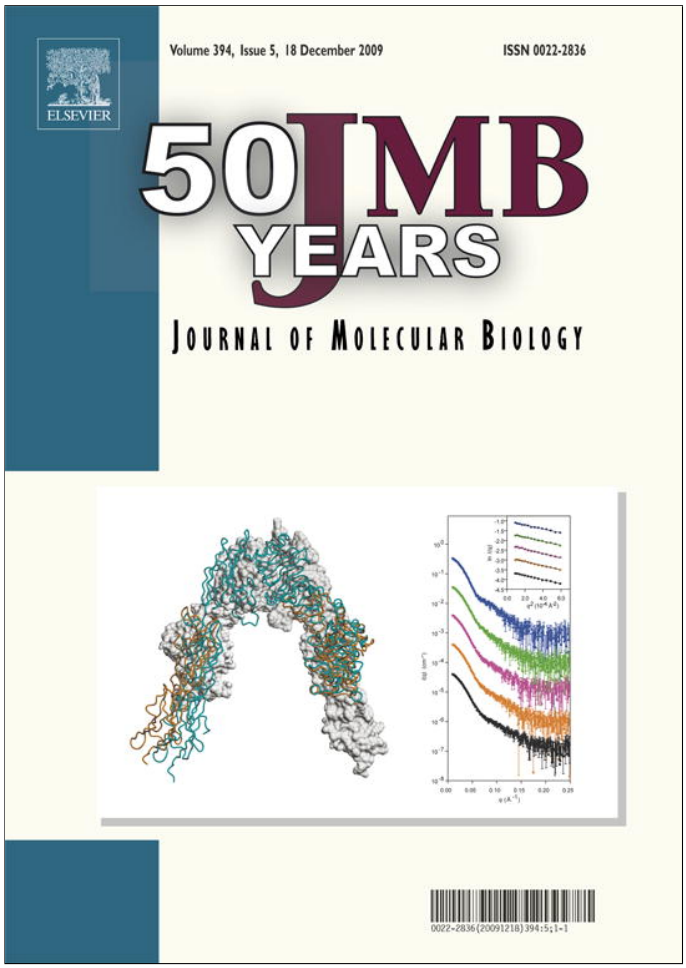


doi:10.1016/j.jmb.2009.09.058 J. Mol. Biol. (2009) 394, 982–993

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Experimentally Approaching the Solvent-Accessible Surface Area of a Protein: Insights into the Acid Molten Globule of Bovine α-Lactalbumin

Patricio O. Craig†, Gabriela E. Gómez†, Daniela B. Ureta†, Julio J. Caramelo† and José M. Delfino*†

Departamento de Química Biológica, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires e Instituto de Química y Fisicoquímica Biológica (Consejo Nacional de Investigaciones Científicas y Técnicas), Junín 956, C1113AAD Buenos Aires, Argentina

Received 18 June 2009; received in revised form 21 August 2009; accepted 25 September 2009
Available online
1 October 2009

Each conformational state of a protein is inextricably related to a defined extent of solvent exposure that plays a key role in protein folding and protein interactions. However, accurate measurement of the solvent-accessible surface area (ASA) is difficult for any state other than the native (N) state. We address this fundamental physicochemical parameter through a new experimental approach based on the reaction of the photochemical reagent diazirine (DZN) with the polypeptide chain. By virtue of its size, DZN is a reasonable molecular mimic of aqueous solvent. Here, we structurally characterize nonnative states of the paradigmatic protein α-lactalbumin. Covalent tagging resulting from unspecific methylene (:CH_2) reaction allows one to obtain a global estimate of ASA and to map out solvent accessibility along the amino acid sequence. By its mild apolar nature, DZN also reveals a hydrophobic phase in the acid-stabilized state of α-lactalbumin, in which there is clustering of core residues accessible to the solvent. In a fashion reminiscent of the N state, this acid-stabilized state also exhibits local regions where increased :CH_2 labeling indicates its nonhomogenous nature, likely pointing to the existence of packing defects. By contrast, the virtual absence of a defined long-range organization brings about a featureless labeling pattern for the unfolded state. Overall, :CH_2 labeling emerges as a fruitful technique that is able to quantify the ASA of the polypeptide chain, thus probing conformational features such as the outer exposed surface and inner cavities, as well as revealing the existence of noncompact apolar phases in nonnative states.

