High-Resolution Probing of Local Conformational Changes in Proteins by the Use of Multiple Labeling: Unfolding and Self-Assembly of Human Carbonic Anhydrase II Monitored by Spin, Fluorescent, and Chemical Reactivity Probes

Per Hammarström,* Rikard Owenius,† Lars-Göran Mårtensson,* Uno Carlsson,* and Mikael Lindgren†
Departments of *Chemistry and †Chemical Physics, Linköping University, SE-581 83 Linköping, Sweden

ABSTRACT Two different spin labels, *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide (IPSL) and (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate (MTSSL), and two different fluorescent labels 5-((((2-iodoacetyl)amino)-ethyl)amino)naphtalene-1-sulfonic acid (IAEDANS) and 6-bromoacetyl-2-dimetylaminonaphtalene (BADAN), were attached to the introduced C79 in human carbonic anhydrase (HCA II) to probe local structural changes upon unfolding and aggregation. HCA II unfolds in a multi-step manner with an intermediate state populated between the native and unfolded states. The spin label IPSL and the fluorescent label IAEDANS reported on a substantial change in mobility and polarity at both unfolding transitions at a distance of 7.4–11.2 Å from the backbone of position 79. The shorter and less flexible labels BADAN and MTSSL revealed less pronounced spectroscopic changes in the native-to-intermediate transition, 6.6–9.0 Å from the backbone. At intermediate guanidine (Gu)-HCl concentrations the occurrence of soluble but irreversibly aggregated oligomeric protein was identified from refolding experiments. At ~1 M Gu-HCl the aggregation was found to be essentially complete. The size and structure of the aggregates could be varied by changing the protein concentration. EPR measurements and line-shape simulations together with fluorescence lifetime and anisotropy measurements provided a picture of the self-assembled protein as a disordered protein structure with a representation of both compact as well as dynamic and polar environments at the site of the molecular labels. This suggests that a partially folded intermediate of HCA II self-assembles by both local unfolding and intermolecular docking of the intermediates vicinal to position 79. The aggregates were determined to be 40–90 Å in diameter depending on the experimental conditions and spectroscopic technique used.

INTRODUCTION

Knowledge about protein structure is of fundamental importance, because the final three-dimensional structure is essential for the function of a protein. To understand the factors that stabilize the native state, it is of interest to be able to characterize the native conformation and intermediate states that can be stabilized during various stages of the folding process.

X-ray crystallography and NMR spectroscopy are the methods most often used to determine the structure of proteins. It is generally very difficult and time consuming to obtain crystals of a protein, especially if the protein is bound to a membrane or if a second protein or a surface is involved. It would, of course, be almost impossible to acquire crystals from conformationally heterogeneous proteins in unfolded states. Solution NMR spectroscopy can provide detailed structural information but can be used to analyze only proteins with relatively small molecular weights. In addition, circular dichroism (CD) measurements can register conformational changes, for example, those occurring during the folding process, but does so with relatively low structural resolution.

Recent advances in protein engineering have opened up exciting new possibilities for studies of protein structure, protein folding, interactions between proteins, and interactions between proteins and surfaces. With this approach, cysteines can be introduced at particular locations in the protein structure. Either the sulfhydryl groups of these Cys residues are used as handles to which spectroscopic labels can be attached (Likhtenshtein, 1993; Lindgren et al., 1993; Svensson et al., 1995; Hubbell et al., 1998), or the chemical reactivity of the sulfhydryl group per se constitutes the probe function (Ballery et al., 1990; Mårtensson et al., 1993). A spin label will provide dynamic information and report on changes in mobility, and a fluorescent label will monitor both mobility and polarity. The chemical reactivity approach measures accessibility and reveals compactness of various substructures. Site-directed spin- and fluorescent-labeling methods have been employed with great success in numerous studies of T4 lysozyme, resulting in a detailed picture of the protein structure, dynamics, and conformational transitions (Hubbell et al., 1998; Mansoor et al., 1999). These site-directed labeling techniques are particularly useful for proteins that cannot be crystallized or analyzed by NMR and for structural characterization of unfolded states. By combining these methods to gain complementary data, it is possible to obtain detailed infor-

Received for publication 27 March 2000 and in final form 5 March 2001.

P. Hammarström's present address: Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037.

M. Lindgren's present address: Division of Sensor Systems, Defence Research Establishment, P.O. Box 1165, SE-581 11 Linköping, Sweden.

Address reprint requests to Dr. Uno Carlsson, IFM-Department of Chemistry, Linköping University, SE-581 83 Linköping, Sweden. Tel.: 46-13281714; Fax: 46- 13281399; E-mail: ucn@ifm.liu.se.

0006-3495/01/06/2867/19 $2.00

mation on both local and global conformational changes associated with the folding process (Svensson et al., 1995). Using a set of different labels of each type should further improve structural resolution, because, in addition to probing changes in different physicochemical characters, the labels will map local conformational rearrangements that occur during folding at different distances from the attachment or target site. We recently applied such a multi-probe technique to study the local environment of the interface between two complexed proteins (tissue factor and factor VIIa) (Owenius et al., 1999) and found that different labels provided varying results in regard to the protein-protein interaction. This could be rationalized by variations in label characteristics such as size, polarity, and flexibility.

As a model protein for protein-folding studies, we used human carbonic anhydrase II (HCA II), which catalyzes the hydration of CO_2 as well as the hydrolysis of aromatic esters. HCA II consists of 259 residues folded into a primary β-sheet protein (Fig. 1 *A*). The central part of the protein comprises 10 β-strands that divide the molecule into two halves. The upper half contains the N-terminal subdomain and the active site cleft (as shown in Fig. 1 *A*), and the lower half contains a large stable hydrophobic cluster that includes the side chains from residues in β-strands 3–7 (Eriksson et al., 1988; Håkansson et al., 1992; Mårtensson et al., 1993; Svensson et al., 1995; Hammarström et al., 1997). During unfolding of HCA II, partially folded states become populated. It was suggested that a molten-globule intermediate between the native and unfolded state is observed during equilibrium unfolding in guanidine (Gu)-HCl (Mårtensson et al., 1993). A recent pyrene excimer study (Hammarström et al., 1999) revealed that a substantial amount of the protein is prone to irreversible aggregation at a concentration of ~2 M Gu-HCl and below, also at relatively low protein concentrations (<1 μM).

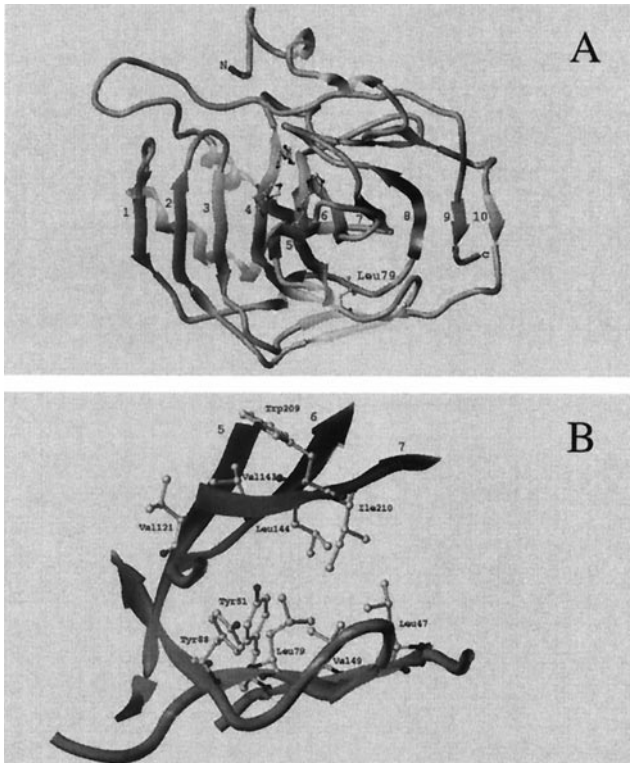

FIGURE 1 (*A*) Overall fold of HCA II with the 10 central β-strands indicated by numbers. The mutated residue Leu79 as well as the three Zn(II) ligands (His94, His96, and His119) in the active site are highlighted. (*B*) A close-up view of the region around position 79 with the local hydrophobic cluster residues indicated. Residues Val121, Val143, and Trp209 in the hydrophobic pocket of the active site are also shown.

In this study we employed the L79 position close to the hydrophobic core to investigate 1) the possible involvement in self-assembly at the aggregation interface, 2) the extension of the residual structure during various unfolding stages, and 3) the probing effect of similar but not identical spin and fluorescent labels.

In our previous EPR study of the unfolding of a spin-labeled HCA II mutant the experimental spectra were fitted with a linear combination of three simulated spectra, a slow component, an intermediate component, and a rapid component (Lindgren et al., 1995). A conspicuous broadening of the most slow component was noted for samples at intermediate Gu-HCl concentrations, and no definite conclusion about this broadening phenomenon could be made. Here it will be shown that the spectral appearance, as viewed from position 79, at intermediate Gu-HCl concentrations can be explained in terms of a soluble self-associated disordered state, and our objective is to gather more structural information about the aggregates. Investigation of protein-protein interactions leading to misfolded proteins in the form of amyloids or aggregates are of the highest importance. Protein aggregation is a major obstacle in biotechnology and biomedicine, as well as in studies in vitro focused on the mechanism of protein folding. During most of the protein-folding era, protein aggregation has been regarded as a mere nuisance that interferes with acquirement of native protein. However, the misfolding of proteins is now considered to be almost as interesting as the folding problem itself. This change in focus has been nourished by the discovery of molecular chaperones, which are known to be essential for the growth of bacteria (Fayet et al., 1989), and also by the development of recombinant protein technology, which often entails formation of inclusion bodies in the prokaryotic expression system. The major reason for the increase in research on protein aggregation was the discovery of the protein-folding diseases, such as Alzheimer's and Parkinson's diseases. It is of interest to map the structure and characteristics of soluble aggregates; for example, it was suggested that soluble forms of oligomeric A-β peptides in Alzheimer's disease can invoke a wide variety of toxic cellular responses (Lambert et al. 1998). In addition, it is not clear whether the amyloid fibrils or the precursors are the toxic particles involved in such kinds of diseases.

In addition, we aimed with this study to demonstrate the benefit of using a variety of labels in folding studies to achieve detailed mapping of a substructure of interest. To tune our detection system we evaluated the sensitivity of the labels to local environments by performing model experiments in various solvents. We also modeled the structure of the labels in extended conformations to complement the constrained structures found in the sterically restricted environments of a protein-protein interface. The effects of the labels on the protein structure were carefully analyzed in terms of biological activity and stability perturbation.

MATERIALS AND METHODS

Chemicals

Gu-HCl, sequential grade, was purchased from Pierce (Rockford, IL), and the concentration was determined by index of refraction (Nozaki, 1972). *N*-(1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidinyl)iodoacetamide (IPSL) was obtained from Sigma (St. Louis, MO); (1-oxyl-2, 2, 5, 5-tetramethylpyrroline-3-methyl) methanethiosulfonate (MTSSL) was obtained from Reanal (Budapest, Hungary); 5-((((2-iodoacetyl)amino)-ethyl)amino)naphtalene-1-sulfonic acid (IAEDANS) and 6-bromoacetyl-2-dimetylaminonaphtalene (BADAN) were purchased from Molecular Probes (Leiden, The Netherlands); 7-chloro-4-nitrobenzofurazan (NBD-Cl) was purchased from Fluka (Stockholm, Sweden); and iodo[2-^{14}C]acetic acid (54 mCi/mmol) was obtained from Amersham Sweden (Solna, Sweden).

Production and purification of mutated protein

Production and purification of the L79C HCA II mutant protein was performed as described previously (Mårtensson et al., 1993; Freskgård et al., 1994; Svensson et al., 1995). This mutant was made with the HCA II_{pwt}, C206S, protein as a template. The C206S mutation was made to remove the only naturally occurring cysteine and thereby produce a pseudo-wild-type protein where a cysteine can be inserted at any position in the sequence.

Activity assay

The specific activity for p-nitrophenyl acetate (pNPA) was determined by measuring the absorbance change at 348 nm, and the apparent second-order rate constant, k', was calculated as described earlier (Armstrong et al., 1966).

CD measurements

CD spectra were recorded in the near-UV region (240–320 nm) on a CD6 spectrodichrograph (Jobin-Yvon Instruments, Longjumeau, France). Samples with a protein concentration of 8.5 μM in 0.1 M Tris-H_2SO_4, pH 7.5, buffer were prepared for all enzyme variants.

Chemical labeling and radioactivity measurements

The chemical labeling approach used resembles the procedures reported previously (Mårtensson et al., 1993; Svensson et al., 1995). In brief, 1.5 ml of the L79C mutant protein (8.5 μM) was denatured for 24 h in various concentrations of Gu-HCl buffered with 0.1 M Tris-H_2SO_4, pH 7.5. Reaction with 24.4 μM iodo[2-^{14}C]acetic acid proceeded for 60 min and was then quenched by addition of 1.4 M β-mercaptoethanol. The protein was then purified on a Sephadex G-25 superfine gel-filtration column equilibrated with 2.2 M Gu-HCl. The protein concentration was determined by UV absorption measurements. A 1-ml vol of solution of labeled protein was added to 15 ml of scintillation cocktail (Aqua safe plus, Zinsser Analytic, Frankfurt, Germany). Measurements of radioactivity (cpm) were performed on a Beckman LS 1801 instrument.

