measuring the recovery of enzymatic activity for 90 min. Refolded and labeled protein was concentrated and dialyzed against 0.1 M Tris−H_2SO_4, pH 7.5, to remove excess reagent. All enzyme variants were purified as described in Lindgren et al. (1993). The degree of AEDANS labeling was determined by measuring the absorbance at 337 nm, using the extinction coefficient 6100 M^{-1} cm^{-1} (Jullien & Garel, 1981). Independent assessment of the degree of labeling was accomplished by adding NBD-Cl in a 40-fold molar excess to the labeled protein in the denatured state. Quantitation was done spectrophotometrically, using ϵ_{420} = 13 000 M^{-1} cm^{-1} for the NBD-sulfur adduct (Birkett et al., 1971). For the labeled mutants (0.85 μM) incubated for 24 h at 23 °C in various concentrations of GuHCl buffered with 0.1 M Tris−H_2SO_4, pH 7.5, fluorescence emission spectra were recorded at 400−550 nm; the excitation wavelength was 350 nm, with bandwidths of 5 nm for both excitation and emission. All measurements were made in a 1-cm quartz cell, and the sample compartment was maintained at 23 °C.

Labeling with a Spin-Probe and EPR Measurements. The labeling of SH-groups with the spin-probe *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide was performed as described earlier (Lindgren et al., 1993). Briefly, the enzyme variants (34 μM) were allowed to react with the spin-probe (680 μM) in 5−6 M GuHCl, except for V68C (8 M GuHCl), buffered with 0.1 M Tris-H_2SO_4, pH 7.5, at 23 °C for 20 h in the dark. An aliquot of the solution was withdrawn, and the degree of labeling was assessed by adding NBD-Cl in a 10-fold molar excess to the labeled protein in the denatured state. Quantitation was done as described above. The EPR measurements were performed as described in Lindgren et al. (1993) with one modification: the double ER102ST cavity was replaced by an ER4103TM cavity.

Distance Calculations. Molecular dynamic simulations of the AEDANS moiety and the spin-label attached to a Cys residue were performed using the HyperChem program (AUTODESK) which was run on a PC 486/66. The system was run at 400 K over a period of 22 ps in vaccum. The OPLS force-field module was used, and a time step of 0.0005 ps was chosen in this simulation. The distances between C_α and the ring-N of the spin-label, C_α and C-9 in the naphthalene moiety of AEDANS (Figure 2), and C_α and $C_{\epsilon 2}$ in the indole side chain of Trp were registered, and from these simulations we obtained maximal and minimal distances between these atoms.

RESULTS

Production of Cysteine Mutants of HCAII. A pseudo-wild-type of HCAII ($HCAII_{pwt}$) was produced, in which the only naturally occurring Cys, no. 206, was replaced by a Ser residue. In addition, single Cys residues were introduced—one per mutant at different positions—into this $HCAII_{pwt}$. A total of eight mutants were produced in this way (Figure 1 and Table 1), and each of these double mutants contained only one Cys residue, which could be specifically modified by using various sulfhydryl reagents. In other words, the Cys

Table 1: Stability of Unmodified and Spin-Labeled HCA II toward GuHCl Denaturation Measured as Absorbance Change and Presented as Midpoint Concentration (C_m) of Denaturation[a]

enzyme variant	C_{mNI} (M)	C_{mIU} (M)	C_{mNI} (M) spin-labeled	C_{mIU} (M) spin labeled	C_{mNI} (M) AEDANS-labeled	C_{mIU} (M) AEDANS-labeled	rel act.[b]	rel act.[b] spin-labeled	rel act.[b] AEDANS-labeled
HCA II_{pwt}	1.0 ± 0.07[c]	2.1 ± 0.3[c]					100[c]		
S56C	0.6 ± 0.09[d]	2.2 ± 0.3[d]	0.4 ± 0.05	2.1 ± 0.2	0.4 ± 0.09	2.0 ± 0.4	65	65	65
V68C	0.7 ± 0.09[d]	2.1 ± 0.2[d]					89		
W97C	0.5 ± 0.06[c]	1.9 ± 0.3[c]	0.4 ± 0.06	1.6 ± 0.3	0.4 ± 0.08	1.6 ± 0.6	51[c]	51	25
L118C	0.9 ± 0.10	1.7 ± 0.3					62		
W123C	0.6 ± 0.09[c]	2.0 ± 0.6[c]	0.5 ± 0.10	1.7 ± 0.6	0.6 ± 0.15	1.8 ± 0.6	56[c]	55	47
F176C	0.6 ± 0.06	2.1 ± 0.2	0.4 ± 0.08	1.8 ± 0.3	0.6 ± 0.14	1.9 ± 0.5	83	65	76
C206	0.9 ± 0.07[d]	2.4 ± 0.3[d]	0.6 ± 0.06	2.0 ± 0.2	0.7 ± 0.10	1.9 ± 0.5	107	100	85
W245C	0.6 ± 0.10[c]	2.1 ± 0.4[c]	0.5 ± 0.10	1.8 ± 0.2	0.6 ± 0.18	1.9 ± 0.7	100[c]	95	77
I256C	0.6 ± 0.10[d]	2.3 ± 0.5[d]			0.4 ± 0.10	1.8 ± 0.5	100		95

[a] C_{mNI} and C_{mIU} represent the transition midpoint concentrations (GuHCl) for the transitions from the native to the intermediate and from the intermediate to the denatured enzyme, respectively. [b] The relative activity in 0 M GuHCl is compared to HCA II_{pwt}. [c] Data previously published by Mårtensson et al. (1995). [d] Data previously published by Mårtensson et al. (1993). The errors shown are standard deviations.

residues could be used as chemical reactivity probes and also as handles to which a fluorescent or an EPR spin-probe (Figure 2) could be attached in order to explore the structural changes that occur in these specific parts of the protein structure upon GuHCl-induced denaturation.

Effect of Mutation on the Stability of the Enzyme. The stability of the various mutants in regard to GuHCl denaturation, was measured as the change in the absorbance ratio A_{292}/A_{260}. This parameter mainly reflects the exposure of Trp residues during unfolding of the protein. There are seven Trp residues distributed in the 3-D structure of HCAII; thus we take these changes in UV absorbance to reflect overall conformational changes in the molecule (global transitions). The data from these transition curves are presented as midpoint concentrations of denaturation, see Table 1.

As has been previously demonstrated, HCAII and the engineered variants thereof unfold into two well-separated transitions, indicating the existence of a stable intermediate (I) with residual structure. The enzymatic activity vanishes during the first global unfolding transition (designated N → I), leading to an inactive intermediate (Mårtensson et al., 1992, 1993). The midpoint concentration of denaturation (C_m) of the N → I transition varies between 0.5 and 0.9 M for the various Cys double mutants, and the corresponding C_m values for HCAII and $HCAII_{pwt}$ are 0.9 and 1.0 M, respectively. The second global unfolding transition (designated I → U) appears to be less sensitive to the investigated mutations with the exception of the L118C mutant.

The noted decreased stabilities of the Cys mutants are not likely to reflect large disturbances of the protein conformation, since the mutations have only minor effects on the catalytic activity. Thus, although W97C is the least active of the mutants, it still exhibits a considerable specific CO_2 hydration activity that corresponds to 51% of the activity of $HCAII_{pwt}$. Furthermore, 2-D NMR spectra of ^{15}N labeled Trp mutants, including W97C and W123C as well as $HCAII_{pwt}$, indicated that the mutations had essentially no long-range effects on structure and that the perturbations of structure in the vicinity of the mutated Trp were small (Mårtensson et al., 1995). Near-UV CD spectra of all Trp mutants also corroborate this conclusion (Freskgård et al., 1994).

Unfolding Measured by Chemical Labeling. In a previous study (Mårtensson et al., 1993) we used a chemical reactivity approach, and in the present report, we used the same technique to examine several new positions in the enzyme structure. The rate of incorporation of the radioactivity labeled reagent, iodoacetate, gives information about local conformational changes in the substructures surrounding each point of mutation. The modification of the Cys mutants was performed after a 24-h incubation in various concentrations of the denaturant. Relative rates of alkylation of the various mutants as a function of GuHCl concentration are shown in Figure 3A,B. In a separate experiment titration of the SH-group with NBD-Cl was performed on protein variants incubated in 5 M GuHCl to ensure that no interfering intermolecular disulfide bond formation had occurred in the various mutants. It was possible to achieve almost complete modification of all SH-groups, which shows that the introduced Cys residues were unoxidized and were reactive toward alkylating reagents.