Keywords: diazirine; protein folding; nonnative states; solvent-accessible surface area; α-lactalbumin

Edited by K. Kuwajima

*Corresponding author. E-mail address: delfino@qb.ffyb.uba.ar.

†P.O.C., J.J.C. and J.M.D. are career investigators of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). G.E.G. was a recipient of a graduate student fellowship from the University of Buenos Aires (UBA) and is currently an Estenssoro fellow from the YPF Foundation. D.B.U. is professional technician of CONICET.

Present addresses: P. O. Craig and J. J. Carmelo, Fundación Instituto Leloir e Instituto de Investigaciones Bioquímicas de Buenos Aires (Consejo Nacional de Investigaciones Científicas y Técnicas), Av. Patricias Argentinas 435, C1405BWE, Buenos Aires, Argentina.

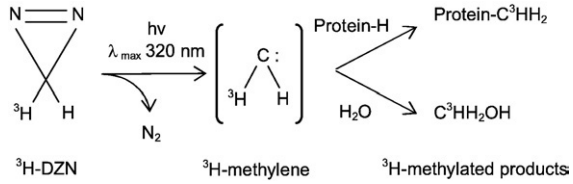
Abbreviations used: ASA, accessible surface area; N, native; DZN, diazirine; :CH_2, methylene; U, unfolded; α-LA, α-lactalbumin; A, acid-stabilized; I, intermediate; [^{3}H]DZN, tritiated derivative of diazirine; CM α-LA, carbamidomethylated α-LA; ANS, 1-anilino-8-naphthalenesulfonic acid; HEWL, hen egg white lysozyme; GdnHCl, guanidine hydrochloride; TFA, trifluoroacetic acid; RP, reversed-phase; PDB, Protein Data Bank; CONICET, Consejo Nacional de Investigaciones Científicas y Técnicas.

Introduction

The polypeptide chain adopts various conformations across a funneled energy landscape en route to the native (N) state.[1] However, not every conformation is likely to be significantly populated; therefore, nonnative states represent a limited repertoire of folded forms that can sometimes be stabilized under appropriate conditions.[2,3] There is much to be learned about the folding process by a detailed characterization of these partially folded forms for which scant information is currently available.

Inextricably associated with the events underlying protein folding or oligomerization is the fact that the accessible surface area (ASA) of the polypeptide chain becomes minimized. Nevertheless, despite being of great theoretical interest,[4,5] there is hardly any experimental method suitable for approaching a direct measurement of ASA. In this regard, parameters such as ΔC_p,[6–9] m value,[10] and ΔR_g,[11] measured along conformational transitions, are thought to be correlated to changes in the accessibility of the whole polypeptide chain, but do not directly yield information along the protein sequence. In a different vein, $^1H/^2H$ exchange of amide protons[12–14] is a powerful technique used to experimentally assess solvent exposure, but data are intrinsically limited to amide protons belonging to the backbone chain, therefore becoming strongly dependent on secondary-structure integrity. Besides, one main drawback arises from the labile nature of the label, a factor that greatly restricts further analytical processing of the sample.

An alternative approach aims at a promiscuous modification of the protein surface by chemical means. Such an approach would expand the set of targets to all solvent-accessible sites, while concurrently obviating chemical selectivity as much as possible, thus generating a map of the topography of the folded chain. Known 'footprinting' reagents belong to this class. A prominent example is the hydroxyl radical (·OH), which has been successfully used to explore the structure and interactions involving nucleic acids[15–17] and proteins.[18–21] However, the ·OH species shows some chemical selectivity in its reaction with peptide targets, thus potentially biasing the analysis to particular spots along the structure. The disruption of the polypeptide chain gives rise to products that are generally stable; however, most often, radical-triggered reactions generate ill-characterized derivatives that complicate ensuing analytical procedures. A related—but essentially different—approach relies on the application of electron-deficient species, such as nitrenes or carbenes, for the modification reaction.[22,23] For the achievement of as close a mimic of the water molecule as geometrically possible, methylene ($:CH_2$), the smallest and most reactive member of this family, emerges as a natural candidate.[24,25] $:CH_2$ inserts readily into any X–H bond (X: C, O, N, or S), generating methylated products that are minimally different in their physicochemical character from the unreacted counterparts (Scheme 1). The best source