Fluorescence labeling

Fluorescence labeling was performed as follows: 10 mg of protein (L79C mutant) was dissolved in 10 ml of 5 M Gu-HCl containing 0.1 M Tris-H_2SO_4, pH 7.5, for 30 min. A 10–20 times molar excess of reagent dissolved in 500 μl of 5 M Gu-HCl containing 0.1 M Tris-H_2SO_4, pH 7.5 (IAEDANS), or in 200 μl of DMSO (BADAN) was added to the denatured protein. The reaction was allowed to proceed for 20 h on a mechanical shaker in the dark for IAEDANS or 72 h for BADAN, as the latter reagent had been reported to react slowly (Haugland, 1996). The reaction was quenched with a 2 times molar excess of β-mercaptoethanol over reagent and the enzyme was refolded by dilution with 310 ml of 0.1 M Tris-H_2SO_4, pH 7.5, buffer to a final Gu-HCl concentration of 0.15 M and a protein concentration of 0.03 mg/ml. After 2–3 h of refolding, enzymatically active labeled protein was purified by affinity chromatography (Khalifah et al., 1977). The concentrations of the IAEDANS and BADAN adducts were determined by absorbance measurements, using $\epsilon_{337\ nm} = 6100\ M^{-1}\ cm^{-1}$ (Birkett et al., 1971) and $\epsilon_{380\ nm} = 21{,}000\ M^{-1}\ cm^{-1}$ (Haugland, 1996), respectively, and the protein concentration from $\epsilon_{280\ nm} = 54{,}800\ M^{-1}\ cm^{-1}$ (Nyman and Lindskog, 1964) after correction of the contribution of the label at 280 nm. The spectrophotometric analyses were recorded on a Hitachi U-2001 spectrophotometer.

Spin labeling

A representative spin labeling scheme is presented below: 15 mg of protein (L79C mutant) was dissolved in 10 ml of 5 M Gu-HCl containing 0.1 M Tris-H_2SO_4, pH 7.5, for 30 min. A 20 times molar excess of IPSL, dissolved in 500 μl of 5 M Gu-HCl containing 0.1 M Tris-H_2SO_4, pH 7.5, was added to the denatured protein. The reaction was allowed to proceed for 20 h on a mechanical shaker in the dark. A small aliquot of the protein solution was withdrawn to test for free cysteines with NBD-Cl as described previously (Svensson et al., 1995). The reaction was quenched with a 2 times molar excess of β-mercaptoethanol over reagent and the enzyme was refolded by dilution with 390 ml of 0.1 M Tris-H_2SO_4, pH 7.5, buffer to a final Gu-HCl concentration of 0.125 M and a protein concentration of 0.0375 mg/ml. After 3 h of refolding, enzymatically active labeled protein was purified by affinity chromatography (Khalifah et al., 1977). The same procedure was used for MTSSL labeling except for a few differences: a 10 times molar excess reagent was used (dissolved in 200 μl of DMSO) and the reaction was not quenched before refolding.

Steady-state fluorescence measurements

Fluorescence spectra from fluorescent-labeled protein were recorded on a Hitachi F-4500 spectrofluorophotometer equipped with a thermostatted cell. The measurements were carried out in the following way: IAEDANS-labeled protein (0.85 μM) or BADAN-labeled protein (1.0 μM) was incubated overnight in varying concentrations of Gu-HCl buffered by 0.1 M Tris-H_2SO_4, pH 7.5; a 1-cm cuvette was used. IAEDANS emission spectra were recorded in the wavelength region 380–600 nm, using 350 nm as excitation wavelength. Fluorescence emission spectra from BADAN-labeled protein were registered between 410 and 650 nm, using 390 nm as excitation wavelength, and 5-nm slits for both excitation and emission were used in all measurements.

The stability of the protein variants was measured by intrinsic Trp fluorescence. Fluorescence emission spectra were recorded in the range 310–450 nm after excitation at 295 nm. Both excitation and emission slits were set to 5 nm. The Trp emission maximum was evaluated by first-derivative calculation of the emission spectrum after subtraction of the appropriate blank. All fluorescence spectra were recorded at 23°C.

Fluorescence spectra from 1 μM solutions of mercaptoethanol adducts of IAEDANS and BADAN were also recorded in water, methanol, ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, 1-heptanol, 1-octanol, and 1-nonanol.

Anisotropy spectra were recorded (at 22°C) for both IAEDANS-L79C and BADAN-L79C at various Gu-HCl and protein concentrations as follows. Different protein samples (1 μM) were made, having a Gu-HCl concentration of 0, 1.0, and 3.0 M, corresponding to the native (N), intermediate (I), and unfolded (U) states, respectively. Protein samples in the range 1–21 μM were examined to vary the aggregate composition. All samples were incubated overnight. The polarization of the exciting and emitted light was changed between vertical (v) and horizontal (h) to obtain all four combinations. The emitted fluorescent light for IAEDANS was collected in the range 400–600 nm and for BADAN in the interval 400–650 nm. The fluorescence anisotropy (r_s) was calculated using the conventional formula:

$$r_s = \frac{I_{vv} - GI_{vh}}{I_{vv} + 2GI_{vh}},$$

where I_{vv} and I_{vh} are the intensities (i.e., the integral) of the emission spectra with the respective polarization configuration and G is an apparatus constant ($G = I_{hv}/I_{hh}$) taking into account the polarization dependence of the detection system. In a previous work (Owenius et al., 1999) we used $G = 1.22$ in anisotropy calculations for IAEDANS. In this study, however, we determined $G(\lambda)$ over the wavelength region of interest for both IAEDANS- and BADAN-labeled L79C. $G(\lambda)$ increased continuously with the wavelength, but the variation was small, with approximately 10% maximal deviation comparing the value of the central spectral position and the extremes on either side. Because the average is used to calculate the anisotropy (r_s) most of this effect will cancel if an average value from the center of the spectral interval is used. The average value of G was found to be 1.25 or 1.28 depending on whether IAEDANS- or BADAN-labeled samples were used (data not shown). The difference in G was due to the slightly different wavelength intervals used for the two fluorophores.

Time-resolved fluorescence measurements

Fluorescence lifetimes of the IAEDANS- and BADAN-labeled L79C mutant were measured using a miniature fluorescence lifetime analyzer, mini-τ (Edinburgh Instruments, Edinburgh, UK) equipped with a PDL 800-B pulsed diode laser (PicoQuant, Berlin, Germany). Measurements were made at 22°C, using 380-nm excitation pulses (10 MHz repetition rate with ~500-ps pulse width), while monitoring the emission through a FCG065 long-pass filter (>475 nm) (Melles Griot, Irvine, CA). The Brewster angle polarizers were set in magic-angle configuration; i.e., the analyzer polarizer was set 54.7° with respect to the excitation polarization. The emitted photons were detected with a Pentium-II PC-based SPC-300 single photon counter (v. 5.2) with a 12-bit ADC (Edinburgh Instruments/PicoQuant) and were collected in 4096 channels. Lifetime measurements were performed on the same samples that were also studied with fluorescence anisotropy (described above) using the same 1-cm cuvettes. The measurement time was 20 min. The instrument response function (IRF) was determined by measuring light scatter from a vesicle solution. In this way adjustment for the lifetime of the excitation could be made. Data were analyzed using the F900 (v. 5.13) analysis toolkit (Edinburgh Instruments). In the case of BADAN data, an n-exponential ($n = 1$–3) decay function was fitted to each experimental data set using a reconvolution fit method, because the measured lifetimes were in the same order of magnitude as the excitation lifetime. In these fits the maximum number of counts for the IRF and sample were included within the fit range (channels 765-4095). The order of the decay function was increased from $n = 1$ when the decay could not be fitted with a single exponential. For IAEDANS both reconvolution fits and tail fits yielded essentially the same results, because the lifetimes generally were longer in this case. The fit proceeded until χ^2 was minimized.

Knowing both the anisotropy, r_s, and lifetime, τ, of the fluorophore makes it possible to calculate the rotational correlation time, ϕ_c, from the Perrin equation:

$$r_s = \frac{r_0}{1 + \tau/\phi_c},$$

assuming r_0 to be 0.4 for both IAEDANS and BADAN.

Refolding experiments

Unfolding/refolding experiments were conducted in a similar manner as described previously (Hammarström et al., 1999). MTSSL-L79C and IPSL-L79C were incubated for 24 h in various concentrations of Gu-HCl (0–3 M). Protein concentrations were 8.5 μM. Refolding was initiated by dilution of the denatured enzyme solution to 0.3 M Gu-HCl and a final protein concentration of 0.85 μM. The Gu-HCl solutions were buffered with 0.1 M Tris-H_2SO_4, pH 7.5. The CO_2 hydration activity of the enzyme was measured after 3–5 h of refolding. The enzyme activity assay has previously been described (Freskgård et al., 1991).

EPR measurements

EPR spectra were recorded from 8.5 μM samples of IPSL-L79C or MTSSL-L79C at different Gu-HCl concentrations ranging from 0–5.0 M. Studies of the aggregated state were performed at 1.2 M Gu-HCl, in which the concentration of IPSL-L79C and MTSSL-L79C was varied in the range 5.0–20 μM. Details will be explained in the following sections.

EPR measurements were performed on a Bruker CW X-band EPR spectrometer composed of a combination of the ER200D-SRC and ESP300 systems. An ER4103TM cavity connected to the 200-mW microwave bridge was used, and the protein samples were introduced in a standard TM_{110} flat cell for aqueous samples (Wilmad Glass, Buena, NJ).

Line-shape distortions that could arise from experimental conditions, such as overmodulation or microwave saturation, were avoided by recording room temperature (20 ± 1°C) spectra using a modulation amplitude less than one-half the linewidth of the $m_I = 0$ nitrogen hyperfine transition and 4-mW microwave power.

Processing and simulations of experimental EPR data

All first-derivative EPR spectra were baseline corrected and normalized to a constant spin concentration (Bruker Win-EPR 2.11 and Matlab 5.2 or 5.3). The amplitude of the $m_I = 0$ nitrogen hyperfine transition of integrated EPR spectra obtained from spin-labeled L79C was used to determine unfolding curves. The central line was chosen because it has the highest signal-to-noise ratio, but any of the transitions in the spectrum can be used.

The published set of simulation programs developed by Freed and co-workers were used to analyze EPR spectra and estimate the dynamic parameters of the spin labels in the various protein conformations (Schneider and Freed, 1989). Our implementation of these programs has been described in detail previously (Owenius et al., 1999). The search for the model describing the local dynamic structure of the spin labels was considerably simplified by using isotropic hyperfine splitting and g values (A_{iso}, g_{iso}) obtained from measurements under conditions where the labels are tumbling freely on the EPR time scale, and anisotropic

parameters (A_{zz}, g_{zz}) from measurements of the labels in the rigid state. In contrast to our previous studies of various spin-labeled HCA II mutants using IPSL (Lindgren et al., 1993, 1995; Svensson et al., 1995) we here allow the magnetic parameters of IPSL and MTSSL to vary and be dependent on the properties of the surroundings at various stages as the protein unfolds and use them as starting values in the simulations. The changes attributable to solvent polarity are examined by inspection of key parameters for a series of solvents (data not shown). The parameters we here use for simulation (for IPSL) can therefore be somewhat different from those presented previously (Lindgren et al., 1995). Thus, when simulating spectra from unfolded L79C, parameters obtained from the spin labels in a polar solvent are used because the labels are believed to be solvent exposed. In parallel, parameters from the spin labels in an apolar solvent are used to simulate spectra from native L79C where the labels are expected to be in a hydrophobic environment.

Experimental EPR spectra from spin-labeled L79C were fitted with a linear combination of three simulated spectra. A unique slow and a rapid component here reflect the overall tumbling rate for the nitroxide moiety of the native and unfolded state, respectively, associated with the unique conformations of the protein at the extreme Gu-HCl concentrations. The aggregated state at intermediate concentrations could be analyzed in terms of a multi-component spectrum made up by a distribution of different rotational diffusion constants. More details about these conformations and the procedure to get the EPR parameters for particularly the aggregated state are described in Results and Discussion.

Computer evaluation of molecular structures

Structural information from HCA II was extracted using the coordinates from the protein data bank operated by Research Collaboratory Structural Bioinformatics (accession code 2cba; Håkansson et al., 1992). Computer modeling and studies of the various side chains were performed on a Silicon Graphics workstation using the Sybyl software from Tripos (SYBYL 6.3, Tripos, St. Louis, MO). Fully extended probing distances and lengths of the side chains were measured from optimized molecular models of each label attached to a cysteine (QUICK MINIMIZE routine in Sybyl), where the torsion angles of selected bonds were rotated to achieve extended conformations of the side chains. Volumes were determined from molecular surfaces (Connolly, 1983) obtained from these side-chain models. Lipophilicity potentials (LPs) describing the hydrophobic properties of the labels were determined from the same molecular surfaces according to Heiden et al. (1993).

RESULTS AND DISCUSSION

L79C mutant

We produced the L79C mutant by using the cysteine-free C206S mutant, designated HCA II_{pwt}, as a template in order to have a unique attachment site for various cysteine-specific labels (Mårtensson et al., 1993; Svensson et al., 1995). A molten globule-like intermediate of HCA II was found to form aggregates, in the form of misfolded soluble oligomeric particles, specifically with the region comprising the central β-strands 4–7 (Hammarström et al., 1999). This implies that HCA II could be a very useful model system for misfolding studies. It was noted that position 79 was not involved in the aggregation surface in that work. The aggregated structure appears, however, ubiquitous for all mutants at intermediate Gu-HCl concentrations. As can be judged from the crystal structure, L79 points inward toward the active site region and is deeply buried in the native state of the protein. L79 is situated in a peripheral β-strand connecting the inner β-strands 3 and 4 (Fig. 1). We address the question of whether the peripheral β-strand (residues 78–82) is also part of the aggregated structure, and to answer that query, we use several different molecular labels to study the L79C mutant.