Unfolding Measured with a Fluorescent Probe. To obtain extrinsic fluorescent probes that were covalently attached to various positions in the protein, the single Cys residues were labeled with the Cys-specific reagent 1,5-IAEDANS, which comprises the dansyl fluorochrome (Figure 2B). In the native state of all protein variants the SH-group, which was the target of modification, was unavailable to the reagent, and these enzyme forms were therefore denatured to expose the buried SH-group prior to labeling. The derivatized protein was then renatured, and by using this strategy AEDANS labeling was successful for the following enzyme forms: S56C, W97C, W123C, F176C, HCAII (C206), W245C, and I256C. However, V68C and L118C could not be reactivated after labeling. According to measurements of the absorbance of the bound AEDANS moiety, complete modification was achieved in all instances. This was also verified by titration of the remaining SH-group with NBD-Cl. The introduction of an AEDANS moiety into various positions in the protein structure has only minor effects on the stability since the C_m values do not differ much between the modified variants and the unmodified. Moreover, the specific enzymatic activities of these AEDANS-labeled enzymes were only moderately affected compared to the unlabeled mutants (Table 1). Thus, these results indicate that the AEDANS moiety only causes minor perturbations of the protein conformation. The labeled enzyme variants were incubated at various concentrations of GuHCl, and the denaturation was monitored by measuring fluorescence emission from the AEDANS probe after excitation at 350 nm. Changes in the environment of the probe were monitored by the wavelength shift of the emission maxima versus GuHCl concentration

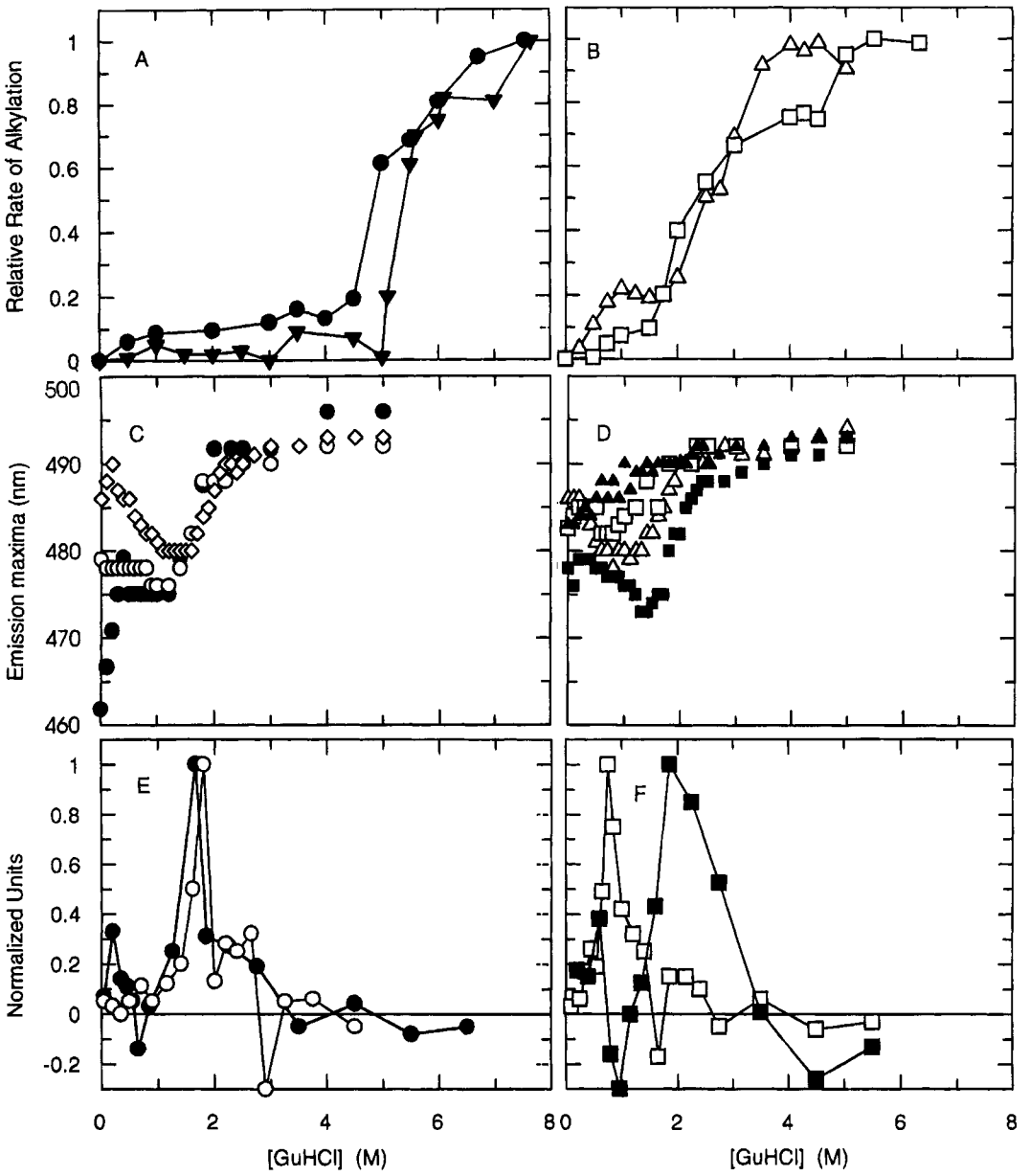

FIGURE 3: Panels A and B: Relative rate of incorporation of radioactive iodoacetate to cysteines in the different protein variants at various GuHCl concentrations. Panel A: W97C (●); L118C (▼). Panel B: F176C (△); W245C (□). The absolute rates of alkylation in cpm $nmol^{-1}$ h^{-1} in the native state and 5 M GuHCl, respectively, are as follows: 70/1840 for W97C, 50/60 for L118C, 80/2750 for F176C, and 130/2760 W245C. Panels C and D: GuHCl denaturation of 1,5-IAEDANS-labeled protein variants monitored as shifts in emission maxima. Panel C: S56C (◇), W97C (●); W123C (○). Panel D: F176C (△); C206 (■); W245C (□), I256C (▲). Excitation was performed at 350 nm. Panels E and F: Relative amplitude of the most narrow component in difference spectra obtained by EPR measurements of spin-labeled protein variants. Panel E: W97C (●); W123C (○). Panel F: C206 (■); W245C (□). The experimental (first derivative) EPR spectrum was normalized to the same total signal before the subtraction, and the difference spectrum thus obtained was divided by the increment in GuHCl concentration.

(Figure 3C,D). A red shift of the fluorescence indicates that the probe has been moved to a more polar environment during the unfolding process, whereas a blue shift indicates the opposite. Additionally, a red shift of the fluorescence is accompanied by a decrease in fluorescence intensity, while a blue shift of the fluorescence is accompanied by an increase in intensity for all variants (data not shown).

When the denaturation is monitored by this probe, we observe for all but one mutant two unfolding transitions. The positions of these transitions are very similar to the corresponding UV absorbance transitions registered for both the unlabeled and labeled enzyme variants, although the labeled W123C mutant lacks the N → I transition (Figure 3C). Interestingly, with AEDANS in positions 56, 176, and 206, the N → I transition is accompanied by a blue shift of the emission spectrum indicating that the environment of the probe is becoming gradually more apolar. Furthermore, an increase of the fluorescence intensity indicates that the fluorophore becomes buried during the N → I transition. It is notable that the total shift in emission maxima was much larger when the fluorescent probe was in position 97 (460 to 495 nm) than when it was in the other examined positions, and that the emission at 5 M GuHCl occurred at essentially the same wavelength for all the AEDANS-labeled variants.

Unfolding Measured with a Spin-Label. As for the fluorescence labeling, it was necessary to spin-label all of the Cys variants while they were in the denatured state. The denatured, modified proteins were renatured, and excess