Scheme 1. Photolysis of DZN.

of $:CH_2$ is diazirine (DZN; CH_2N_2) because of the ready solubility of this gas in aqueous solutions and its chemical inertness (unless irradiated at $\lambda \sim 320$ nm[26]). Due to its extremely short half-life, $:CH_2$ will react with the target groups belonging to its immediate molecular cage. Therefore, the pattern of labeled sites represents the distribution of DZN prior to the photolysis event. In regard to geometry (size and shape), DZN surveys essentially the same accessible surface probed by the water molecule.

In our laboratory, we have advanced the $:CH_2$ labeling technique to address the unfolding transitions of single-domain proteins[25,27] and the contact surface of an antigen–antibody complex.[28] Other researchers have also applied this method to the conformational study of a heterotrimeric multidomain protein.[29]

In this work, we show how we can exploit the capabilities of $:CH_2$ labeling to explore the structure of a partially folded form of a model protein and to draw comparisons with the N state and the unfolded (U) state. The molten globule state of bovine α-lactalbumin (α-LA) obtained at acidic pH [acid-stabilized (A) state] represents a paradigmatic system that has been characterized by several biochemical and biophysical methods.[2,3,30] Accurate quantification of the $:CH_2$ labeling yield along the protein sequence and its correlation with the extent and nature of the ASA allow us to derive insights into the fine structure of this nonnative state.

Results and Discussion

Conformational-dependent $:CH_2$ labeling of α-LA

α-LA was labeled in the N, U, or intermediate (I) state by reaction with [^{3}H]methylene produced by photolysis of the tritiated derivative of DZN ([^{3}H]DZN) (Table 1). The use of the ^{3}H label allows the measurement of the extent of the reaction even at very low modification levels. Under this condition, the modified protein is mostly monolabeled, minimizing any possible conformational perturbation and limiting the heterogeneity of the products. Under all conditions known to populate the U state, the $:CH_2$ labeling yield is uniformly higher than that observed for the N state, in agreement with the expected larger solvent exposure of the polypeptide chain. The increment is the same, regardless of the nature of the denaturant used or the pH value, a corollary of the indiscriminate chemical reactivity of the $:CH_2$ species. On the other hand, reduction of

Table 1. Conformational-dependent :CH_2 labeling of α-LA

Conformational state	Labeling yield[a]	Ratio[b]
N	3.5±0.5 (44)	1.0
U (8 M urea, pH 7)	5.0±0.3 (8)	1.4
U (8 M urea, pH 2)	4.8±0.1 (2)	1.4
U (8 M urea CM α-LA, pH 7)	5.3±0.4 (2)	1.5
U (6 M GdnHCl, pH 7)	5.0±0.6 (2)	1.4
I (apo state, pH 7)	6.2±1.0 (2)	1.8
I (A state, pH 2)	6.5±0.2 (6)	1.9

Samples of α-LA were incubated under strong denaturing conditions (U state), at pH 2 (A state), or in the apo state (Ca^{2+} depleted) stabilized at neutral pH, in the absence of monovalent cations and at low ionic strength [5 mM ethylenediaminetetraacetic acid and 50 mM Tris/HCl buffer (pH 7.0)]. Reduction and carbamidomethylation of α-LA (CM α-LA sample) were performed as described previously.[31] All samples were photolysed at 25 °C.

[a] The labeling yield is expressed as millimoles of CH_2 incorporated per mole of protein at 1 mM [^{3}H]DZN concentration. Values are expressed as mean±standard deviation. The values in parentheses indicate the number of independent experiments.