Stability and global unfolding

The introduction of artificial side chains, such as molecular reporter groups, often induces structural strain in the protein, leading to a destabilized folded structure. Therefore, we recorded unfolding curves of the variants to investigate the impact of the labeling and the mutation on the stability of the protein. The unfolding process was monitored by the wavelength shift of the intrinsic tryptophan fluorescence (Mårtensson et al., 1995). The global unfolding of HCA II, as well as unfolding of the L79C mutant and the derivatives thereof, is a three-stage process that proceeds via formation of an intermediate structure (N→I→U). The stability of HCA II_{pwt} and of the L79C variants are summarized in Table 1. As can be seen from the midpoint concentrations of unfolding, L79C substitution resulted in only a minor destabilization of the native state, whereas labeling with the spectroscopic probes caused larger decreases in stability. The second unfolding transition (I→U) appears to be less affected by the chemical modification, probably because the tertiary interactions are already broken in the intermediate state (I). Interestingly, the stability of the protein variants was found to depend on the volume of the side chain in position 79 (Fig. 2). When labeled, that side chain became larger than the wild-type side chain (Table 2), which probably induced structural perturbation of close packing in the interior of the protein. That, in turn, may have affected tertiary interactions, leading to lower stability of the native state. Plots of parameters representing local structural changes in the vicinity of the spin labels, the fluorescent labels, and the chemical reactivity probe are shown in Fig. 3; each plot includes the global unfolding curve (measured as the intrinsic Trp fluorescence shift) of the same derivative as a reference.

Evidence of native-like enzyme

The esterase activity was measured of HCA II_{pwt}, the L79C mutant, and all the spin- and fluorescence-labeled derivatives. The second-order rate constants (k') for pNPA hydrolysis were determined and are presented in Table 1. Removing the leucine side chain and replacing it with the smaller cysteine side chain resulted in decreased enzyme activity as well as structural stability. Compared with HCA II_{pwt}, the fluorescent- and spin-labeled variants had almost the same enzyme activity but showed a further decrease in stability

TABLE 1 Global and local stability and enzyme activity of different HCA II variants

HCA II derivatives	Global unfolding*		Local unfolding†		Residual structure persistence, [GuHCl] (M)	Enzyme activity, k' (%)
	C_{mNI},‡ [GuHCl] (M)	C_{mIU},‡ [GuHCl] (M)	C_{mNI},‡ [GuHCl] (M)	C_{mIU},‡ [GuHCl] (M)		
HCA II_{pwt}	0.96	2.2				100§
L79C	0.82	2.1	0.90	None¶	2–7¶	71
L79C-IAEDANS	0.25	1.7	0.30	1.6	2.5–4	96
L79C-BADAN	0.55	1.7	None‖	1.9	3–4	102
L79C-IPSL	0.38	2.1	0.31	2.0	ND	96
L79C-MTSSL	0.50	2.2	0.95	2.1	ND	99

ND, not done.

*Determined from intrinsic Trp fluorescence measurements.

†Determined from the spectral change of the corresponding label in position 79 and from the rate of incorporation of iodo[2-^{14}C]acetic acid for unlabeled L79C.

‡C_{mNI} and C_{mIU} represent the transition midpoint concentrations (GuHCl) for the transitions from the native (N) to the intermediate (I) and from the intermediate (I) to the unfolded (U) state, respectively. All stability measurements were fitted to a three-state model and calculated as described previously (Mårtensson et al., 1993).

§Corresponding to a rate constant of 2023 M^{-1} s^{-1} for the esterase activity.

¶Decreased accessibility observed between 1.5 and 2.5 M GuHCl that remained up to 7 M GuHCl.

‖A blue-shift of the fluorescence was detected for L79C-BADAN with a minimum at 0.6 M GuHCl.

(as discussed above). In addition, near-UV CD measurements showed that all variants had intact tertiary structure (data not shown).

Obviously, removal of the buried side chain would cause structural changes that would propagate to the active-site region and thereby affect the enzyme activity. Interestingly, introducing a new hydrophobic side chain (i.e., a fluorescent label or a spin label) restored the lost activity of the mutant enzyme to the HCA II_{pwt} level. This implies that preservation of hydrophobicity in position 79 is important for enzyme activity. Fierke and co-workers were the first to increase side chain hydrophobicity in HCA II by using protein engineering to introduce hydrophobic residues in positions around the active site (Fierke et al., 1991; Alexander et al., 1991). In those studies, a hydrophobic pocket defined by the residues L198, W209, and, in particular, V121 and V143 were shown to be important for maintaining enzymatic activity. Residues V121 and V143 (Fig. 1 *B*) are in the vicinity of L79 (closest distance ~6 Å). Although the labels are larger than the leucine side chain it is possible that the restoration of enzymic activity is due to refilling of the cavity created by the Leu→Cys mutation (Table 2).

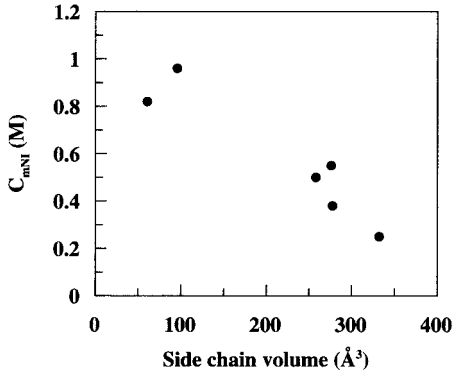

FIGURE 2 The dependence of the stability on the side chain volume for the L79C mutant and its labeled variants. The midpoint concentration of unfolding by Gu-HCl for the N→I transition is plotted versus the side-chain volume in position 79.

Characteristics of the different labels

In the native state, the local environment surrounding position 79 in HCA II should be rather fixed. This assumption from the crystal structure was verified by the spectroscopic data obtained from the spin and fluorescent labels (to be discussed more in the following sections). It is reasonable to assume that steric hindrance around the labels would lead to a constrained structure, limiting the number of possible

TABLE 2 Structural properties of the side chains

Side chain*	Probing distance†		Volume‡ (Å^3)
	Constrained (Å)	Extended (Å)	
Leucine			95.7
Cysteine	3.6	3.6	60.6
IAEDANS	8.7	11.2	331.8
BADAN	6.6	9.0	274.3
IPSL	7.4	10.0	277.9
MTSSL	7.1	8.8	255.5

*Either amino acid residue or label attached to a Cys residue.

†Distance from C_α to active probing part of the side chain, i.e. the center of the fluorophore for IAEDANS and BADAN and the nitroxide nitrogen for IPSL and MTSSL. Constrained distances were calculated from molecular models of labels buried at a protein-protein interface (Owenius et al., 1999). Extended distances were determined from optimized models of extended side chains, i.e. with no external restrictions (see Materials and Methods).

‡Obtained from optimized models of extended side chains (see Materials and Methods). Volumes were determined from surface models made with a probe radius of 1.4 Å (Fig. 4), corresponding to the surface accessible to a water molecule (Connolly, 1983).

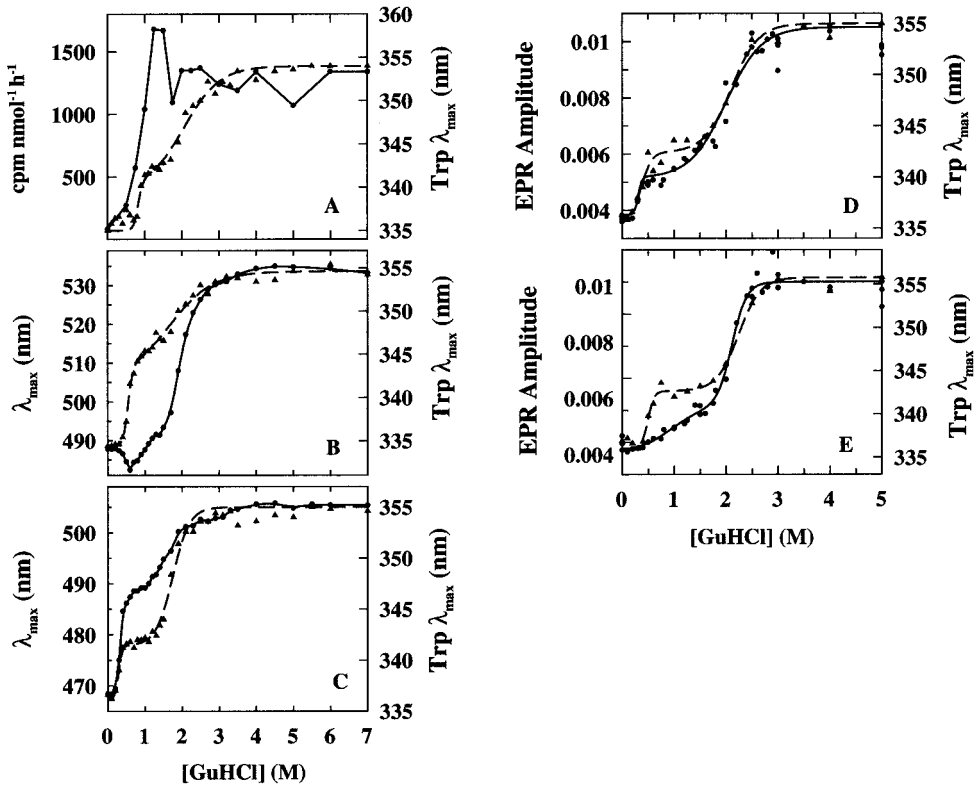

FIGURE 3 Local and global unfolding of L79C variants. (*A*) Incorporation of iodo[2-^{14}C]acetic acid; (*B*) BADAN-L79C fluorescence emission wavelength maximum; (*C*) IAEDANS-L79C fluorescence emission wavelength maximum; (*D*) IPSL-L79C EPR amplitude; (*E*) MTSSL-L79C EPR amplitude. Symbols represent parameters describing local unfolding (●, ——) and global unfolding measured by intrinsic tryptophan fluorescence (▲, — — —). In *A–C*, the solid line was drawn to guide the eye. In *D* and *E*, the solid line represents a fitted curve. All stability measurements were fitted to a three-state model and calculated as described previously (Mårtensson et al., 1993).

conformations of the label. Here we use the term constrained probing distance for such a label, employing the computer-modeled probing distances obtained for the labels in the interface between two proteins (Owenius et al., 1999); the constrained probing distances are given in Table 2. Furthermore, the local environments of the spectroscopic labels change continuously as the protein unfolds, which will make the labels less restricted by surrounding structures. Therefore, we also determined the length of the labels in a maximally extended conformation, as described in Materials and Methods. Surface contour plots of the side chains in extended conformations (Fig. 4) reveal differences in both structure and distribution of lipophilicity potential. Evidently, the moieties of this conformation also display large variation in geometry and distribution of polarity. The probing distance of the labels in the extended form is elongated by 1.7–2.6 Å compared with the constrained conformation. Thus, the spectroscopic labels appear to probe their environment at different distances from the backbone during various stages of the unfolding process.

Monitoring aggregation by refolding experiments

The unfolding/refolding experiments revealed similar results as reported previously for wild-type HCA II (Hammarström et al., 1999). The refolding yield of the spin-labeled variants was found to be 70% when the protein was denatured in high concentrations of Gu-HCl (3 M; Fig. 5). However, the refolding yields were significantly lower when refolding was performed on enzyme that was denatured in lower concentrations of Gu-HCl. The reactivation yields are mirrored by the two unfolding curves N→I and I→U, respectively, and form a refolding trough, implying that the amount of protein that can be reactivated decreases in parallel with the increase in formed I (molten globule) at both transitions (Fig. 5). As was shown previously, the width of the refolding trough depends on the protein concentration, which indicates that aggregation is the cause of the low recoveries of active enzyme upon refolding (Hammarström et al., 1999).