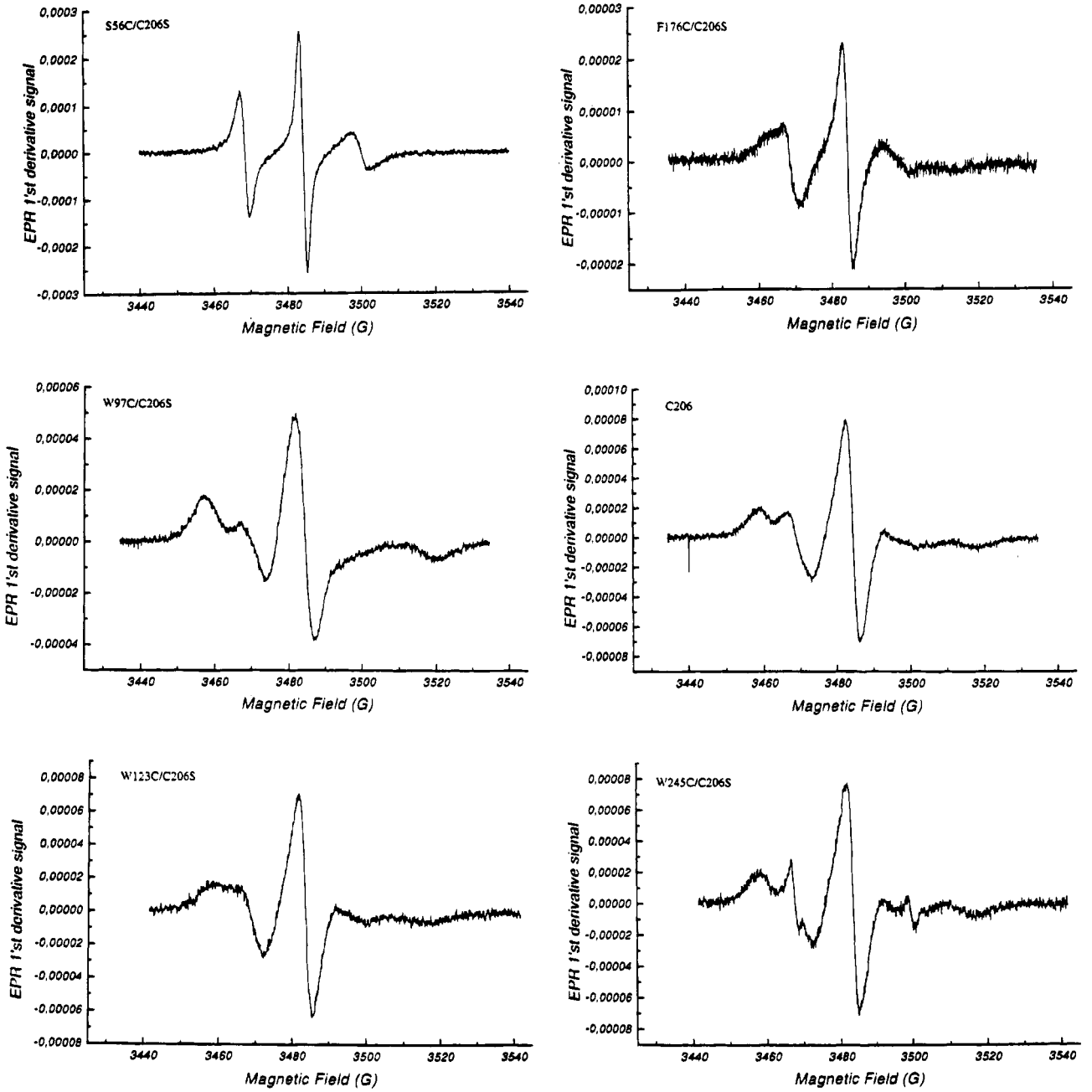

FIGURE 4: EPR spectra of native HCAII variants spin-labeled at a unique cysteine.

reagent was removed by immobilization on an affinity chromatography gel. One advantage of this procedure was that only active, spin-labeled enzyme molecules were recovered after elution regardless of the degree of renaturation. The degree of labeling was at least 95% for all mutants, judged by titration of free SH-groups with NBD-Cl after the modification. Using this procedure, the spin-label (Figure 2A) was successfully attached to Cys residues in positions 56, 97, 123, 176, 206, and 245. The introduction of the spin-label leads to no or minor decreases in specific enzymatic activities compared to the corresponding mutant (Table 1). Furthermore, the stability of the enzyme variants toward GuHCl denaturation was only little affected by the introduction of the spin-label (Table 1). Thus, the moderate decrease in catalytic efficiency and stability indicates that attachment of the spin-label only causes minor perturbations of the protein structure. Spin-labeling of the V68C, L118C, and I256C variants in the denatured state led to adducts that were impossible to reactivate; hence it was impossible to introduce the spin-label into the native states of these mutants.

The EPR spectra of the spin-labeled enzyme variants in the native state are shown in Figure 4. There is a marked difference in the appearance of spectral anisotropic features due to the differing mobility of the spin-label in the various enzyme forms: for slow motions the spectral bands are broader and reveal contributions from the anisotropic hyperfine interaction between the unpaired electron and the nitrogen nucleus of the nitroxide radical moiety of the spin-label, whereas for a rapidly tumbling spin-probe one observes an EPR signal consisting of narrower peaks with a characteristic three-line hyperfine pattern associated with the isotropic hyperfine interaction.

Raising the concentration of denaturant causes a dramatic change in the EPR line shape, and this is illustrated in Figure 5A for the spin-labeled W97C mutant. Here we show

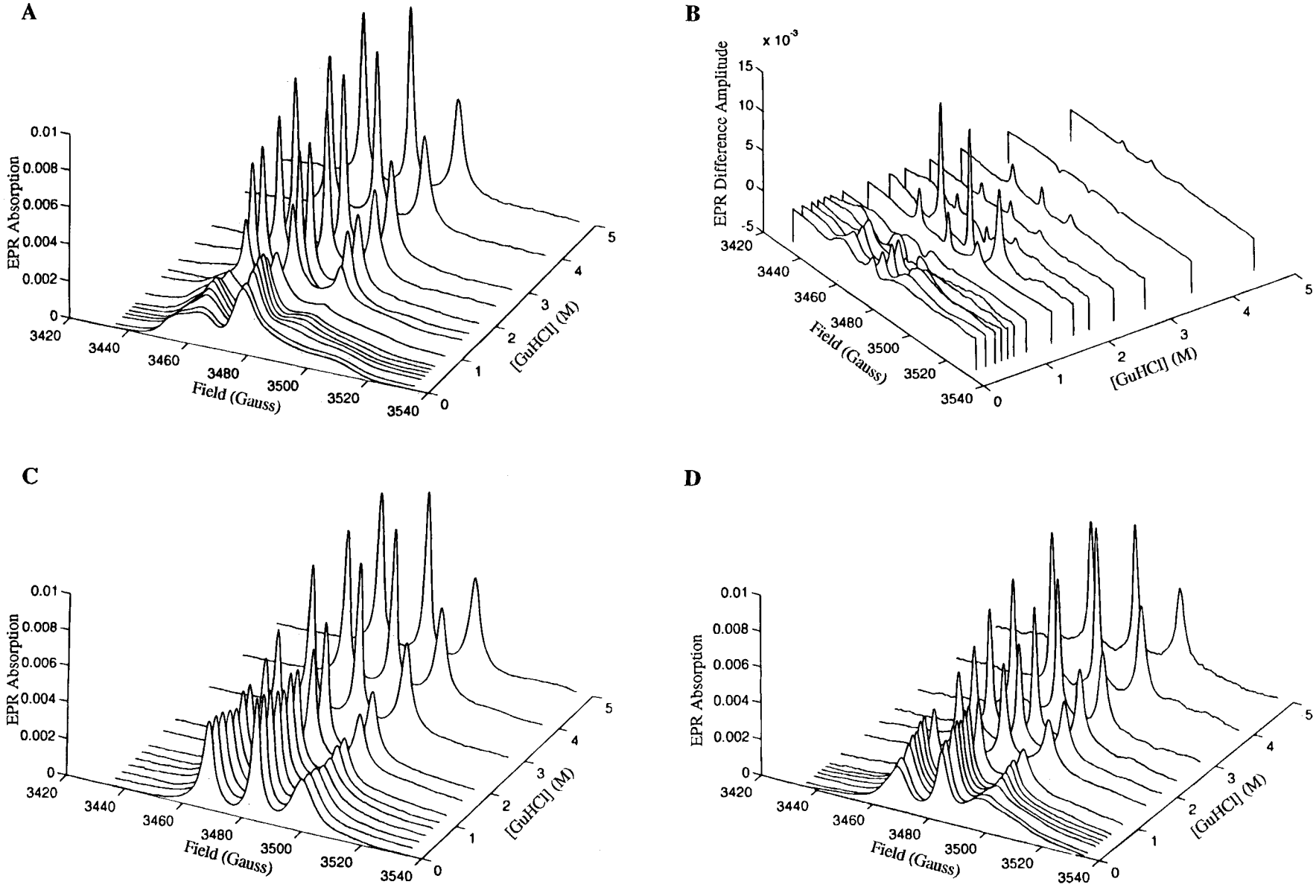

FIGURE 5: Panel A: Absorption EPR spectra of the spin-labeled W97C/C206S mutant at different GuHCl concentrations. (Absorption mode spectra were obtained by integrating the experimental first-derivative spectra.) Panel B: Absorption mode EPR difference spectra of the spin-labeled W97C/C206S mutant at different GuHCl concentrations. Difference spectra were calculated between each consecutive EPR spectrum shown in panel A. Each spectrum is divided by the increment in GuHCl concentration. Panel C: Absorption EPR spectra of the spin-labeled S56C/C206S mutant at different GuHCl concentrations. Panel D: Absorption EPR spectra of the spin-labeled F176C/C206S mutant at different GuHCl concentrations.

absorption EPR spectra, which is the best way to present broad components superimposed onto a more rapidly diffusing component in the same EPR signal. At 0 M GuHCl, a spectrum characteristic of a spin-probe in slow motion is observed. When comparing EPR spectra recorded at different GuHCl concentrations, one observes the gradual increase of components with narrower lines with the increase of GuHCl concentration, and a distinct three-line hyperfine structure has become dominant at ca. 2.5−3.0 M GuHCl. The line shape of the EPR spectra obtained at intermediate GuHCl concentrations indicates the presence of a spin-probe with medium mobility. In this intermediate GuHCl concentration range, the enzyme has been shown to exist as a stable folding intermediate (Mårtensson et al., 1993).