[b] Labeling yields are expressed relative to the value measured for the N state.

disulfide bridges and blockage of nascent thiol groups [carbamidomethylated α-LA (CM α-LA) sample] do not cause further enhancement of the :CH_2 labeling. As regards the I states, we assayed (i) that which is prevalent at acidic pH (A state) or (ii) that which is stabilized after removal of structural Ca^{2+} [apo state (Ca^{2+}-free form of α-LA); at neutral pH, at low ionic strength, and in the absence of added monovalent cations].[32] Samples prepared in this fashion show the characteristic features of the molten globule state, as ascertained by the (near) absence of aromatic CD spectra, but preserving far-UV signals (data not shown). In both cases, the :CH_2 labeling yields observed are the highest. In principle, from a geometrical standpoint, one would not expect the I states to present a higher extent of solvent exposure than the U state. However, this result points to the fact that another factor added to the accessibility of the polypeptide chain should be taken into account in interpretation (i.e., the appearance of a 'condensed liquid phase' where DZN becomes concentrated; see below).

The N–A transition of α-LA

The conformational transition N→A is characterized by (i) a cooperative disruption of the tertiary structure of the protein, (ii) an increasing solvent exposure of hydrophobic residues belonging to the core, and (iii) preservation of the secondary-structure content. From an experimental standpoint, these structural changes can be evidenced by (i) a loss of the near-UV CD signal, (ii) an increase in 1-anilino-8-naphthalenesulfonic acid (ANS) binding, and (iii) the conservation of the far-UV CD signal, respectively. Along this transition, the change in the yield of the photoreaction with :CH_2 follows a course parallel with those observed by the techniques described above (Fig. 1a). The midpoint of the conformational change measured with all the methods was around pH 3.3. At or below pH 3, the :CH_2 labeling yield levels off to a value of ~80% above that measured for the N state and, strikingly, also higher than that observed for the U state (cf. Table 1). Thus, the augmented :CH_2 labeling measured for the A state may be the outcome of both (i) a greater accessibility of the polypeptide chain to DZN brought about by its minimal size and (ii) the favorable partition of the reagent (due to its mild hydrophobic nature) into a molten core that might represent a 'newly exposed' condensed fluid phase, thus increasing the local concentration of the :CH_2 precursor.

An important control is provided by hen egg white lysozyme (HEWL), a protein that is closely related in its native conformation to α-LA but does not undergo any conformational change in the same pH range.[33] Consistently, no significant change in the :CH_2 labeling yield occurs (Fig. 1a, inset), thus demonstrating that pH bears no direct influence on the chemistry of the labeling reaction. This observation guarantees that differences in the extent of labeling can be safely interpreted in terms of the conformational characteristics of the protein.

Partition of DZN into the hydrophobic core of the molten globule state would not be impeded because of the fluid nature of this environment. By contrast, this phenomenon would be lower in the N state because of the compactness of the hydrophobic core that would impose a steric restriction to DZN. Thus, the :CH_2 labeling technique appears to be particularly sensitive to the extent of compactness of the polypeptide. On the other hand, the disruption of a contiguous hydrophobic phase in the U state would prevent any partitioning effects of the reagent.

The A–U transition of α-LA

The A state of α-LA can evolve into the U state by the addition of a chaotropic agent such as urea. This A→U transition is characterized by its noncooperative nature and can be monitored by the gradual loss of secondary structure, as measured by far-UV CD.[34] In this regard, the changes in :CH_2 labeling were measured alongside the variation in ellipticity at 222 nm and the binding of the fluorescent probe ANS (Fig. 1b). Here, a steady decrease in labeling is observed in the range 1–5 M urea, amounting to an ~30% loss between states A and U. This observation is consistent with a favorable partition of DZN into the fluid hydrophobic nucleus of state A, a structure that noncooperatively unfolds upon challenge with urea. As evidenced in the range 0–5 M urea, :CH_2 labeling responds primarily to the disruption of the hydrophobic phase, a phenomenon occurring concurrently with the gradual loss of both ANS binding and secondary structure.