Probing accessibility of the sulfhydryl group during unfolding

The accessibility of the thiol group of the engineered Cys-79, was probed by chemical labeling with iodo[2-^{14}C]acetic acid, which is a rather small and highly amphiphilic molecule. In Fig. 3 *A*, the accessibility is plotted as the rate of incorporation of the radioactive label during 1 h of reaction

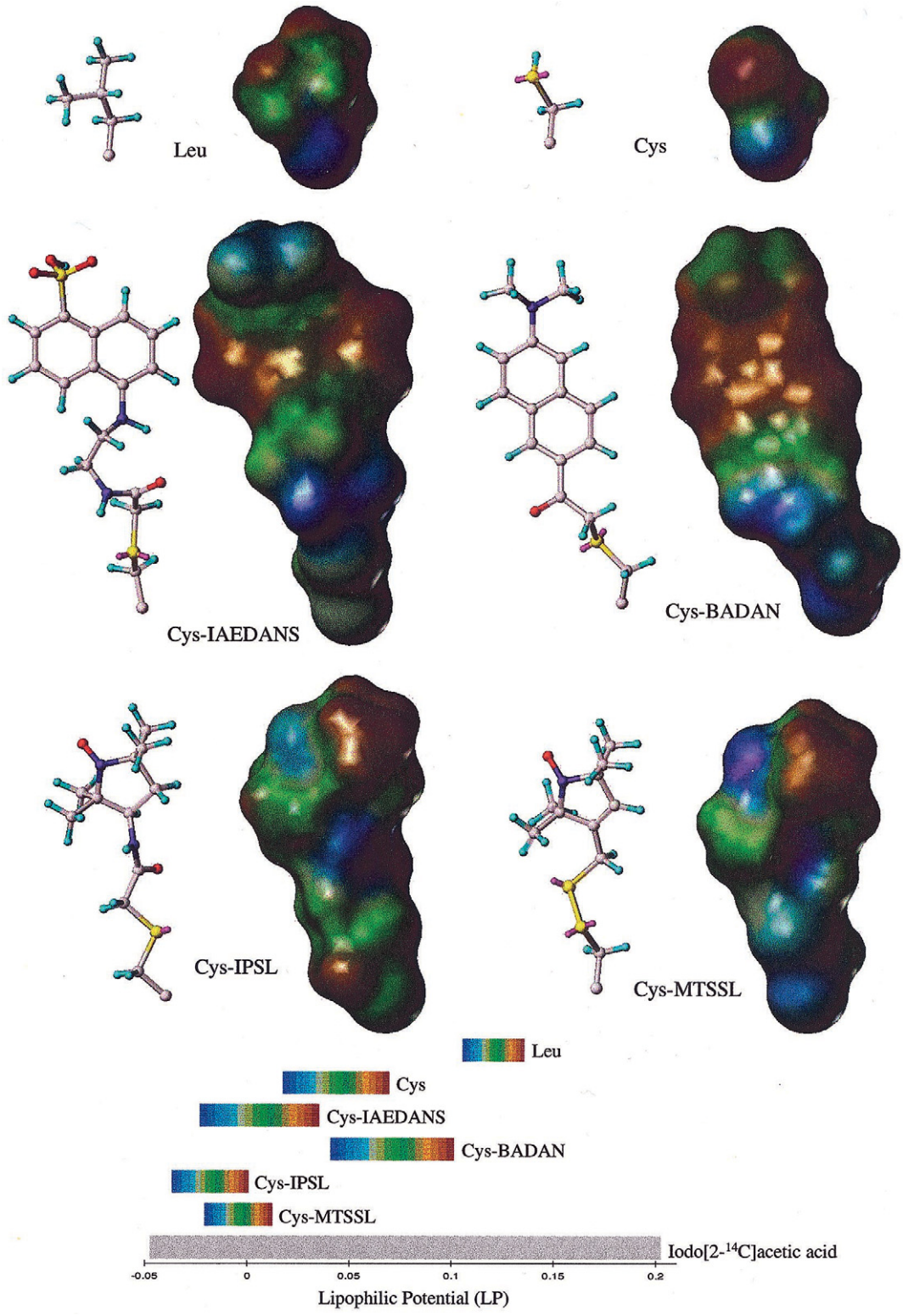

FIGURE 4 The different side chains illustrated both with ball-and-stick models and lipophilicity surface models. Each side chain is plotted from the backbone of the protein, i.e., beginning with the C_α, to be able to compare the size of the side chains. The side chains are color coded with different lipophilic potential (LP) scales and can therefore not be compared directly. Instead, the range in LP for each side chain is given in a separate graph at the bottom of the figure, where blue and brown represent the lowest and highest LP of the side chain, respectively. The extreme LP values for the chemical reactivity probe iodo[2-^{14}C]acetic acid is also included for comparison. The models are constructed as described in Table 2 using the Sybyl software from Tripos Inc. on a Silicon Graphics workstation.

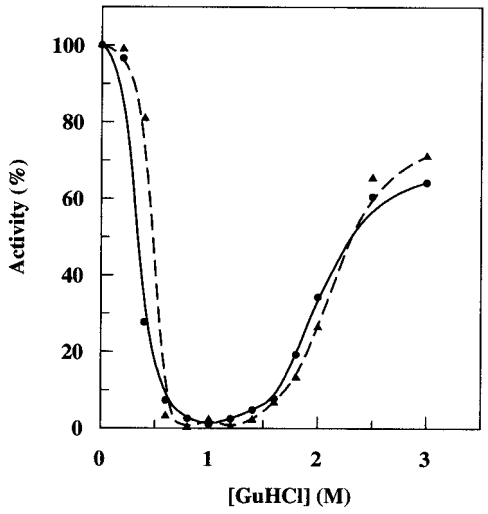

FIGURE 5 Refolding yield of spin-labeled L79C. Reactivation of the enzyme was performed after incubation of 8.5 μM protein for 24 h in various concentrations of Gu-HCl as indicated on the *x* axis. Refolding was induced by dilution of denatured enzyme to 0.2 M Gu-HCl in 0.1 M Tris-H_2SO_4, pH 7.5, and a final protein concentration of 0.85 μM. The enzyme activity was recorded after 3–5 h of refolding. Symbols represent IPSL-L79C (●, ——) and MTSSL-L79C (▲, – – –).

as a function of Gu-HCl concentration; as a reference, the global unfolding of L79C is shown in the same figure. In the native state, the Cys side chain is completely buried in the protein. Upon unfolding, we found that the thiol group became highly accessible to iodo[2-^{14}C]acetic acid after the first unfolding transition (native→molten globule), which is not very surprising, because residue 79 is located in a peripheral β-strand in the protein. Further unfolding showed that the rate of alkylation decreased in the interval 1.5–2.0 M Gu-HCl and reached a somewhat lower level than that attained between 1.2 and 1.4 M Gu-HCl. The decreased reactivity suggests that the side chain is shielded in the unfolded state of the protein, and such behavior has been observed in several HCA II single-cysteine mutants (unpublished results). It is also possible that the amphiphilic nature (Fig. 4) of the iodoacetate molecule will render the iodo part of the molecule to bind to the local hydrophobic cluster surrounding position 79 and thereby achieve maximum reactivity in the molten-globule intermediate. However, the maximal incorporation of iodoacetate at 1.2–1.4 M coincides with the maximal population of protein aggregates. Thus, these results indicate that the sulfhydryl moiety at 3.6 Å from the protein backbone in position 79 is not hidden in the aggregation interface.

Fluorescent labeling to probe local polarity during unfolding

Fluorecent labels, IAEDANS and BADAN, were attached to Cys-79 to monitor local changes in polarity. The degree of labeling was found to be 0.9 for IAEDANS and 0.7–1.0 for BADAN. For comparison, the fluorescence of β-mercaptoethanol adducts of IAEDANS and BADAN in water and in unbranched primary alcohols was measured to model the transition from the nonpolar protein interior to a polar aqueous phase. Fluorescence emitted by both fluorophores underwent red-shifts in the transition by changing the solvent from alcohols to water, due to the difference between the dielectric constants at 20°C, which is 80.4 for water and 33.6 for methanol, and gradually decreasing to 10.3 for octanol (Weast, 1983). Using these solvents, BADAN was sensitive to polarity changes over the entire region, with a red-shift of the fluorescence emission with increased polarity. The emission peak of IAEDANS was most blue-shifted using butanol, but interestingly, it was red-shifted when using both more hydrophobic and hydrophilic solvents (Fig. 6). Thus, it can be difficult to interpret fluorescence spectra in terms of polarity in hydrophobic environments when probing only with IAEDANS. This is a somewhat unexpected result, but it further emphasizes the importance of multiple probing of protein conformational changes.

Fig. 7, *A* and *B*, illustrates representative fluorescence emission spectra from BADAN-L79C and IAEDANS-L79C, respectively, in 0, 1.0, and 5.0 M Gu-HCl. For the native state (0 M Gu-HCl), the emission peak is found at 488 nm (Fig. 7 *A*; BADAN) and 467 nm (Fig. 7 *B*; IAEDANS). This is consistent with an environment equiv-

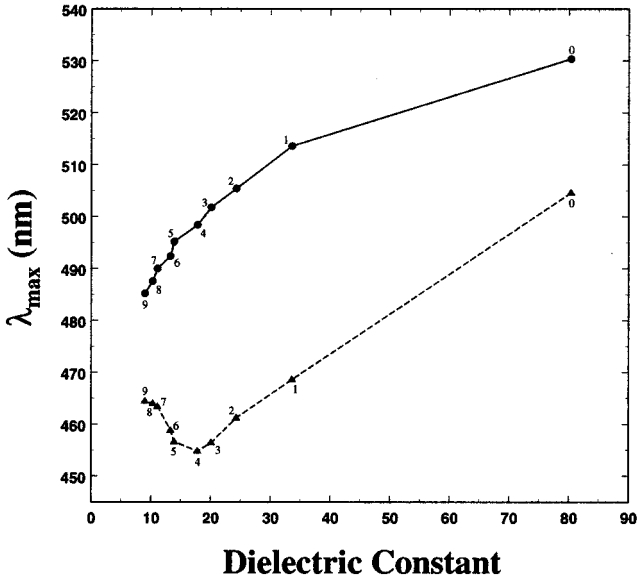

FIGURE 6 The influence of the dielectric constant on the fluorescence emission wavelength of model compounds of BADAN and IAEDANS. The model fluorophores (1 μM) were mixed in water, and 1-alcohols and fluorescence spectra were recorded as described in Materials and Methods. All samples were purged with N_2 gas before measurements to remove dissolved oxygen. Fluorescence shift of BADAN-mercaptoethanol (●, ——) and IAEDANS-mercaptoethanol (▲, – – –) were plotted versus solvent type. The numbers in the figure correspond to the solvents as follows: 0, water; 1, methanol; 2, ethanol; 3, 1-propanol; 4, 1-butanol; 5, 1-pentanol; 6, 1-hexanol; 7, 1-heptanol; 8, 1-octanol; and 9, 1-nonanol.

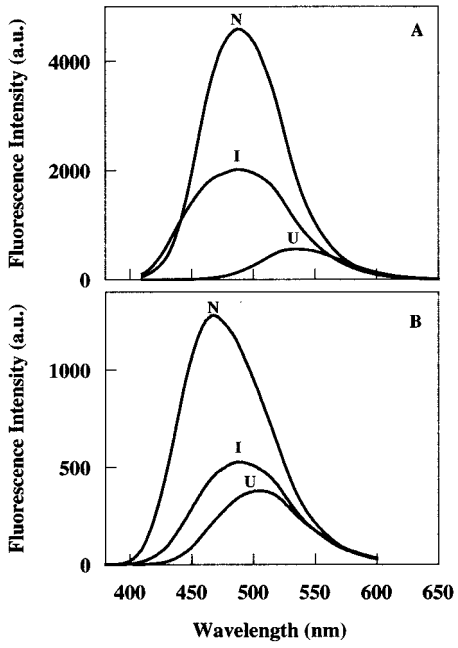

FIGURE 7 Fluorescence emission spectra from labeled L79C in various conformational states. (*A*) BADAN-L79C in 0 M (native), 1.0 M (intermediate), and 5.0 M Gu-HCl (unfolded); (*B*) IAEDANS-L79C in 0 M (native), 1.0 M (intermediate), and 5.0 M Gu-HCl (unfolded).

alent to octanol to nonanol, comparing the solvent dependence for the two labels (Fig. 6). For the intermediate state (1.0 M Gu-HCl) the spectrum of BADAN is slightly blue-shifted, with a tendency of broadening of the entire spectrum. For IAEDANS the spectrum is substantially red-shifted to 488 nm. The peak for the unfolded state is shifted to 535 nm for BADAN and to 505 nm in the case of IAEDANS. As expected, both fluorescent probes indicated a water environment for the unfolded form (Fig. 6).

In the unfolding curve shown in Fig. 3 *B*, the emission wavelength maximum of BADAN-L79C is plotted versus the Gu-HCl concentration. When unfolding was probed as a change in polarity by the BADAN moiety at a site further away from the backbone than the thiol group, which entailed probing the interior of the protein, a different picture emerged than that obtained by measuring thiol accessibility. The length of the BADAN side chain was found to be 6.6 Å in the constrained conformation, which means that the probed site was 3.0 Å from the thiol group in the chemical labeling measurements (Table 2). During the first global unfolding transition, the emission maximum of the label was blue shifted by 5 nm (Fig. 3 *B*), which indicates a more hydrophobic environment in the intermediate state. One explanation for this behavior is that, during the first global transition, the rigid tertiary structure was distorted, leading to exposed hydrophobic patches prone to aggregate with HCA molecules of the same structure. Apparently, this gave rise to an even more hydrophobic environment at distances 6.6–9.0 Å from the backbone of position 79. This structure is compact as will be shown by the results of other experiments below. Notably, the entire spectrum broadens at intermediate Gu-HCl concentrations, both toward the blue and red ends of the emission spectrum. This indicates that populations of BADAN labels experience different structures of quite varying polarity. Upon further unfolding, in 1.5–4.0 M Gu-HCl, there was a large red-shift in fluorescence, concomitant with the second global unfolding transition.

Fig. 3 *C* shows a similar unfolding curve for IAEDANS-L79C. IAEDANS and BADAN differed in regard to the behavior of the fluorescence shift seen upon unfolding of the protein. The second local unfolding transition probed by the IAEDANS label was concurrent with the second global unfolding of the intermediate state. The label also revealed a possible third unfolding transition in the interval 2.5–4.0 M Gu-HCl, indicating unfolding of local residual structure. The overall IAEDANS-L79C fluorescence behavior is similar to what was previously reported for IAEDANS-W97C, located in β-strand 4 (Svensson et al., 1995).