EPR spectra can be used to calculate difference EPR spectra and thereby visualize the unfolding transitions (Figure 5B); as will be shown below, this technique gives under certain circumstances one single parameter useful for a qualitative direct comparison with the other spectroscopic characterization methods used in this study. A difference spectrum is obtained by subtracting the EPR spectrum of a lower GuHCl concentration from the spectrum of the next higher concentration. This procedure is described in detail in Lindgren et al. (1995) and is therefore only briefly explained here. Provided the EPR signatures of the spectral components are different with regard to mobility and changes in relative populations can be expected as the GuHCl concentration is varied, the difference spectrum constructed for small concentration intervals will reveal an increase of a component with positive phase, and a decrease of a component with negative phase. Furthermore, since EPR spectra of the mobile species usually dominate in amplitude over the slower component, approximate information is also obtained on the rotational diffusion from the EPR line shape of the individual spectra.

For the spin-labeled W97C mutant, growth of a distinct, narrow component is observed in the EPR difference spectra at higher GuHCl concentrations (Figure 5B), with maximum growth at ca. 1.6 M, and a slightly broader component appears at lower GuHCl concentrations, with a maximal growth at ca. 0.3 M (Figure 5B). The former component represents the spin-probe in a mobile and essentially unstructured environment, since essentially the same spectrum can be obtained at higher GuHCl concentrations. The spectral component evolving at 0.3−0.5 M represents the spin-probe in a state of medium mobility.

The difference spectra amplitude is a useful (single) parameter for the comparison of the EPR results with the information obtained from other methods used for characterization of protein structure. Thus, the amplitudes (measured at the $m_I = 0$ transition) of the most narrow component of the EPR difference spectra (constructed from the experimental first-derivative spectra) were plotted as a function of GuHCl concentration. In principle any of the transitions in the EPR spectrum can be used. The advantage of using the center lines is that this portion of the spectra has the largest signal to noise ratio, which is critical at low GuHCl concentrations for our experimental conditions. The results obtained for W97C, W123C, C206, and W245C are shown in Figure 3E,F. For the spin-labeled W97C mutant and C206, two transitions can be seen, when considering EPR-monitored unfolding. For the spin-labeled W123C and W245C mutants, however, there is only one main transition. For the W123C variant, the distinct amplitude maximum of the difference spectra is associated with the I → U transition around 2 M GuHCl, whereas for W245C the transition coincides with the N → I transition at ca. 0.8 M GuHCl. This can be interpreted as the EPR signatures of the intermediate forms of these two mutants being similar to those of the native (W123C) and unfolded (W245C) forms, respectively.

For S56C and F176C the changes in rotational diffusion of the spin-probe, going from the native state to the unfolded state, are not large or distinct enough to resolve any further details in a series of difference spectra. Instead we here discuss the EPR spectra recorded for the spin-labeled variants S56C and F176C at various GuHCl concentrations (Figure 5, panels C and D, respectively). The EPR spectra of spin-labeled S56C show characteristics of a spin-label that already in the native state is relatively mobile and do not reveal any significant change during the N → I transition ($C_{mNI} = 0.4$ M; Table 1). Some mobilization of the spin-label can, however, be observed in the intermediate state before the I → U transition ($C_{mIU} = 2.1$ M; Table 1) is reached. After completion of the latter transition, the spin-label becomes fully mobile. In the case of F176C a broad component in the EPR signal disappears in the GuHCl range 0.3−1.0 M. Meanwhile a spectrum evolves that resembles an "unfolded" form. Thus, the structural changes monitored by the spin-probe at position F176C seem to follow the first of the two transitions (N → I), whereafter only minor further mobilization of the spin-label occurs.

To quantify the spectral changes in terms of rotational mobility and relative populations of different spectral components, simulations associated with slow, medium, and rapid rotational diffusion can be used in superposition to fit the experimental data. The W97C mutant will serve as an example to give an indication of the rotational correlation times; the full analysis of all spin-labeled mutants at various GuHCl concentrations is lengthy and will appear elsewhere.

The EPR spectra along with simulations using components of slow, medium, and rapid mobility are shown in Figure 6, with the parameters used in the simulations shown in the figure caption. The spectrum of the native state (Figure 6 A) is associated with a single-component spin-label, apparently tumbling isotropically, with a rotational correlation time of ca. 9.4 ns. This corresponds to the overall motion of the protein structure, and thus the spin-label seems to be held in a "rigid" structure at the position of the engineered cysteine. Above ca. 3 M GuHCl a rapidly rotating component is dominating, and its simulation together with the experimental spectrum is shown in Figure 6C. The rotational correlation time is here in the range 0.4−0.8 ns indicative of a spin-label probing the local motions of the polypeptide chain in a relatively flexible environment. To make a good fit of the spectra recorded at intermediate GuHCl concentrations, it was necessary to simulate superpositions using three spectral components. An example of such a fit is shown in Figure 6B where the medium mobility component shows its strongest contribution to the total signal (ca. 40%). Essentially the same quality of the fit was obtained throughout the whole GuHCl range using only these three components.

The spectra obtained at 3.0 M GuHCl and above were essentially the same for all spin-labeled variants of the protein. Therefore we can regard this as the "unfolded" form, as far as the EPR spin-labeling technique can tell.