In this conformational transition, the labeling yield would be influenced by two phenomena with opposing effects. The increase produced in the accessibility of the polypeptide chain would produce an increase in the labeling yield, whereas the unfolding of the hydrophobic core of the molten

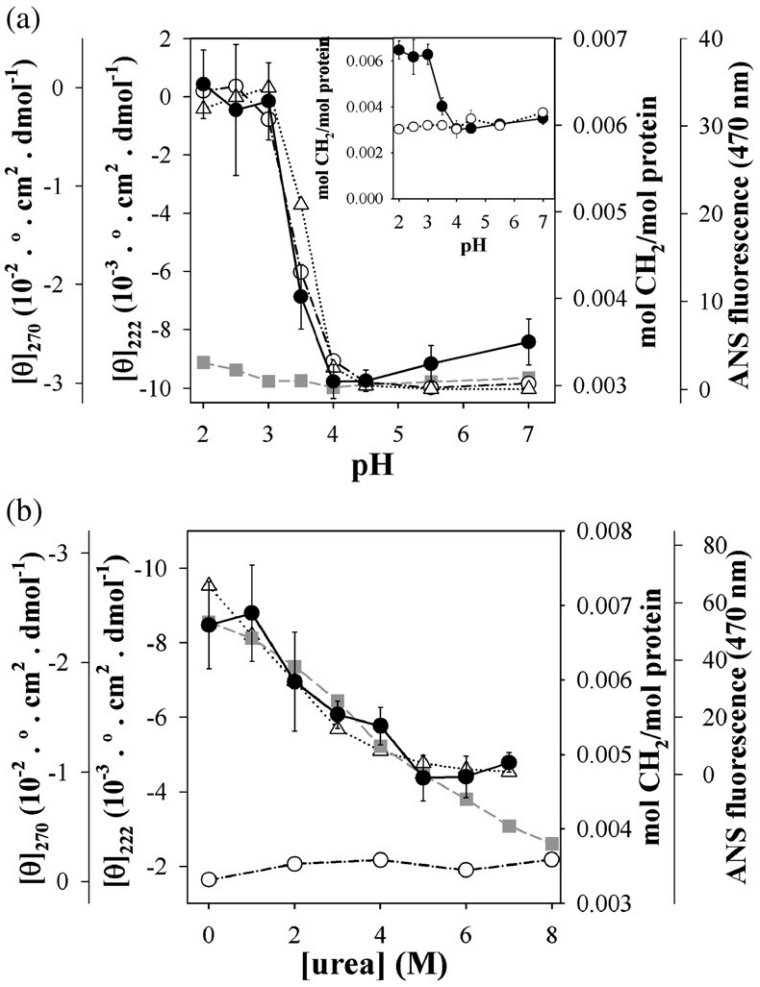


Fig. 1. Conformational transitions of α-LA. α-LA (0.14 mM) was photolabeled with [^{3}H]DZN (0.35 mM [^{3}H]DZN, 0.5 mCi/mmol) in buffer A adjusted to different pH values [pH 2–7; (a) the N→A transition] or adjusted to pH 2 and added with urea up to different final concentrations [0–8 M; (b) the A→U transition] and further processed as indicated in Materials and Methods. The extent of reaction was expressed as the average number of moles of CH_2 per mole of protein at 1 mM [^{3}H]DZN (black filled circles). The molar ellipticity measured at 222 nm (gray filled square) and 270 nm (black open circles) and the fluorescence intensity at 470 nm associated with ANS binding to the protein (black open triangles) are also plotted. The inset in (a) shows the :CH_2 labeling yield of HEWL (open circles) side by side with that measured for α-LA (filled circles) as a function of pH.

globule would decrease [^{3}H]DZN binding to the protein and its contribution to the labeling yield. We observed a decrease in the labeling yield of the protein as a function of urea concentration. This does not rule out the existence of an accessibility increment but indicates that this phenomenon is coupled to the unfolding of the hydrophobic core[35,36] and that this last effect dominates the change in the labeling yield observed during this transition.