The fluorescence unfolding curves of BADAN-L79C and IAEDANS-L79C revealed transitions from a hydrophobic protein interior (equivalent to octanol-nonanol) to water. However, the labels had very different unfolding signatures, which can be attributed to the characteristics of the labels. Because the same type of covalent bond is formed between the protein and the different fluorophores, this discrepancy should depend mainly on the distance of the label from the protein backbone and the particular polar characteristic of the label that promoted solvation in either polar water or hydrophobic clusters. From the calculated distribution of lipophilicity of the two labels (Fig. 4), it is evident that BADAN is more hydrophobic than IAEDANS, which probably contributes to the different behavior of the two fluorophores (Owenius et al., 1999). In addition, the larger number of bonds in the IAEDANS linker, compared with the BADAN linker, would also render the IAEDANS label more flexible. Considering these characteristics, IAEDANS should prefer to be more exposed to the solvent. The spectral shifts of IAEDANS-L79C and BADAN-L79C accompanying the first global unfolding transition occurred at different Gu-HCl concentrations due to the different global stabilities of the native states of these protein variants (Table 1; Fig. 3 *B*). However, the different spectral responses (a blue- and a red-shift for BADAN-L79C and IAEDANS-L79C, respectively) should not be due to varying different global stabilities, but instead reflect different probing of the vicinal regions at the N→I transition.

Fluorescence life-time and anisotropy measurements

Well known from the Perrin equation the fluorescence anisotropy in combination with lifetime measurements can be used to determine molecular dynamics, for example, the motion of a fluorophore attached to a protein. In an attempt

to vary the size and structure of the protein aggregates, samples were made at different protein concentrations by adding unlabeled L79C to the solution. Thus, all samples had the same fluorophore concentrations.

The fluorescence anisotropy spectra for the N, I, and U states were measured at 0, 1.0, and 3.0 M Gu-HCl, respectively (data not shown). All spectra showed traces of sidebands due to vibrational substructure with a downward slope toward the low-energy end. This was most pronounced for the anisotropy spectra of the intermediate states. All anisotropy data are collected together with lifetime data and the calculated rotational correlation times in Table 3. Three different protein concentrations were examined for each Gu-HCl concentration and fluorescent label. For IAEDANS essentially the same curve was obtained for the N state, regardless of protein concentration. This was also the case for the U state, however, with an anisotropy close to zero, indicating that the IAEDANS label tumbled rapidly in the unfolded protein. The anisotropy of IAEDANS in the native protein is still quite low; however, part of this may be ascribed to the increased lifetime of the label in the excited state. This is verified indirectly by the increase in fluorescence quantum yield (Fig. 7 *B*) and it was also verified in the lifetime measurements that gave approximately double the lifetime for IAEDANS in the native compared with the unfolded state. The largest fluorescence anisotropy was obtained for IAEDANS in the intermediate state, considerably larger than in the native conformation. The lifetime associated with this state was, however, intermediate between that of the native and the unfolded structure. The associated rotational correlation times were calculated by using both the anisotropy and lifetime data. Interestingly, it was found to increase in the order of unfolded (1.9 ns), native (9.3 ns), intermediate (11.5 ns). Thus, the observed increase in anisotropy is due to more pronounced slower overall tumbling of the label in the intermediate state, i.e., consistent with aggregation at 1 M Gu-HCl (refolding experiments; Fig. 5). This was further supported by increasing the protein concentration to 21 μM where a difference in anisotropy was observed for IAEDANS only in the intermediate state, resulting in a dramatic increase of the rotational correlation time (ϕ_c = 17.7 ns). Traces of the fluorescence lifetime of IAEDANS in the intermediate state along with their fits are shown in Fig. 8 *A*.

Similar results as for IAEDANS-L79C were obtained for BADAN-L79C (Table 3; Fig. 8 *B*). It was evident also for BADAN-L79C that an increased protein concentration at 1 M Gu-HCl increased the fluorescence anisotropy. However, the short fluorescence lifetime of BADAN rendered the total fluorescence anisotropy to be much larger than for IAEDANS-L79C, and consequently it also decreased the sensitivity to reveal structural changes. Except for the native state, the lifetime traces of BADAN-L79C were not possible to analyze with a single exponential. This indicates, and as will be further supported by the results of EPR line-shape simulations, that several dynamic structures of the label may

TABLE 3 Results of fluorescence anisotropy and lifetime measurements

	IAEDANS			BADAN			
Sample*	r_s	τ (ns)	ϕ_c (ns)†	r_s	τ (ns)	ϕ_c (ns)†	
N-state	0.12	21.6	9.3	0.27	4.7	9.6	
I-state (1 μM)	0.18	14.8	11.5	0.28	1.4	3.1	26%
					3.6	8.2	74%
I-state (21 μM)	0.21	15.4	17.7	0.34	1.4	8.6	23%
					3.8	22.2	77%
U-state	0.05	11.9	1.9	0.24	0.5	0.8	24%
					1.5	2.2	59%
					3.1	4.7	17%

*N corresponds to 0 M GuHCl, I to 1.0 M GuHCl, and U to 3.0 M GuHCl.
†ϕ_c was calculated using r_0 = 0.4. For BADAN, weights are given for the multi-exponential fits.

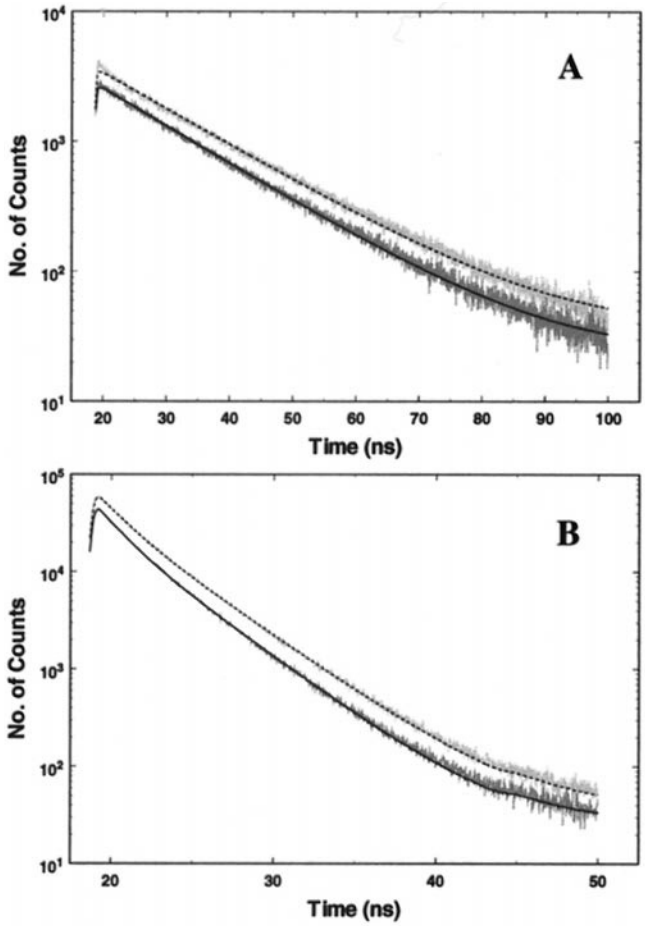

FIGURE 8 Fluorescence lifetime traces of IAEDANS-L79C (*A*) and BADAN-L79C (*B*) at 1 M Gu-HCl corresponding to the aggregated state. The lower trace in both figures corresponds to a protein concentration of 1 μM. The upper traces are obtained by adding 20 μM unlabeled protein to the solution (see Materials and Methods for more details). Solid and dashed curves are fits using the parameters in Table 3.

co-exist. Interestingly, the rotational correlation time is essentially identical for both the IAEDANS and the BADAN label in the native conformation. Similar values were obtained for the unfolded and intermediate conformations if we used a weighted average rotational diffusion constant for the ill-defined lifetime traces associated with BADAN in the unfolded and intermediate states.

The diameter of the protein calculated from the rotational correlation times of the native conformation is 41–42 Å for both the IAEDANS and BADAN cases, in agreement with earlier studies and the crystal structure of HCA II ($39 \times 42 \times 55$ Å^3; Eriksson et al., 1988). The increase of ϕ_c of the dominating contributions in the time traces associated with the aggregated state corresponds approximately to a doubling of the apparent molecular volume.

Spin labeling to probe local rigidity during unfolding

Over 95% labeling of the L79C mutant was achieved with both MTSSL and IPSL, as determined from NBD-Cl titration of free thiol groups. Representative EPR spectra of the folded to unfolded forms of MTSSL-L79C and IPSL-L79C are shown in Fig. 9, *A* and *B*, respectively. The mobility of the spin label is influenced by its interactions with neighboring side chains and solvent molecules. In the native state of the protein (0 M Gu-HCl), a spin label attached to position 79 will rotate slowly (or not at all) due to its compact hydrophobic environment. This means that the motion reflects the overall tumbling of the protein. In such a case, a spectrum will have broad lines, which is characteristic of a buried spin label (upper spectra in Fig. 9, *A* and *B*). When the Gu-HCl concentration is successively increased, the interaction of the label with the local environment is changed, because the protein starts to unfold. This causes a dramatic change in the EPR line-shape. Above ~2–3 M Gu-HCl, the spectrum changes into a characteristic three-line hyperfine pattern of a rapidly rotating spin label (lower spectra in Fig. 9, *A* and *B*). At intermediate Gu-HCl concentrations, several components of different mobility contribute to the spectrum (middle spectra in Fig. 9, *A* and *B*).

To conveniently compare the unfolding process as determined by EPR with the other techniques, the most narrow component of each EPR absorption spectrum was plotted versus the Gu-HCl concentration (Carlsson et al., 1975). Fig. 3 *D* shows the resulting curve for IPSL-L79C, and Fig.

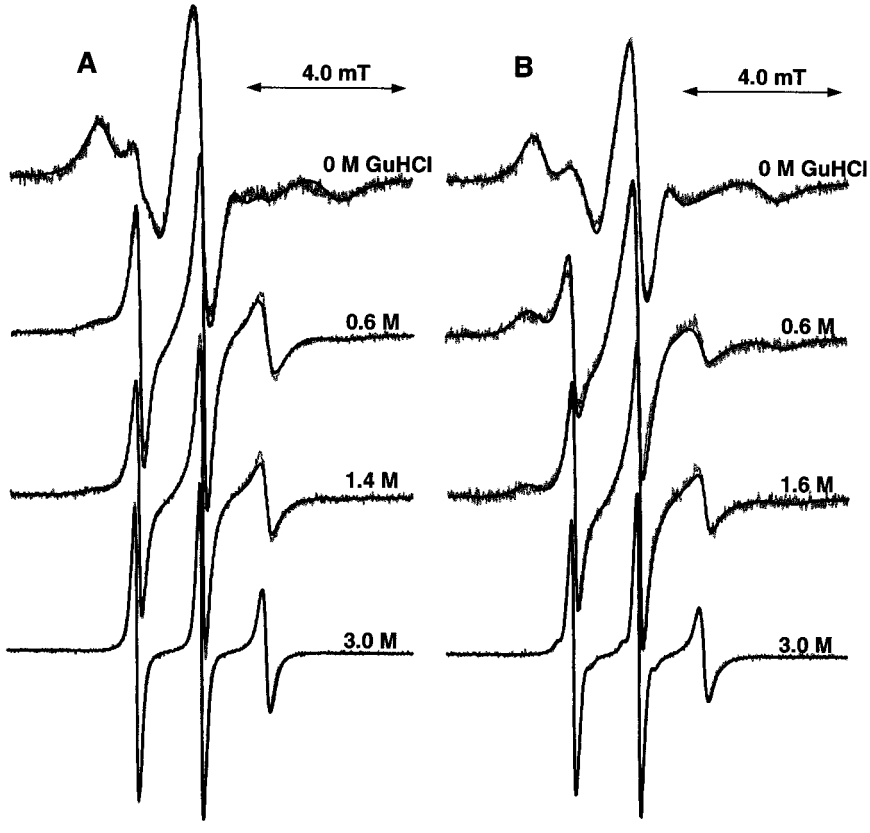

FIGURE 9 EPR spectra obtained from spin-labeled L79C at different degrees of unfolding. Experimental (*gray line*) and simulated (*black line*) EPR spectra of MTSSL-labeled L79C (*A*) and IPSL-labeled L79C (*B*) at different Gu-HCl concentrations. Each simulated spectrum is built up of a superposition of three mobility components associated with three conformational states of the protein: native, intermediate, and unfolded. The relative weights of each component are plotted in Fig. 12. For details concerning the parameters of each component, see Tables 4 and 5 and detailed discussions in the text.

3 *E* illustrates the corresponding curve for MTSSL-L79C. IPSL is greatly immobilized in the native state of the protein, as indicated by the broad lines in the spectrum (Fig. 9 *B*). During unfolding, the mobility of IPSL is increased dramatically, as indicated by narrowing of the spectra with increasing Gu-HCl concentration, resulting in the steep slopes of the first and second unfolding transitions seen in Fig. 3 *D*. In between the two transitions, a stable intermediate state can be seen in the IPSL-L79C unfolding curve. The C_m values of these transitions do not deviate much from the corresponding values of the global transitions (Table 1). At ~3 M Gu-HCl, IPSL reaches its maximal mobility, indicating that the protein has reached its unfolded state. In regard to overall behavior, IPSL-L79C shows some similarities with the IPSL-W97C mutant, in which the mutation site is situated within the hydrophobic core of HCA II, as documented in a previous study (Svensson et al., 1995). MTSSL-L79C displayed a similar response to unfolding (Fig. 9 *A*), although the first transition and the intermediate state are not evident (Fig. 3 *E*). The C_m value for the broad N→I transition is 0.95 M Gu-HCl for MTSSL-L79C (obtained from EPR amplitude data) as compared with 0.31 M Gu-HCl for IPSL-L79C. C_m values for the I→U transition are approximately equal. However, MTSSL-L79C reaches its maximal mobility at a lower Gu-HCl concentration (~2.5 M Gu-HCl). Nevertheless, in the unfolded state, MTSSL-L79C is less mobile than IPSL-L79C, as indicated by the lower EPR amplitude.