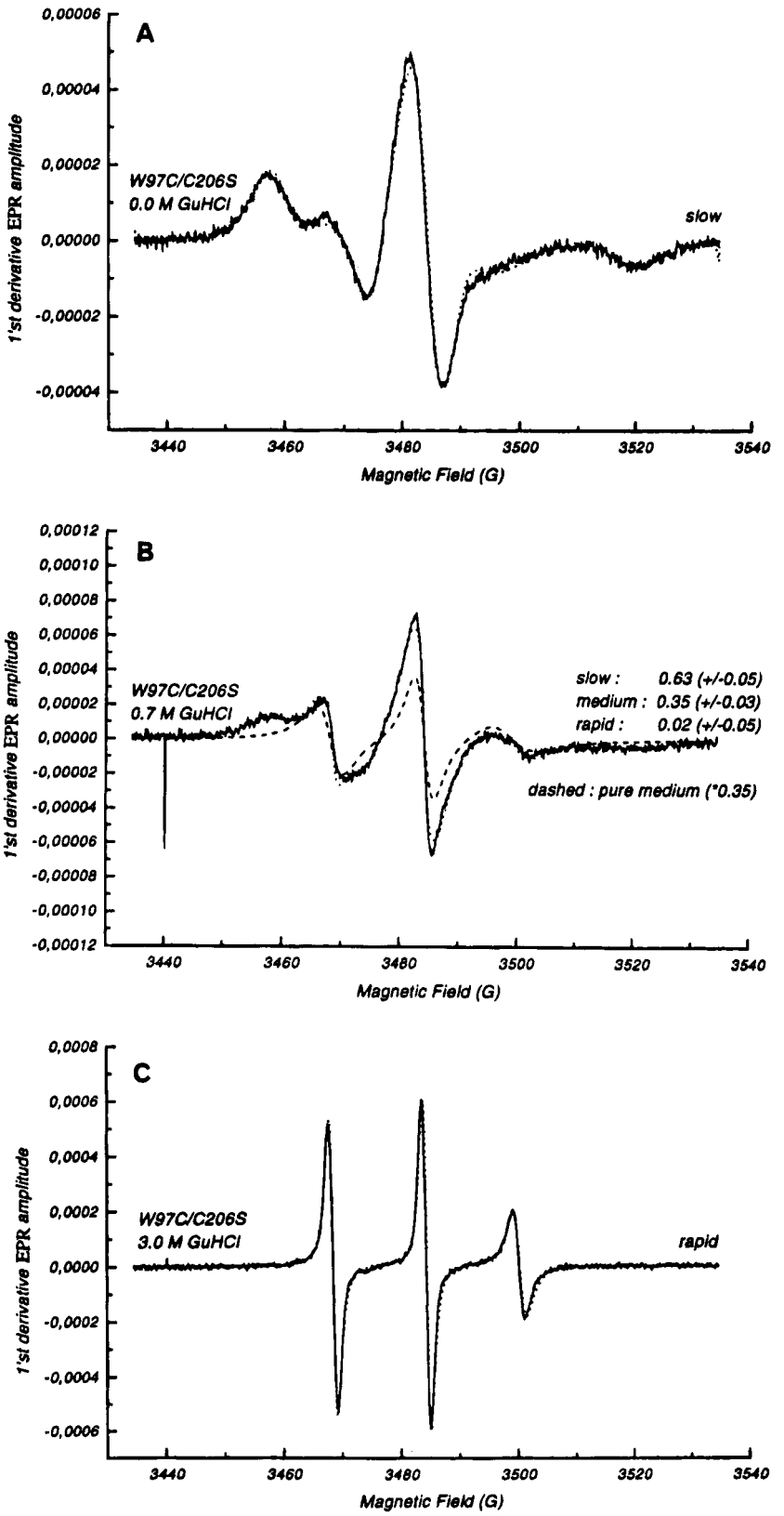

FIGURE 6: EPR spectra and simulations of spin-labeled W97C at selected GuHCl concetrations. EPR hyperfine and g-tensor parameters in all simulations: $g_x = 2.0054$, $g_y = 2.0088$, $g_z = 2.0020$, $A_x = 0.56$ mT, $A_y = 0.665$ mT and $A_z = 3.55$ mT. Rotational diffusion and line-width parameters were as follows. Slow component: $d_{xy} = d_{zz} = 0.175 \times 10^8$ s^{-1}; lw = 0.85 G; gw = 0.95 G. Medium component: $d_{xy} = 0.42 \times 10^8$ s^{-1}; $d_{zz} = 0.14 \times 10^9$ s^{-1}; diffusion tilt, 75°; lw = 0.45 G; gw = 0.75 G. Rapid component: $d_{xy} = 0.22 \times 10^9$ s^{-1}; $d_{zz} = 0.39 \times 10^9$ s^{-1}; diffusion tilt, 75°; lw = 0.4 G; gw = 0.75 G. lw = intrinsic (Lorentzian) line width; gw = Gaussian line width (convoluted). Panel A: Dotted line, simulation of slow component. Panel B: Dotted line, simulation using superposition of slow, medium, and rapid components. The relative weights (insets) were obtained by a least-square fitting procedure with the standard deviation taken as the square root of the variances associated with each weight. Dashed line, simulation of the medium component used in making up the superposition. Panel C: Dotted line, simulation of rapid component. The experimental spectra (lines) were obtained at GuHCl concentrations as indicated in the insets. We have used a program developed by Schneider and Freed (1989). See also Lindgren et al. (1993, 1995).

Table 2: Mobility of the Spin-Probe at Different Positions

position	conformational state[a]		
	native	intermediate	unfolded
56	M	M/R	R
97	S	M	R
123	S	S	R
176	M	R	R
206	S	M	R
245	S	R	R

[a] Notation: S = slow, M = medium, and R = rapid mobility. For details, see text.

The qualitative results of the EPR analysis of the spin-labeled enzyme forms of HCAII are summarized in Table 2. The terms slow (S), medium (M), and rapid (R) are used to indicate the mobility of the spin-label simply, although slow and medium mobilities can differ considerably between the various spin-probe positions.

DISCUSSION

We have previously shown that HCAII forms a stable and compact folding intermediate in GuHCl (Mårtensson et al., 1993). To map the borders of the compact regions of this folding intermediate and of the residual structure in strong denaturing conditions, we extended the cited study by producing another set of Cys mutants and again applying the chemical reactivity strategy; in addition, these SH-groups were used as handles for spectroscopic probes to extract complementary information. This combined approach will allow us to resolve various aspects of local structural and physical changes accompanying the folding process around the mutation site.

Some Characteristics of the Probes. When various types of labels are used to probe structural changes accompanying folding steps, it is important to understand the intrinsic properties of each label to correlate the observed effects to specific features of the protein structure.

For these experiments we introduced Cys residues at buried positions for use as reactive spots. The SH-group most likely occupies a position that is very similar to that of the γ-atom of the wild-type amino acid. Since replacement by Cys residues leads, in all instances, to incorporation of a side chain that is smaller or of similar size compared to the original one, the substitution by itself is not likely to induce rearrangement of the surrounding structure. However, for Cys residues introduced at buried positions near the surface, replacement of larger side chains might cause exposure of the SH-group to the solvent even after small structural perturbations.

The reactivity of the introduced Cys residues was used to monitor local conformational changes occurring during the folding process. The rate of incorporation of a radioactively labeled reagent, in this case iodoacetate, should reflect the accessibility and exposure of the inserted Cys probe. Since the alkylating reagent is comparatively small, changes in the relative alkylation rate should with good precision reflect changes in the compactness of the substructure in the vicinity of the mutation site during the folding process. In addition, there are also other factors that can affect the absolute rate of incorporation of the alkylating reagent, *e.g.*: (1) a different intrinsic pK_a of the thiol group, which is caused by differences in polarity; (2) interference from the closest neighboring amino acid residues in the sequence; and (3) the presence of a compact structure nearby but not encompassing the cysteine that is exposed might limit the diffusional space for the reagent and lead to a decreased alkylation rate. These factors can probably explain the difference in absolute rates of alkylation that was observed between mutants even when the positions seemed to be fully exposed (see values in 5 M GuHCl given in the caption to Figure 3, panels A and B). Therefore, a direct comparison of the absolute alkylation rates of the mutants has not been exploited. On the other hand, the clear dependence of the alkylation rate on the concentration of GuHCl for a given mutant demonstrates that structural exposure of the mutated position is reflected by the changes in the relative alkylation rates regardless of the intrinsic reactivity of the thiol group.

The spectroscopic probes used in this study are environment-sensitive. Thus, the spin-label will report on local changes in mobility, and the fluorescent probe will on the other hand report on local changes in polarity during the folding process. Because of the buried locations of the introduced Cys attachment sites in the native state of the protein, modification with the spectroscopic probes had to be performed in the denatured state followed by renaturation. When probes the size of the spin-label or AEDANS are covalently attached to the thiol group in the denatured state, it is important that the protein structure has the ability to accommodate the added moiety after renaturation. Interestingly, such accommodation proved possible in the native state of all mutants in which a Trp or Phe residue had been replaced by the target Cys residue. When a smaller amino acid residue, situated in the core, was replaced, the labeled mutants could not on the other hand be reactivated. Thus, the size of the created cavity seems to be important. It must be noted that all of these mutants when unlabeled could be renatured after denaturation. This strongly indicates that the covalently attached probe utilizes the cavity created by the mutation, but since the spacer arms of the introduced spectroscopic reporter groups are considerably longer than those of the substituted amino acid residues, these probes must occupy a larger volume. Calculated maximal and minimal distances are 9.7−5.5 Å for the spin-label and 13.5−4.4 Å for the AEDANS moiety (see Figure 2). For comparison the corresponding distances for the Trp indole group were calculated as 5.8−2.8 Å (C_α−$C_{\epsilon 2}$).

Both the spin-label and the fluorescent probe have apolar properties, which favor a position in an apolar environment.

Probing the Native State. Since the three-dimensional structure of the native state is known, each position will not be discussed in detail. However, some observations made from the spectroscopic probes at various locations in the native state of the enzyme molecule are discussed in relation to the X-ray structure. This will roughly indicate the sensitivity of the probes to detect differences in the environment.

All of the mutation sites, except Ser56 and Trp245, were selected to be buried in the native state (Table 3) in order to be unavailable to iodoacetate, the alkylating agent used in the chemical reactivity tests. Accordingly, the labeling rate for the native state of most of the mutants was low, *i.e.*, equal to or less than 80 cpm $nmol^{-1}$ h^{-1} (see caption to Figure 3, panels A and B), which is at least 1 order of magnitude lower than the rate observed for all mutants except L118C in 5 M GuHCl. A somewhat high rate of labeling

Table 3: Solvent Accessibility Surface Area for the Mutated Side Chains in HCA II

residue	Å^2
S56	86
V68	0
W97	0
L118	0
W123	6
F176	7
C206	1
W245	47
I256	9

was, however, noted for W245C (130 cpm nmol^{-1} h^{-1}) and Trp 245 is, according to X-ray data, the second most exposed residue of the substituted residues (Table 3).

The EPR spectra of the spin-labeled enzyme variants also clearly show that the engineered Cys residues are mainly located in the interior of the enzyme mutants. In the native conformation, all attached spin-labels gave rise to spectra characteristic of anisotropic slow-motional features, which indicates that the spin-probe is embedded in the protein structure (Figure 4). Trp97 is deeply buried in the interior of the protein structure and is completely inaccessible from the surface (Table 3). The spin-label in this position also seemed to be the most immobilized of the spin-labels attached to the variants.

In the native state, the fluorescence wavelength maxima of the AEDANS-labeled enzymes varied considerably, implying that the environment of the positions occupied by the dansyl moiety differs in character in the various mutants. The AEDANS-labeled W97C mutant produced the most blue-shifted fluorescence spectrum (λ_{max} = 460 nm), indicating that the surrounding of the probe is apolar (Figure 3C). It has been reported that Trp97 is also fully shielded from solvent and located in an apolar milieu (Eriksson et al., 1988), and these findings are corroborated by the individual fluorescence spectrum of this Trp residue (Mårtensson et al., 1995).

The AEDANS adducts that emitted with longest wavelengths were the labeled S56C, F176C, W245C, and I256C mutants, whose red-shifted spectra must reflect a polar environment. Ser56 and Trp245 are much exposed to the solvent (Table 3), and Phe176 and I256 belong to the peripherally located residues, even though their side chains are almost completely buried (Table 3). Thus, the dansyl moiety, which resides on a fairly long side chain (Figure 2B), is likely to point toward the surrounding solvent in these positions.

Mapping the Intermediate State. The positions used to characterize the structure of the equilibrium intermediate obtained during denaturation are shown in Figure 1. These positions, indicated by the mutants used, are discussed below in an attempt to describe what parts of the protein might retain a native-like structure and what parts are structurally ruptured. In this context, the relative stabilities of various substructures are also compared.

A general observation is that most of the mutations produced have only a minor effect on the stability of the intermediate state, as compared to the unfolded state. The native state, however, is more sensitive to these substitutions (Table 1), which likely reflects the fact that the replaced amino acid side chains are involved in specific tertiary interactions in the native conformation of the enzyme. These tertiary interactions are then lost in the intermediate state in most of the mutants.

Flexible Parts of the Intermediate: S56C and I256C. Both of these positions are located peripherally in the enzyme molecule. Ser56 is situated in β-strand 2, and Ile256 is located on the opposite rim of the β-core, in the C-terminal β-strand 9 (Figure 1). Chemical reactivity data have earlier indicated that these outer strands of the β-sheet are less stable in the intermediate than those located deeper in the core (Mårtensson et al., 1993).

Both position 56 and 256 could be labeled with IAEDANS. The fluorescence spectra emanating from the attached dansyl group in position 56 exhibit an interesting behavior: in the native state, the emission maximum, together with that of the AEDANS-F176C derivative, is the most red-shifted of all studied AEDANS-labeled mutants (Figure 3D), and during the N → I transition, the spectrum is blue-shifted, *i.e.*, the label is sensing a more apolar environment in the intermediate than in the native state. Furthermore, the fluorophore seems to be less quenched in the intermediate than in the native state, since the intensity of the fluorescence increases during the N → I transition. Thus, the fluorophore seems to become buried in an apolar environment during this transition. One plausible explanation for these effects is that, due to the peripheral location of position 56 and the relatively long attachment arm of the dansyl group (Figure 2B), the fluorophore is pointing toward the surface and is in contact with the solvent in the native state. In the native state, the protein structure can probably not be adjusted to accommodate the bulky reporter group. Further evidence for this is that the native conformation of the S56C mutant is only marginally destabilized by AEDANS labeling (C_{mNI}, 0.6 and 0.4 M GuHCl, respectively). When the structure becomes more flexible, during formation of the molten globule intermediate, the AEDANS moiety swings into the interior, where the environment is less polar.

For the AEDANS-I256C derivative most of the spectral change observed from the fluorophore accompanied the N → I transition, which corroborates the earlier report that the C-terminal region is exposed to a high degree in the folding intermediate (Mårtensson et al., 1993).

The N → I transition is not registered in the EPR spectra of the spin-labeled S56C mutant, probably because the spin-label is relatively free to rotate already in the native state. In the GuHCl region between the two unfolding transitions some loosening of the substructure around position 56 seems to occur, as indicated by a gradual increase of the mobility of the attached spin-label (Figure 5C).

It was impossible to reactivate the denatured spin-labeled I256C mutant, but it is perhaps not surprising that labeling of Cys256 in the C-terminal β-strand 9 interferes with the folding process, since this strand is involved in an unusual knot topology. Introduction of AEDANS into this position was possible, probably because its longer attachment arm allows partial location of the probe on the outside of the protein, indicated by the relatively high wavelength maximum of the emission spectrum (Figure 3D).

F176C. Phe176 is also located near the surface close to β-strand 1 (Figure 1) and is almost completely buried because its side chain points toward the interior of the molecule (Table 3). The chemical labeling data for position 176 reveals two well-defined transitions (Figure 3B) that occur

at similar denaturing strengths as the N → I and I → U transitions observed by UV measurements (Table 1). Thus, it seems reasonable to believe that when the protein molecule undergoes the transition from the native to the intermediate state, this is accompanied by a corresponding loosening of the substructure close to position 176. That the magnitude of the N → I transition is large when measured by the relative alkylation rate, which is similar to that previously noted for position 56 (Mårtensson et al., 1993), also indicates that this substructure is relatively flexible in the intermediate state.

The N → I transition is also accompanied by a pronounced increase in mobility of the attached spin-label at this position, while essentially no change occurs during the I → U transition (Figure 5D).

The fluorescence spectra from the attached AEDANS label registered during various stages of the unfolding process behave in a way that is very similar to what was observed and discussed for the S56C-labeled mutant (Figure 3D). Therefore, it is likely that the dansyl group also in this peripheral position undergoes a movement from the outside to the interior of the protein molecule during the N → I transition.

During the I → U transition, the fluorescent label appears to be fully exposed to the solvent. Consequently, all data indicate that the substructure around position 176 is totally ruptured during the I → U transition.

Compact Part of the Intermediate: W97C. Trp97 is found in β-strand 4 and is included in a hydrophobic cluster which consists of 32 amino acid residues, eight of which are aromatic. This β-strand is also involved in a consecutively folded antiparallel β-structure (β-strands 2−6; Figure 1). The chemical labeling data indicated that the substructure around position 97 is inaccessible to external reagents, as demonstrated by the alkylation experiments shown in Figure 3A. This position remains buried in up to 4 M GuHCl, and it appears as if this substructure is not totally ruptured, even in 8 M GuHCl, the same behavior as previously noted for the area around position 68 in the neighboring β-strand 3 (Mårtensson et al., 1993).

From the EPR spectra and associated difference spectra recorded in the range 0.5−1.0 M GuHCl (Figure 5A,B), it is also evident that the introduced spin-label is relatively immobilized in the folding intermediate. The first transition monitored by EPR (Figure 3E) is of a relatively low magnitude, which indicates that the spin-label is only partially mobilized during the N → I transition. The appearance of the difference spectra amplitude also between 2 and 3 M GuHCl might indicate that the spin-label is not fully mobilized during the second transition (I → U) around 2 M GuHCl (Figures 3E and 5B). Furthermore, the incorporated dansyl group at position 97 senses a polarity change that accompanies the N → I transition when the protein conformation is transformed to the intermediate state. Judging from the observed emission red shift, the fluorophore, with its long attachment arm, reaches a less apolar environment in the folding intermediate than in the native state. This might be expected since calorimetric studies have indicated that water molecules penetrate a substantial part of the molten globule structure (Ptitsyn, 1992), but the fluorescence maximum of the position 97 AEDANS-labeled mutant indicates that its environment in the folding intermediate is among the most apolar of the examined substructures (Figure 3C,D). Furthermore, the difference spectrum showing the individual Trp97 fluorescence spectrum shows that this residue is still located in a very apolar surrounding in the folding intermediate (Mårtensson et al., 1994). A pronounced red shift concomitant with the second transition (I → U) is also noted for the emission maximum of the dansyl group, and even a third transition might be present between 3 and 5 M GuHCl.

Together, all these data give a picture of a very stable and compact substructure in the vicinity of position 97. Since far-UV CD spectra demonstrate that the β-structure of the intermediate state is substantially intact (Mårtensson et al., 1993), the substructure around position 97 thus appears to be native-like in the folding intermediate. It also exhibits residual structure under strong denaturing conditions.

For the W97C mutant, it is also clearly illustrated that the different probes give information about different aspects of the folding process. For instance, around 2 M GuHCl the spin-label gradually becomes more mobilized and the AEDANS probe is transferred to a more polar environment, while the engineered SH-group is not rendered more accessible to iodoacetate. Hence, the increased mobility and polarity sensed by the reporter groups most likely reflect the conformational changes which occur when most of the folding intermediate unfolds, leading to a less compact and a more polar surrounding for the outer parts of the attached spin-label and fluorophore and thus increased possibilities to rotate and to come into contact with water, whereas the central core including the engineered SH-group of position 97 is not substantially unfolded.