EPR simulations of the spin-labeled native and unfolded states

In Fig. 9, line-shape simulations are included with the experimental EPR spectra of MTSSL-L79C and IPSL-L79C, respectively. To perform line-shape simulations that closely matched experimental EPR spectra, it was of interest to determine magnetic tensor components of the different environments that the spin labels are likely to be exposed to during unfolding of the protein. In a previous study (Owenius et al., 1999), we noticed possible solvent dependence of some of the isotropic and anisotropic EPR parameters (especially for A_{iso} and A_{zz}) obtained by measurement of mercaptoethanol adducts of IPSL and MTSSL in Tris-HCl and octanol. A more detailed study of the influence of the solvent characteristics on the magnetic tensors of IPSL and MTSSL, using high-field EPR and quantum chemical calculations, will be published elsewhere.

The EPR spectrum associated with the native and unfolded form was each readily simulated using a unique dynamic model (top and bottom spectra of Fig. 9, *A* and *B*). The parameters used for the simulations are summarized in Table 4. There is a small contribution (~5%) from a more rapid component in the IPSL spectrum of the native form (discussed more in the following section). The simulations are, however, indicative of a homogeneous and uniform structure of the protein in the native state. The unfolded state is a very heterogeneous ensemble of structures; however, due to the extensive dynamics of the unfolded state the spectroscopic response will appear homogeneous.

TABLE 4 Simulated EPR parameters of IPSL-L79C and MTSSL-L79C for the native (N) and unfolded (U) protein conformations

	IPSL†		MTSSL†	
Parameter*	Slow (N)	Rapid (U)	Slow (N)	Rapid (U)
A_{xx} (G)	6.5	6.5	7.1	7.2
A_{yy} (G)	5.5	5.5	6.0	6.1
A_{zz} (G)	35.1	35.9	34.7	35.5
d_{xy} (s^{-1})	2.1×10^7	2.0×10^8	1.7×10^7	1.4×10^8
d_{zz} (s^{-1})	2.1×10^7	6.0×10^9	1.7×10^7	3.3×10^9
ϕ (°)	27		28	
lw (G)	0.60	0.35	0.60	0.26

*Parameters are defined as follows: principal components of the hyperfine splitting tensor (A_{xx}, A_{yy}, A_{zz}), principal components of the rotational diffusion tensor (d_{xy}, d_{zz}), also known as ($R_\perp$, $R_\parallel$), diffusion tilt angle (ϕ), and intrinsic (Lorentzian) line-width (*lw*). A definition of coordinate axis systems and the relation between molecular magnetic frame and diffusion frame is depicted in Fig. 2 in Schneider and Freed, 1989.

†EPR *g*-tensor components used in simulations of IPSL spectra were g_{xx} = 2.0084, g_{yy} = 2.0056, and g_{zz} = 2.0022 and of MTSSL spectra were g_{xx} = 2.0080, g_{yy} = 2.0059, and g_{zz} = 2.0023. The truncation values L^e_{max}, L^o_{max}, and K_{max} were set to 44. The components of different mobility correspond to slow (native, N) 0 M GuHCl and rapid (unfolded, U) 3.0 M GuHCl. All spectra were convoluted with a Gaussian function having a peak-to-peak line-width (gw) of 1.0 G. For MTSSL a 4% spectral contribution due to ^{13}C-satellites was convoluted (7.0 G) together with the Gaussian line-shape.

Two models for the aggregated state

As indicated by the refolding, fluorescence anisotropy, and lifetime measurements, HCA II forms aggregates at intermediate Gu-HCl concentrations. At this degree of unfolding some protein concentration dependence of the aggregation was found by comparing EPR spectra obtained from spin-labeled L79C at 5 and 20 μM (data not shown). Lower protein concentrations could not be studied due to the poor signal-to-noise ratio. At 2–3 min after adding 1.2 M Gu-HCl to the samples most of the aggregates are already formed, as further incubation of the samples did not result in more than a marginal broadening of the EPR lines.

A spectrum as shown in Fig. 10 *A* was obtained for MTSSL-L79C at 1.0 M Gu-HCl, and it clearly reveals contribution from a broad component, similar to the spectrum of the native structure. The important difference is that the slowest component in the spectrum covers a considerably wider range than that of the native structure. (The lower dotted spectrum of Fig. 10 *A* corresponds to the simulated spectrum of the native state.) The spin labels in the native state of IPSL-L79C and MTSSL-L79C possess almost no intrinsic freedom, but rotate as locked inside the protein with the tumbling rate of the protein as a whole (see

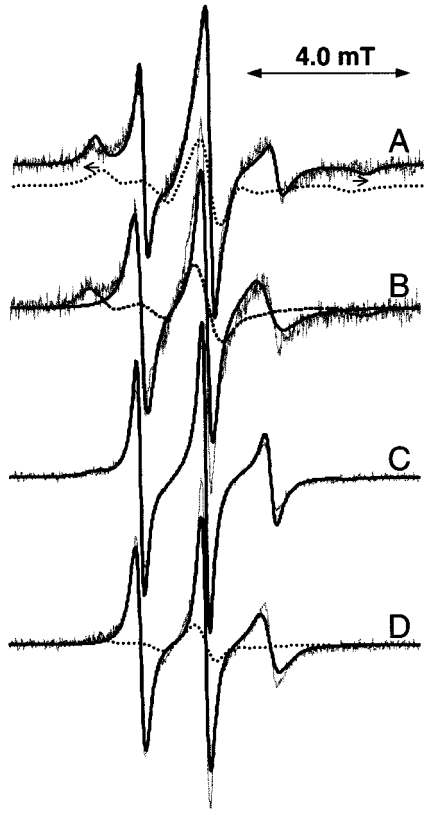

FIGURE 10 EPR spectra of spin-labeled L79C in the aggregated state. Experimental (*gray line*) and simulated EPR spectra of MTSSL-labeled L79C (*A* and *B*) and IPSL-labeled L79C (*C* and *D*) at 1.0 M Gu-HCl. (*A*) Experimental spectrum together with simulation according to the disordered model (*black line*). The dotted line is the simulation of the native state. (*B*) Experimental spectrum together with simulations according to two-state model, i.e., slow-aggregate (*dotted line*) and medium-aggregate (*black line*) components. (*C*) Experimental spectrum together with simulation according to the disordered model (*black line*). (*D*) Experimental spectrum together with simulations according to two-state model, i.e., slow-aggregate (*dotted line*) and medium-aggregate (*black line*) components. Further details about the models are given in the text. EPR parameters of each simulated spectrum are given in Tables 4 and 5.

below). Thus, broader lines than the native EPR lines can only be a spectral component indicating slower global tumbling of the protein, i.e., aggregated protein with spin labels that are equally locked inside the aggregated structure as in the native protein. In addition, there is a contribution from a faster tumbling spin label as indicated by the three distinct sharper peaks, however, considerably broader than observed for the unfolded form. The overall spectrum cannot be simulated using a dominating homogeneous spectral component as in the case of the native or unfolded states. At intermediate Gu-HCl concentrations (0.8–1.2 M Gu-HCl) more than 95% of the protein was not recovered in refolding experiments, indicating irreversible formation of aggregates. To reconstruct the MTSSL-L79C spectrum obtained at 1.0 M Gu-HCl, we started by assuming that the entire spectrum was due to aggregates. The parameter that has the greatest influence on the line-shape is the rotational diffusion constants. To form a composed spectrum, a series of EPR signals associated with rotational diffusion constants lower than those of the unfolded form was calculated. In the graph in Fig. 11 this series is depicted as open triangles. The parameter set was chosen to follow a straight line from the rotational diffusion parameters of the unfolded form to those of the native form. Because the line-shape at the very slow region is not sensitive to rotational anisotropy, d_{xy} and d_{zz} were set equal (isotropic motion) for values smaller than for the native form, thus lining up along the line representing $d_{xy} = d_{zz}$ (isotropic motion). As a first attempt to make a composed spectrum, each spectrum in the series was superimposed with the same weight. The simulation, normalized to the same apparent free radical concentration as the experimental spectrum, is shown as the black solid line in Fig. 10 *A*. The similarity to the experimental spectrum is striking, considering the simplicity of the model, and this supports the fact that the structure around position 79 at 7.1–8.8 Å distance from C_α in the aggregated protein can be interpreted as an inhomogeneous disordered structure. This model is referred to as the disordered model. The quality of the simulation can be readily improved by adjusting the weights of the components making up the superimposed

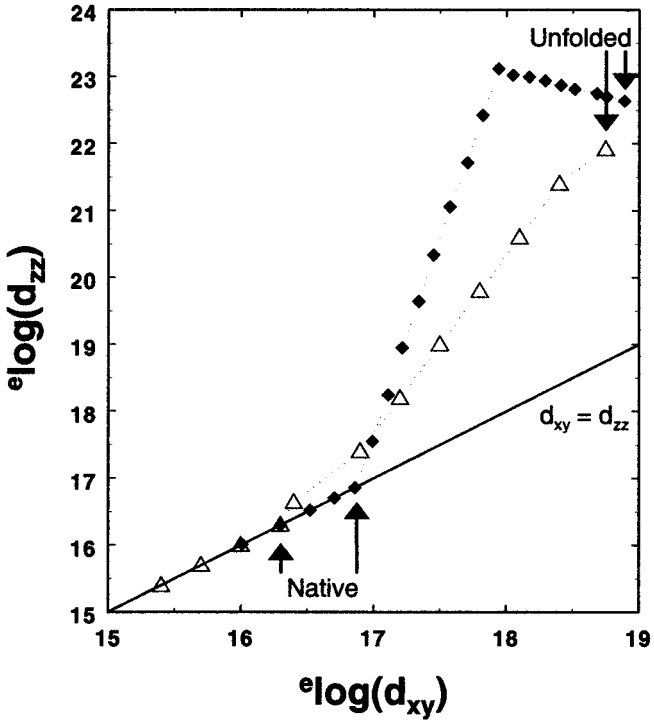

FIGURE 11 Map of the rotational diffusion tensor (d_{xy}, d_{zz}) values (disordered model) used to make up the mobility components associated with the aggregated state of MTSSL-L79C (△) and IPSL-L79C (◆). The components of the native and unfolded states are denoted by arrows. Rotational diffusion tensor values at the solid line describe isotropic rotational diffusion ($d_{xy} = d_{zz}$), and tensor values above the solid line correspond to anisotropic rotational diffusion ($d_{xy} < d_{zz}$).

spectrum; however, this can be made in a variety of ways. Moreover, the uniqueness of certain parameters, such as line-width for each component, is not possible to deduce. To search for a model that uses a minimal number of components, the same series of calculated spectra was used to make non-negative least-square fits to experimental spectra in the Gu-HCl concentration interval 0.8–1.2 M. It was found that the relative weights determined from these fits always were larger for two distinct components making up ~80% of the total intensity: one of slowest motion and one component with a rotational diffusion being intermediate to the simulation of the native and unfolded forms. Thus, as an alternative model the experimental spectrum at 1.0 M Gu-HCl was used to find a composed spectrum of only two components, a two-state model, where the anisotropy and magnitude of the rotational diffusion constant were varied to fit the experimental spectrum. The *g*- and *A*-tensor components as well as the line-widths were similar or intermediate to those of the simulation of the native and unfolded forms. The resulting calculated spectra are shown in Fig. 10 *B*, and the parameters used in the simulations are summarized in Table 5. The relative weights of these slow-aggregate and medium-aggregate components, to make up a simulation for the experimental spectrum at 1.0 M Gu-HCl, was found to be 0.52 and 0.48, respectively. Conclusively, EPR spectra simulations of the MTSSL-labeled aggregated protein clearly proves that at least two distinct components are involved, one associated with a spin label tumbling slower than in the native state, the other corresponding to a spin label in a dynamic environment between the native and unfolded states. The simulations also indicate that the aggregated state could be described by a more wide distribution of spin-label environments; however, it is not possible to deduce the details of these from solely the EPR data.

TABLE 5 Simulated EPR parameters of IPSL-L79C and MTSSL-L79C for the aggregated state

	IPSL†		MTSSL‡	
Parameter*	SA	MA	SA	MA
A_{xx} (G)	6.5	6.5	7.1	7.2
A_{yy} (G)	5.5	5.5	6.0	6.1
A_{zz} (G)	35.1	35.9	34.7	35.2
d_{xy} (s^{-1})	0.9×10^7	6.2×10^7	1.8×10^6	4.7×10^7
d_{zz} (s^{-1})	0.9×10^7	1.1×10^{10}	1.8×10^6	6.5×10^9
ϕ (°)		30		28
lw (G)	0.60	0.64	0.60	0.30

For parameter definitions and for EPR *g*-tensor components used in simulations, see Table 4. The truncation values L^c_{max}, L^o_{max}, and K_{max} were set to 44. The components of different mobility correspond to slow-aggregate (SA) and medium-aggregate (MA) at 1.0 M GuHCl. See the text for details of the procedure. All spectra were convoluted with a Gaussian function having a peak-to-peak line-width (gw) of 1.0 G.