L118C. Leu118 is completely buried in the enzyme molecule (Table 3). This residue is contained in the N-terminal part of β-strand 5 (Figure 1), and its side chain is part of the large hydrophobic cluster. The chemical labeling experiments show that this position is remarkably inaccessible to alkylation in the intermediate state (Figure 3A). Thus, the folding intermediate is compact and seems to be structurally largely intact also in the region surrounding position 118. In addition, this substructure is very stable and does not start to unfold until 5 M GuHCl is used. Similar to the positions 68 and 97, in the neighboring β-strands, position 118 is not completely exposed even at about 8 M GuHCl. The absolute alkylation rate under these extremely strong denaturing conditions is by far the lowest noted (510 cpm $nmol^{-1}$ h^{-1} in 7.7 M GuHCl), also indicating that residual structure is preserved in this region. The L118C mutant was labeled with the spin-label and the fluorescent probe in the denatured state, but was then impossible to renature. The same observation was made for position 68, whose surrounding substructure previously has been shown to be extremely stable and compact (Mårtensson et al., 1993). Together this indicates that the hydrophobic core which comprises positions 68 and 118 is very rigid and densely packed and cannot adjust to accommodate a bulky group of the size of the probes.

W123C. Trp123 is situated fairly close to the surface and, like Leu118, is located in β-strand 5, although in the C-terminal part (Figure 1). From the results of previous chemical labeling experiments it was suggested that the region surrounding position 123 was relatively rigid (Mårtensson et al., 1993).

The folding intermediate also appears to retain its compact structure in the substructure around position 123, as evidenced by both EPR and fluorescence measurements of the labeled W123C mutants. Thus, no significant spectral

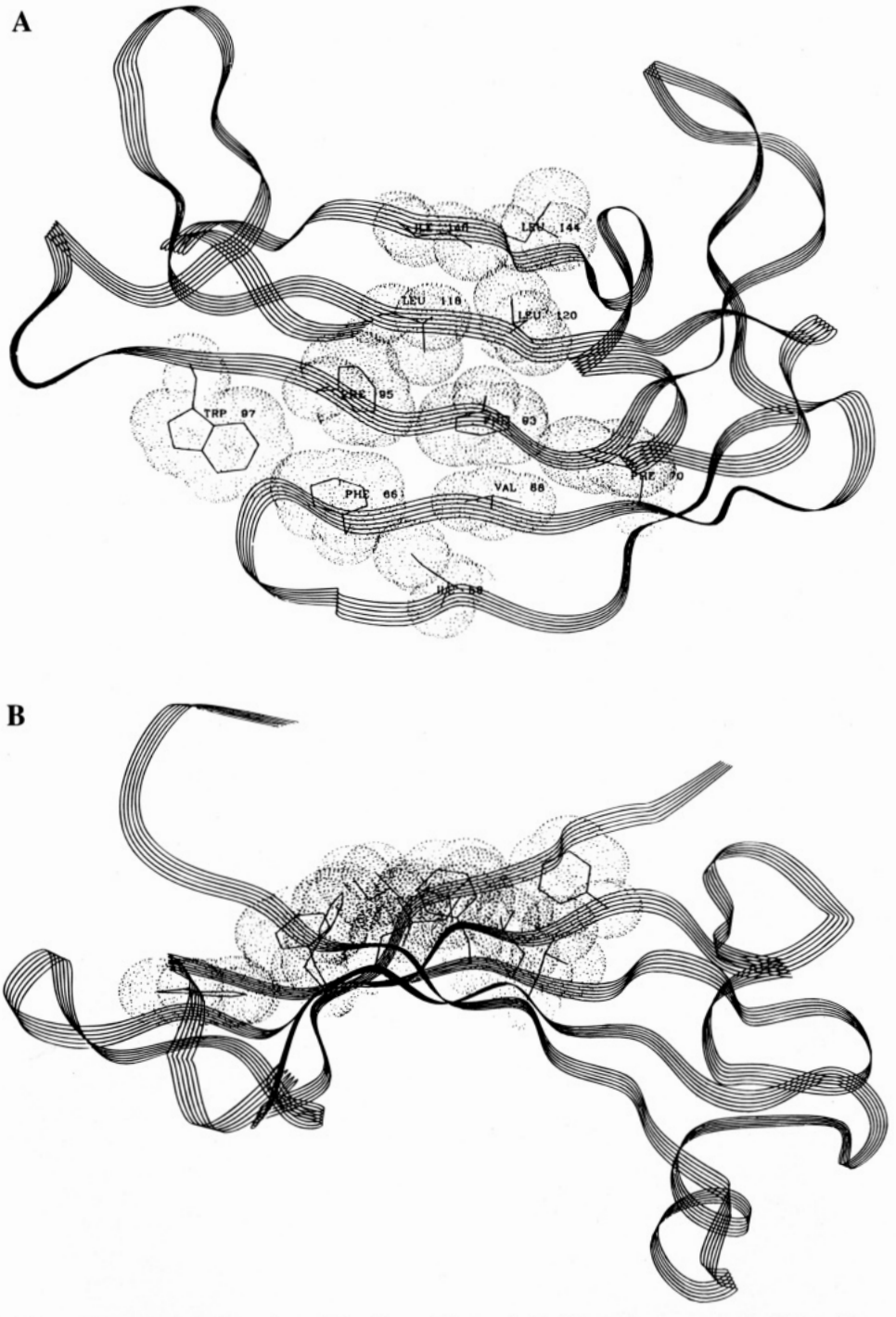

FIGURE 7: Panel A: View of the most stable region of the β-core (β-strands 2−6 from bottom to top) with residues within 5 Å of V68, W97, and L118 indicated. Panel B: The same region but viewed from another angle (side view of panel A) to show the clustering of the apolar residues. The figures were made using the Insight II program from BIOSYM Inc.

changes are observed during the N → I transition, neither for the spin-labeled (Figure 3E) nor for the AEDANS-labeled (Figure 4C) enzyme variants. The EPR spectrum is typical for a comparatively immobilized nitroxide radical, which also indicates that this substructure is rather rigid in the intermediate. The environment of the fluorophore is relatively polar in the native state probably due to partial exposure to the solvent, and the polarity is not increased as a result of the N → I transition. During the I → U transition the environment of the AEDANS moiety in position 123 becomes successively more polar, and this situation does not seem to change significantly after this transition has been completed. The mobility of the spin-label also increases during the I → U transition, but this increase does not cease until 3 M GuHCl is reached. This indicates that position 123 is rather exposed to the solvent, but that some residual structure persists in this area even after the I → U transition has occurred, in accordance with previous chemical labeling results (Mårtensson et al., 1993).

C206. Cys206 in the wild type is located at the edge of β-strand 7 (Figure 1). This strand and the vicinal 6th strand represent the most hydrophobic parts of the molecule (Bergenhem et al., 1989). The region surrounding position 206 forms a relatively thin wall between an apolar part of the active site cavity and the outside surface of the protein. The spin-probe introduced in position 206 undergoes only a

partial mobilization during the N → I transition (Figure 3F). Thus, the environment around position 206 seems to retain rather a compact structure in the intermediate state. Evidence supporting this was obtained from previous chemical accessibility data (Mårtensson et al., 1993) as well as from the blue-shifted fluorescence spectrum of Trp209, a residue situated in the same β-strand as Cys206 (Mårtensson et al., 1995). Nevertheless, this substructure appears to have some flexibility in the folding intermediate, since the AEDANS label in this position behaves like the label in positions 56 and 176, *i.e.*, it moves from relatively polar to less polar surroundings when the structure is loosened (Figure 3D). After the I → U transition no residual structure seems to exist, as evidenced by the spin-label and the fluorescent label data.

W245C. Trp245 is situated in a long loop on the surface of the enzyme connecting the C-terminally located β-strand 9 with an α-helical segment (residues 220−226) on the opposite side of the molecule. Despite the location on the surface, the immediate vicinity of position 245 has crystallographic *B*-values below 5 Å^2 and is situated close to β-strand 4 (Eriksson et al., 1988; the $C_\alpha-C_\alpha$ distance between Trp245 and Trp97 in the C-terminus of β-strand 4 is 6.7 Å), indicating that this part of the structure is involved in the underlying rigid structure comprising the hydrophobic core (Figure 1). Trp245 is one of the most exposed of the amino acid residues engineered in this study (Table 3). However, the inserted Cys245 is rather inaccessible to alkylation in the folding intermediate formed during the N → I transition and is not rendered significantly accessible until the I → U transition (Figure 3B), indicating that, in the intermediate, the loop containing position 245 is held together with the rest of the hydrophobic core. Above 4.5 M GuHCl, a third transition is noted; thus position 245 probably does not become fully accessible until the residual structure of the vicinal hydrophobic cluster begins to loosen (as probed by the V68C, W97C, and L118C mutants).

According to the EPR spectrum, the spin-label attached to position 245 is rather immobile in the native state of the molecule (Figure 4). Moreover, there are two populations of the spin-probe; the minor of these gives rise to a distinct three-line pattern, very similar to the pattern observed for the unfolded structure (abundance ca. 3.5%, as estimated from line shape simulations). The major population exhibits restricted motion due to a rigid position in the protein structure, and it therefore yields a component with broad bands. Upon reaching the intermediate state, the former population increases gradually and becomes dominant at GuHCl concentrations of 1 M and above (Figure 3F). The minor enhancement of the alkylation rate in the N → I transition shows only minor conformational changes around the main chain and the side chain out to the β-carbon, while the spin-probe indicates that there are larger changes closer to the surface, since the probe is almost as mobile as in the unfolded state.

The fluorescence spectrum of the dansyl group on Cys245 shows that the probe is in a relatively polar environment in the native state, indicating that the probe is probably positioned toward the surface of the protein. The fluorescence shifts that are observed during the N → I and I → U transitions are relatively small (Figure 3D), which can be expected if the fluorophore is substantially exposed already in the native conformation.

Concluding Remarks. The principal objectives of the present study were to investigate the compactness and stability of the existing folding intermediate, and to determine whether native-like structure might persist in the intermediate. It can be concluded that, in the vicinity of positions 68, 97, 118, 123, 206, and 245 (Figure 1), the folding intermediate is compact, and since far-UV CD spectra indicate that the folding intermediate retains much β-structure (Mårtensson et al., 1993), it probably contains some native-like structure elements. All of these positions, except no. 245, are located in the predominant β-structure of the protein, which consists of a 10-stranded, twisted β-sheet. The mentioned positions are found in β-strands 3, 4, 5, and 7, implying that it is the interior of the β-structure (*i.e.*, β-strands 3−7) that appears to retain a native-like fold (Figure 7). β-Strands 3−5 are part of a hydrophobic cluster in the core of the molecule. The substructure close to position 245 is rather rigid, indicating that in the intermediate the loop containing this position is held tightly together with the underlying hydrophobic core. Furthermore, β-strands 6 and 7 represent the most hydrophobic part of the molecule, emphasizing the importance of hydrophobic interaction in stabilizing this structure.

On the other hand, a more flexible structure in the folding intermediate is found around positions 56, 176, and 256 (Figure 1), which are all located in peripheral β-strands (nos. 1, 2, and 9). Thus, the stability of the secondary structure in the intermediate state is less in the outer parts of the protein than in the central parts.

After the I → U transition a completely unfolded (random coil) structure is not formed; *i.e.*, residual structure is retained in some parts of the molecule. The hydrophobic region containing β-strands 3−5, as probed at positions 68, 97, and 118, appears to be astonishingly stable and is not totally ruptured until extremely strong denaturing conditions are applied (Figure 7). The environments around positions 123 and 245 also contain some residual structure after the I → U transition. These positions are located closer to the surface than positions 68, 97, and 118. Residue 123 is situated in the C-terminal part of β-strand 5, and residue 245 is in the loop outside of β-strands 3−5. Thus, the stability of this β-sheet appears to increase toward the center of the core. The observed stable region is contained in the middle of the sequentially continuous antiparallel structure that spans β-strands 2−6 (Figure 7), which might represent an initiation site of folding.

ACKNOWLEDGMENT

We thank Dr. Nils Bergenhem, University of Michigan, Ann Arbor, for valuable discussion of the manuscript, and Katarina Wallgren for excellent technical assistance.

REFERENCES

Anfinsen, C. B., & Scheraga, H. A. (1975) *Adv. Protein Chem. 29*, 205−300.

Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 84−91.

Ballery, N., Minard, P., Desmadril, M., Betton, J.-M., Perahia, D., Mouawad, L., Hall, L., & Yon, J. M. (1990) *Protein Eng. 3*, 199−204.

Bergenhem, N., Carlsson, U., & Karlsson, J.-Å. (1989) *Int. J. Pept. Protein Res. 33*, 140−145.

Birkett, D. J., Dwek, R. A., Radda, G. K. Richards, R. E., & Salmon, A. G. (1971) *Eur. J. Biochem. 20*, 494−508.

Calciano, L. J., Escobar, W. A., Millhauser, G. L., Miick, S. M., Rubaloff, J., Todd, P., & Fink, A. (1993) *Biochemistry 32*, 5644−5649.
Carlsson, U., Aasa, R., Hendersson, L. E., Jonsson, B.-H., & Lindskog, S. (1975) *Eur. J. Biochem. 52*, 25−36.
Dolgikh, D. A., Kolomiets, A. P., Bolotina, I. A., & Ptitsyn, O. B. (1984) *FEBS Lett. 165*, 88−92.
Edsall, J. T., Mehta, S., Myers, D. V., & Armstrong, J. McD. (1966) *Biochem. Z. 345*, 9−36.
Eriksson, A. E., Jones, T. A., & Liljas, A. (1988) *Proteins: Struct., Funct., Genet. 4*, 274−282.
Freskgård, P.-O., Mårtensson, L.-G., Jonasson, P., Jonsson, B.-H., & Carlsson, U. (1994) *Biochemistry 33*, 14281−14288.
Garvey, E. P., & Matthews, C. R. (1989) *Biochemistry 28*, 2083−2093.
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, D. B., Lottich, S. C., Stein, P. J., & Williams, T. J. (1982) *Biochemistry 21*, 5918−5925.
Jennings, P. A., & Wright, P. E. (1993) *Science 262*, 892−896.
Julienne, M., & Garel, J.-R. (1981) *Biochemistry 20*, 7021−7026.
Khalifah, R. G., Strader, D. J., Bryant, S. H., & Gibbons, S. M. (1977) *Biochemistry 28*, 4914−4922.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87−103.
Lindgren, M., Svensson, M., Freskgård, P.-O., Carlsson, U., Jonsson, B.-H., Mårtensson, L.-G., & Jonasson, P. (1993) *Perkin Trans. 2 11*, 2003−2007.
Lindgren, M., Svensson, M., Freskgård, P.-O., Carlsson, U., Jonasson, P., Mårtensson, L.-G., & Jonsson, B.-H. (1995) *Biophys. J.* (in press).
Liu, J., Rutz, J. M., Klebba, P. E., & Feix, J. B. (1994) *Biochemistry 33*, 13274−13283.
Mårtensson, L.-G., Jonasson, P., Freskgård, P.-O., Svensson, M., Carlsson, U., & Jonsson, B.-H. (1995) *Biochemistry 34*, 1011−1021.
Mårtensson, L.-G., Jonsson, B.-H., Freskgård, P.-O., Kihlgren, A., Svensson, M., & Carlsson, U. (1993) *Biochemistry 32*, 224−231.
Mårtensson, L.-G., Jonsson, B.-H., Andersson, M., Kihlgren, A., Bergenhem, N., & Carlsson, U. (1992) *Biochim. Biophys. Acta 1118*, 179−186.
Murry-Brelier, A., & Goldberg, M. E. (1989) *Proteins: Struct., Funct., Genet. 6*, 395−404.
Neri, D., Billeter, M., Wider, G., & Wuthrich, K. (1992) *Science 257*, 1559−1563.
Nozaki, Y. (1972) *Methods Enzymol. 26*, 43−50.
Nyman, P. O., & Lindskog, S. (1964) *Biochem. Biophys. Acta 85*, 141−151.
Ohgushi, M., & Wada, A. (1983) *FEBS Lett. 164*, 21−24.
Oliveberg, M., Vuilleumier, S., & Fersht, A. R. (1994) *Biochemistry 33*, 8826−8832.
Ptitsyn, O. B. (1992) in *Protein folding* (Creighton, T. E., Ed.) pp 243−300, W. H. Freeman and Company, New York.
Ptitsyn, O. B. (1995) *Curr. Opin. Struct. Biol. 5*, 74−78.
Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20−24.
Rickli, E. E., Ghazanfar, S. A. S., Gibbons, B. H., & Edsall, J. T. (1964) *J. Biol. Chem. 236*, 1065−1078.
Schneider, D. J., & Freed, J. H. (1989) in *Biological and Magnetic Resonance, Volume 8* (Berliner, L. J., & Reuben, J., Eds.) pp 1−76, Plenum Press, New York.
Shortle, D. (1993) *Curr. Opin. Struct. Biol. 3*, 66−74.
Todd, A. P., Cong, J., Levinthal, F., Levinthal, C., & Hubell, W. L. (1989) *Proteins: Struct., Funct., Genet. 6*, 294−305.

BI950049V