†In the two-state model the SA and MA spectra were added with weights 0.38 and 0.62, respectively. This defines the intermediate state (I).

‡In the two-state model the SA and MA spectra were added with weights 0.52 and 0.48, respectively. This defines the intermediate state (I).

Similar results were obtained analyzing the spectra of IPSL–L79C. As discussed in previous sections, IPSL has a more flexible linker and has a stronger propensity to stick out in the more flexible regions further away from the protein backbone. The main spectral difference from MTSSL-L79C is that the contribution from the native component disappears at lower Gu-HCl concentrations. Moreover, the relative contribution from a slow component of the aggregated state is also smaller. The analysis was performed in a similar manner as for MTSSL. Spectra of a disordered model are shown in Fig. 10 *C*. Here we show a slightly modified model introducing more anisotropy in the intermediate region (Fig. 11, filled diamonds). There is, however, little difference from a model that follows the procedure used to choose the series of calculated spectra for MTSSL-L79C (data not shown). The components of the alternative two-state model for the aggregate are shown in Fig. 10 *D*. Notably, the relative weights of the slow-aggregate and medium-aggregate are different from the case of MTSSL-L79C (0.38 and 0.62, respectively). We ascribe this difference to the intrinsic properties of the spin labels as will be discussed further below. The parameters are summarized in Table 5.

Interpretation of the EPR parameters

What is characteristic for the dynamic state of the spin label in the three defined states? The three simulated spectra of the native state (0 M Gu-HCl), the intermediate/aggregated state (1.0 M Gu-HCl) obtained from either the disordered or two-state model, and the unfolded state were used to make non-negative least-square fits to experimental spectra recorded in the Gu-HCl range 0–3 M. The relative weights of the three components in the population analysis for both MTSSL-L79C and IPSL-L79C are presented in Fig. 12. They reflect the fraction of the corresponding conformational states of the substructure around each label present at different Gu-HCl concentrations. Thus, the slow component curve shows how much folded structure that is retained locally around position 79 in various stages of the unfolding process. Similarly, the intermediate and rapid component curves reflect the amounts of aggregated and unfolded conformations, respectively.

It can be seen in Fig. 9 that the spin labels in both IPSL-L79C and MTSSL-L79C were immobilized in the native state. This is presented in Fig. 12 as relative weights of almost 1.0 for the slow components of both HCA II variants. The slow mobility component of IPSL-L79C was obtained from a simulation using an isotropic rotational diffusion of 2.1×10^7 s^{-1} (Table 4). The immobilization of MTSSL-L79C was even more pronounced, with an isotropic rotational diffusion of 1.7×10^7 s^{-1}. The time scale of this motion, with a rotational correlation time of 7.9 ns (IPSL-L79C) or 9.8 ns (MTSSL-L79C), is associated with the global tumbling of the protein itself (cf. 9.5 ns for

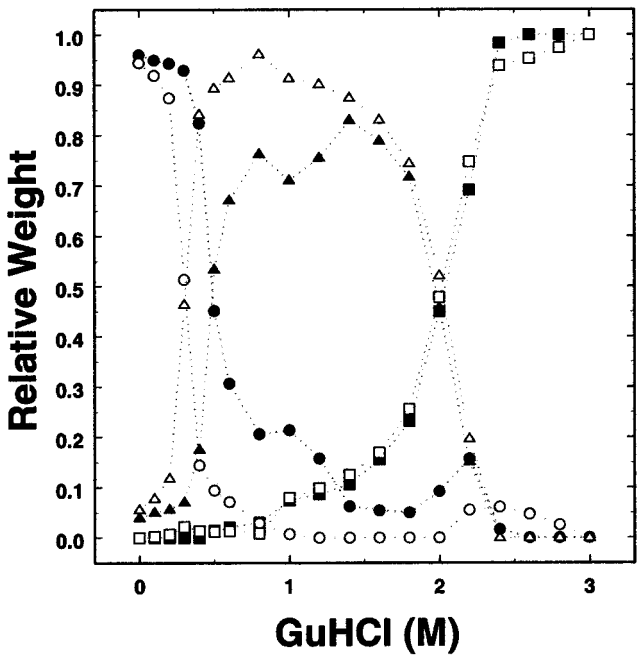

FIGURE 12 Relative abundance of slow, intermediate, and rapid mobility components obtained by non-negative least-square fitting to experimental data using superpositions of the simulated spectra associated with 0, 1.0 (two-state model), and 3.0 M Gu-HCl, corresponding to the native (○), intermediate (△), and unfolded (□) states, respectively. Filled symbols, MTSSL-L79C; open symbols, IPSL-L79C.

IPSL-W97C located in the core of HCA II; Svensson et al., 1995), in excellent agreement with results of the combined anisotropy and lifetime measurements presented in a previous section. Thus, both IPSL-L79C and MTSSL-L79C seem to be held in a rigid position by the highly structured surrounding environment. In such a case it is possible to determine the diameter of HCA II (assuming that the protein is spherical and in a water solvent at room temperature) from the rotational diffusion constant (Goldman et al., 1972). By doing so, we found the diameter to be 39–42 Å, depending on the spin label used, which agrees fairly well with the crystal structure data (Eriksson et al., 1988).

Unfolding of the protein resulted in a decrease of the slow component and a concomitant increase of the intermediate component (Fig. 12). The midpoints of unfolding of these transitions were 0.50 and 0.30 M Gu-HCl for MTSSL-L79C and IPSL-L79C, respectively, reflecting the differences in global stability of the two HCA II variants (Table 1). The relative weights of the slow and intermediate components after the first unfolding transition (at 1.0 M Gu-HCl) were, respectively, 0.21 and 0.71 for MTSSL-L79C and 0.01 and 0.91 for IPSL-L79C. This shows that the mobility of MTSSL, compared with that of IPSL, was more restricted at that particular Gu-HCl concentration (as discussed below). Interestingly, due to aggregation of the protein the intermediate state has to be described by at least two components (e.g., the two-state model), where one component exhibited increased rotational diffusion (medium-aggregate) and the other component exhibited decreased rotational diffusion (slow-aggregate) compared with the slow component of the native state (Table 5). From the rotational diffusion tensor of the slow-aggregate component an estimation of the maximum size of the protein aggregates could be made as described above. Because MTSSL has a less flexible structure than IPSL, and easily becomes locked within a restricted environment, it will be the better reporter of global motion of the two. This is also evident from Fig. 10, *B* and *D*, where the slow-aggregate component appears more clearly in MTSSL-L79C than IPSL-L79C. The diameter determined for MTSSL-L79C was 89 Å, corresponding to an increase in molecular volume of ~9.4 times relative to the calculated protein volume in the native state (Table 4). The IPSL probe, on the other hand, will underestimate the aggregate size (52 Å) due to its more flexible structure. This value is similar to the diameter calculated from the fluorescent anisotropy and lifetime data, 52 Å for IAEDANS-L79C and 56 Å for BADAN-L79C, in the latter case referring to the dominant contribution in the multi-component time-resolved trace (Table 3). Thus, these labels all have a similar flexibility. The estimation from MTSSL-L79C is in fairly good agreement with previous results obtained from dynamic light-scattering measurements on HCA II_{pwt} at 2.0 M Gu-HCl, revealing an aggregate size of 135 Å, however, at twice the protein concentration used in the EPR measurements (Hammarström et al., 1999). We also noted a slight protein concentration dependence of the line-widths of the EPR spectrum indicating that larger aggregates were formed at higher concentrations. However, the size is difficult to determine accurately because such large aggregates tumble with a rotational correlation time (>100 ns) for which the sensitivity of the CW EPR spectrum to the nitroxide motion is low (Millhauser et al., 1995). Other EPR methods have to be employed to study the tumbling of larger aggregates. Line-shape simulations of the medium-aggregate component of the spin-labeled HCA II variants showed that the overall motion ($d = (d_{xy}d_{zz})^{0.5}$) of the probe is ~50% larger in IPSL-L79C compared with MTSSL-L79C (Table 5). Note also the large anisotropy introduced in the rotational motion ($d_{xy} < d_{zz}$) of both spin labels, which indicates that the diffusion around the bond connecting the probe to the protein is less restricted than the diffusion perpendicular to the bond.

Unfolding of the aggregated protein resulted in a drastic increase of the rapid component due to the disruption of secondary structure (Fig. 12). Considering the profile of the relative weight of the rapid component, MTSSL-L79C and IPSL-L79C were very similar in the 0–2.0 M Gu-HCl range. Also in the unfolded state, IPSL is more mobile than MTSSL (~60%), but the anisotropy of the motion is reduced compared with the medium-aggregate component due to exposure of the spin labels to the solvent.

Differences in the probing abilities of the spin labels

As observed when using the fluorescent labels, the specific properties of the spin labels give rise to a different spectral response at various stages of unfolding. Measuring global stability by use of intrinsic Trp fluorescence showed that, due to its larger volume, IPSL destabilized the L79C mutant more than MTSSL did. This general difference should also be apparent in the local unfolding curves, which was confirmed by comparing the C_m values of the local and global unfolding curves (Table 1). However, this global difference should have had only a small effect on the local unfolding curves of the two spin labels; nevertheless, the N→I transition of the protein was probed in very different ways by these labels (Fig. 3, *D* and *E*). Instead, we believe that the difference between IPSL and MTSSL with respect to how they probe their local environment arose from differences in structure, polarity, and/or flexibility of the labels. In the native state of the L79C mutant, both spin labels were almost totally immobilized in the local hydrophobic cluster surrounding position 79, and the motion indicated by the EPR spectra was close to the rotational diffusion of the whole protein molecule (Table 4). In the first transition, IPSL-L79C was rapidly mobilized as the surrounding structure loosened up, whereas MTSSL-L79C remained more immobilized as judged from the components used to explain the composed aggregate spectrum (Fig. 10; Table 5). In the native protein, the approximate probing distance is 7.4 Å for IPSL-L79C and 7.1 Å for MTSSL-L79C, i.e., about the same, if the constrained conformation noted in previous molecular modeling experiments on a protein-protein interface is assumed (Owenius et al., 1999). Even though IPSL is slightly more polar than MTSSL, the distribution of lipophilic potential is almost the same for the two spin labels compared with the fluorescent labels (Fig. 4). Therefore, it is more likely that the difference seen in the first unfolding transition was caused by the lower flexibility of MTSSL compared with IPSL. This can be due to its rather constrained disulfide bond at the attachment site and/or the smaller number of atoms in the MTSSL linker (Richardson, 1981; Owenius et al., 1999). The second transition, on the other hand, occurred at approximately the same Gu-HCl concentration for both spin labels (Table 1). In this transition, there were fewer motional restrictions; thus, the effects of the constrained flexibility of MTSSL decreased. However, the lower EPR amplitude (Fig. 3, *D* and *E*) and lower overall motion (Table 4) of MTSSL-L79C than of IPSL-L79C in the unfolded state indicate that that the structure of MTSSL is more constrained than that of IPSL. When the motion of the spin labels becomes less restricted, the length of the side chain is probably gradually increased. In the most extended conformations, the modeled probing distance of IPSL is ~1.2 Å longer than that of MTSSL (Table 2).

Summary of the unfolding behavior of HCA II around position 79

The various unfolding transitions detected by the specific labels are summarized in Table 1. The local hydrophobic cluster (Fig. 1 *B*), which includes the side chain of position 79, consists of the amino acids L47, V49, Y51, Y88, L144, and I210, which are all within 6 Å of residue 79 in the crystal structure. In the native state, the side chain would be buried, because the accessible surface area of L79 in the crystal structure was calculated to be 0 $Å^2$, as determined by a probe radius of 1.4 Å (Connolly, 1983). The results show that C79 is not accessible to iodo[2-^{14}C]acetic acid in the native state, the polarity of the local structure is very hydrophobic, and the spin labels have a very immobilized EPR signature. During the first global unfolding transition (N→I), the outer part of the cluster becomes accessible; i.e., the thiol group is accessible for reagent. The inner part of the cluster, however, is shielded as shown by the BADAN fluorescence and the immobilized MTSSL spin label. The IAEDANS moiety, on the other hand, becomes exposed to the solvent, and the IPSL spin label enters a state of increased mobility. The cluster around position 79 in the molten globule is large and hydrophobic enough to house the BADAN and MTSSL moieties, but not the longer and more polar/flexible IAEDANS and IPSL labels. The longer linkers of the IAEDANS and IPSL labels would enable IAEDANS and IPSL to reach the outer parts of the large central hydrophobic core, but the shorter labels are situated inside the local cluster. However, the outer part of the β-strand is exposed, because the thiol group of C79 was accessible to iodo[2-^{14}C]acetic acid.

The intermediate state of HCA II forms aggregates and the aggregation interface was previously assigned to β-strands 4–7 (Hammarström et al., 1999). Because the maximal level of iodo[2-^{14}C]acetic acid reactivity was achieved in the aggregate, position 79 is not likely to be completely buried in the aggregation interface. It was also previously shown that position 79 was not involved in a homologous aggregation interface due to the lack of excimer fluorescence from a pyrene maleimide-labeled L79C (Hammarström et al., 1999). In addition, in a recent fluorescence resonance energy transfer study, no evidence was found for involvement of position 79 in the aggregation interface both in Gu-HCl and thermally induced aggregates (Hammarström et al., 2001). However, the proximity of position 79 to the aggregation interface showed that the spectroscopic labels respond to the aggregation structure. L79 is located in a peripheral β-strand joining β-strands 3 and 4. The structure around position 79 could be characterized at increased resolution due to the multitude of labels used. The following picture of the aggregated structure emerged. The local hydrophobic cluster surrounding position 79 is largely intact. A population of the BADAN and MTSSL labels are therefore locked into this cluster as the

protein aggregates; hence the local structure is hydrophobic and rigid. On the other hand, another population of the labels is partially exposed showing that the aggregate can be described as an ensemble of different structures. Increased protein concentration during aggregate formation affects global tumbling measured by both fluorescence anisotropy and EPR. The increased protein concentration therefore results in larger conglomerates of molten-globule HCA II molecules. The hydrophobicity of the environment surrounding BADAN seems to increase with increased protein concentration, which indicates that the packing of the structure around position 79 is also affected, probably due to the proximity to the aggregation interface. This phenomenon was also found for IAEDANS-L79C, but was detected as an increase in fluorescence anisotropy of labels with blue-shifted emission. The IAEDANS and IPSL labels are larger, more flexible and polar than the BADAN and MTSSL labels and will therefore be more solvent exposed in the aggregate. Notwithstanding, the decreased global tumbling is evident from the IAEDANS-L79C fluorescence anisotropy. This can also be observed to some extent from the line broadening of the EPR spectrum of IPSL-L79C.

Unfolding of the aggregation-prone intermediate (unfolding of the secondary structure and disruption of the aggregate (Hammarström et al., 1999)) of the protein would separate the local cluster around position 79 from the central part of the HCA II molecule. After the second transition the fluorescent labels become almost fully exposed, and the spin labels become fully mobilized. However, some indications of residual structure are evident. The region comprising β-strands 3–7 has been found to persist under strong denaturing conditions in previous studies (Mårtensson et al., 1993; Svensson et al., 1995; Hammarström et al., 1997). The previous findings were mostly based on inaccessibility of thiol side chains to iodo[2-^{14}C]acetic acid. For L79C, the chemical reactivity experiments showed a highly accessible side chain after the first unfolding transition, which decreased at high concentrations of Gu-HCl. This indicates a shielding of the thiol group that would require a compactness of the surrounding structure. The spectroscopic labels also detected structural rearrangements in the denatured state. A possible third unfolding transition was detected for IAEDANS-L79C between 2.5 and 4.0 M Gu-HCl, and BADAN was not fully exposed until 4.0 M Gu-HCl was applied. Unfolding of the unfolded state is non-cooperative and reflects disruption of hydrophobic clusters. Restrictions on accessibility and polarity in the unfolded state, as viewed from position 79, can most likely be interpreted as connections with the main hydrophobic cluster. Based on the present findings, we conclude that the peripheral β-strand connecting β-strands 3 and 4, where L79 is situated, is also part of this stable residual structure, although through looser interactions. The inner part of the β-strand is, however, connected to the main cluster of the protein via hydrophobic interactions.

Summary of the various probing functions

In summary, we conclude that a resolved picture of conformational changes in a protein can be obtained by probing the environment around a specific site at varying distances from the target site in the protein. Our results show that it is important to know the length, flexibility, and polarity of the different labels when interpreting structural changes in a protein. Some labels are more suitable than others, depending on the system being studied. A longer linker will enable the IAEDANS and IPSL labels to monitor a larger volume of space in the surroundings of the labeled site than will be monitored by the BADAN and MTSSL labels. The more polar character of IAEDANS compared with BADAN and the larger flexibility of IPSL compared with MTSSL will contribute to the different locations the probes reside in. The chemical reactivity approach is straightforward and does not require modification of the protein to be able to monitor accessibility and compactness. Regarding the fluorophores, BADAN provides higher quantum yield compared with IAEDANS in hydrophobic environments and gives better resolution of the surrounding polarity in a wider range of hydrophobic to polar solvents. On the other hand, IAEDANS yields a higher degree of labeling and is more sensitive than BADAN in more polar environments. The longer fluorescence lifetime of IAEDANS compared with that of BADAN will also render increased sensitivity of fluorescence anisotropy for mid-sized proteins. Considering the spin labels, it is evident that MTSSL and IPSL provide complementary information. In compact environments, MTSSL is more restricted due to lower flexibility of the linker, and in more dynamic surroundings the longer probing range of IPSL can resolve interactions with neighboring structures at a greater distance from the attachment site than MTSSL can. Concerning buried positions, it is important to keep in mind that different labels will perturb protein stability to different extents; introduction of a side chain with a larger volume will cause more substantial perturbations. By using a wide variety of methods, instead of a single technique, to monitor conformational changes in proteins, the risk of artifacts can be minimized.

We thank Camilla Eklund for excellent technical assistance, Dr. Malin Persson for valuable discussions, and Dr. Bengt-Harald Jonsson (Umeå University) for comments on the manuscript.

This work was supported by grants from the Swedish Natural Science Research Council (M.L. and U.C.), the Foundation for Strategic Research (U.C.), the Wenner Gren Foundations (P.H.), and the Swedish Medical Research Council (U.C.).

REFERENCES

Alexander, R. S., S. K. Nair, and D. W. Christianson. 1991. Engineering the hydrophobic pocket of carbonic anhydrase II. *Biochemistry.* 30: 11064–11072.

Armstrong, J. M., D. V. Myers, J. A. Verpoorte, and J. T. Edsall. 1966. Purification and properties of human erythrocyte carbonic anhydrases. *J. Biol. Chem.* 241:5137–5149.

Ballery, N., P. Minard, M. Desmadril, J. M. Betton, D. Perahia, L. Mouawad, L. Hall, and J. M. Yon. 1990. Introduction of internal cysteines as conformational probes in yeast phosphoglycerate kinase. *Protein Eng.* 3:199–204.

Birkett, D. J., R. A. Dwek, G. K. Radda, R. E. Richards, and A. G. Salmon. 1971. Probes for the conformational transitions of phosphorylase b. Effect of ligands studied by proton relaxation enhancement, fluorescence and chemical reactivities. *Eur. J. Biochem.* 20:494–508.

Carlsson. U., R. Aasa, L. E. Henderson, B.-H. Jonsson, and S. Lindskog. 1975. Paramagnetic and fluorescent probes attached to "buried" sulfhydryl groups in human carbonic anhydrases. Application to inhibitor binding, denaturation and refolding. *Eur. J. Biochem.* 52:25–36.

Connolly, M. L. 1983. Analytical molecular surface calculation. *J. Appl. Crystallogr.* 16:548–558.

Eriksson, E. A., A. T. Jones, and A. Liljas. 1988. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins Struct. Funct. Genet.* 4:274–282.

Fayet, O., T. Ziegelhoffer, and C. Georgopoulos. 1989. The GroES and GroEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J. Bacteriol.* 171:1379–1385.

Fierke, C. A., T. L. Calderone, and J. F. Krebs. 1991. Functional consequences of engineering the hydrophobic pocket of carbonic anhydrase II. *Biochemistry.* 30:11054–11063.

Freskgård, P.-O., U. Carlsson, L.-G. Mårtensson, and B.-H. Jonsson. 1991. Folding around the C-terminus of human carbonic anhydrase II. Kinetic characterization by use of a chemically reactive SH-group introduced by protein engineering. *FEBS Lett*. 298:117–122.

Freskgård, P.-O., U. Carlsson, L.-G. Mårtensson, and B.-H. Jonsson. 1994. Assignment of the contribution of the tryptophan residues to the circular dichroism spectrum of human carbonic anhydrase II. *Biochemistry.* 33:14281–14288.

Goldman, S. A., G. V. Bruno, and J. H. Freed. 1972. Estimating slow-motional rotational times for nitroxides by electron spin resonance. *J. Phys. Chem.* 76:1858–1860.

Håkansson, K., M. Carlsson, L. A. Svensson, and A. Liljas. 1992. Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes. *J. Mol. Biol.* 227:1192–1204.

Hammarström, P., B. Kalman, B.-H. Jonsson, and U. Carlsson. 1997. Pyrene excimer fluorescence as a proximity probe for investigation of residual structure in the unfolded state of human carbonic anhydrase II. *FEBS Lett.* 420:63–68.

Hammarström, P., M. Persson, and U. Carlsson. 2001. Protein compactness measured by FRET: human carbonic anhydrase II is considerably expanded by the interaction of GroEL. *J. Biol. Chem*. In press.

Hammarström, P., M. Persson, P.-O. Freskgård, L.-G. Mårtensson, D. Andersson, B.-H. Jonsson, and U. Carlsson. 1999. Structural mapping of an aggregation nucleation site in a molten-globule intermediate. *J. Biol. Chem.* 274:32897–32903.

Haugland, R. P. 1996. Handbook of Fluorescent Probes and Research Chemicals, 6th ed. Molecular Probes, Eugene, OR.

Heiden, W., G. Moeckel, and J. Brickmann. 1993. A new approach to analysis and display of local lipophilicity/hydrophilicity mapped on molecular surfaces. *J. Comput. Aided Mol. Design.* 7:503–514.

Hubbell, W. L., A. Gross, R. Langen, and M. A. Lietzow. 1998. Recent advances in site-directed spin labeling of proteins. *Curr. Opin. Struct. Biol.* 8:649–656.

Khalifah, R. G., D. J. Strader, S. H. Bryant, and S. M. Gibson. 1977. Carbon-13 nuclear magnetic resonance probe of active-site ionizations in human carbonic anhydrase B. *Biochemistry.* 16:2241–2247.

Lambert, M. P., A. K. Barlow, B. A. Chromy, C. Edwards, R. Freed, M. Liosatos, T. E. Morgan, I. Rozovsky, B. Trommer, K. L. Viola, P. Wals, C. Zhang, C. E. Finch, G. A. Krafft, and W. L. Klein. 1998. Diffusible, nonfibrillar ligands derived from A1-42are potent central nervous system neurotoxins. *Proc. Natl. Acad. Sci. U.S.A.* 95:6448–6453.

Likhtenshtein, G. I. 1993. Biophysical Labeling Methods in Molecular Biology. Cambridge University Press, New York.

Lindgren, M., M. Svensson, P.-O. Freskgård, U. Carlsson, B.-H. Jonsson, L.-G. Mårtensson, and P. Jonasson. 1993. Probing local mobility in carbonic anhydrase. EPR of spin-labeled SH groups introduced by site-directed mutagenesis. *J. Chem. Soc. Perkin. Trans*. 2:2003–2007.

Lindgren, M., M. Svensson, P.-O. Freskgård, U. Carlsson, P. Jonasson, L.-G. Mårtensson, and B.-H. Jonsson. 1995. Characterization of a folding intermediate of human carbonic anhydrase II: probing local mobility by electron paramagnetic resonance. *Biophys. J.* 69:203–213.

Mansoor, S. E., H. S. Mchaourab, and D. L. Farrens. 1999. Determination of protein secondary structure and solvent accessibility using site-directed fluorescence labeling. Studies of T4 lysozyme using the fluorescent probe monobromobimane. *Biochemistry.* 38:16383–16393.

Mårtensson, L.-G., B.-H. Jonsson, P.-O. Freskgård, A. Kihlgren, M. Svensson, and U. Carlsson. 1993. Characterization of folding intermediates of human carbonic anhydrase II: probing substructure by chemical labeling of SH groups introduced by site-directed mutagenesis. *Biochemistry.* 32:224–231.

Mårtensson, L.-G., P. Jonasson, P.-O. Freskgård, M. Svensson, U. Carlsson, and B.-H. Jonsson. 1995. Contribution of individual tryptophan residues to the fluorescence spectrum of native and denatured forms of human carbonic anhydrase II. *Biochemistry.* 34:1011–1021.

Millhauser, G. L., W. R. Fiori, and S. M. Miick. 1995. Electron spin labels. *Methods Enzymol*. 246:589–610.

Nozaki, Y. 1972. The preparation of guanidine hydrochloride. *Methods Enzymol.* 26:43–50.

Nyman, P. O., and S. Lindskog. 1964. Amino acid composition of various forms of bovine and human erythrocyte carbonic anhydrase. *Biochim. Biophys. Acta.* 85:141–151.

Owenius, R., M. Österlund, M. Lindgren, M. Svensson, O. H. Olsen, E. Persson, P.-O. Freskgård, and U. Carlsson. 1999. Properties of spin and fluorescent labels at a receptor-ligand interface. *Biophys. J.* 77: 2237–2250.

Richardson, J. S. 1981. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 34:167–339.

Schneider, D. J., and J. H. Freed. 1989. Calculating slow motional magnetic resonance spectra. *In* Biological Magnetic Resonance, Vol. 8: Spin Labeling - Theory and Applications. L. J. Berliner and J. Reuben, editors. Plenum Press, New York.1–76.

Svensson, M., P. Jonasson, B.-H. Jonsson, M. Lindgren, L.-G. Mårtensson, M. Gentile, K. Borén, and U. Carlsson. 1995. Mapping the folding intermediate of human carbonic anhydrase II. Probing substructure by chemical reactivity and spin and fluorescence labeling of engineered cysteine residues. *Biochemistry.* 34:8606–8620.

Weast, R. C. 1983. CRC Handbook of Chemistry and Physics, 64th ed. CRC Press, Boca Raton, FL. E49–E52.