In the range 5–7 M urea, the invariance of the labeling yield (i) parallels the virtual disappearance of any fluorescence signal due to bound ANS, supporting the absence of any hydrophobic phase in this condition, and (ii) does not accompany the observed further decrease in ellipticity, indicating that loss of secondary structure by itself does not influence much the solvent accessibility of the polypeptide chain. This goes in line with the fact that the overall exposure of the polypeptide chain is dominated by the exposure of the side chains of amino acids, regardless of the kind of organization imposed on the backbone (ASA calculations not shown).

Peptide analysis of α-LA labeled in the N, A, or U state

To extract site-specific structural information, we measured the radioactivity associated with tryptic peptides derived from α-LA modified under experimental conditions where the N, A, or U state becomes populated (Fig. 2). This protease is an advantageous choice because the mean peptide length (~10 amino acids) allows for adequate averaging over the chemical composition, providing a useful window for discriminating variations due to conformational features rather than variations due to specific chemistry (i.e., avoiding the slight preference of :CH_2 for nucleophilic groups).[24,37] Thus, we obtained a collection of 11 nonoverlapping peptides that span ~93% of the protein sequence.

In a scenario dominated by the unspecific reaction of the :CH_2 species with the polypeptide chain, for differences in the labeling yield of each peptide to be evaluated, results should be expressed as the label incorporated per mass of peptide (mol CH_2/g peptide; alternatively, one could have chosen to express results per mole of amino acid). In this fashion, labeling data will become independent of peptide length. After this normalization, differences in the extent of reaction at a given spot due to a particular microenvironment will become evident. Proof of the unspecific behavior of the labeling event is that all peptides become labeled for any conformational state being considered. For peptides derived from the U state, an almost featureless pattern, showing a constant level of methylation (3.4±

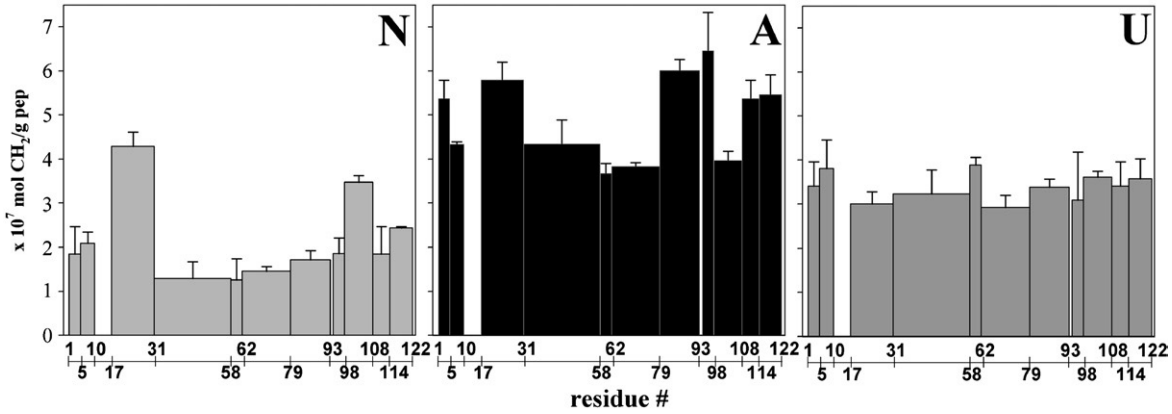


Fig. 2. CH_2 labeling yield of tryptic peptides of α-LA labeled in state N, state A, or state U. α-LA (0.35 mM) was photolabeled with [^{3}H]DZN (3.3 mM, 1 mCi/mmol) in buffer A at pH 7 (N state), at pH 2 (A state), or at pH 7 in the presence of 8 M urea (U state) and further processed as indicated in Materials and Methods. For a direct comparison among the :CH_2 labeling of peptides of different lengths, the label associated with each tryptic peptide is expressed as moles of CH_2 incorporated per gram of peptide at 1 mM [^{3}H]DZN.

0.3×10^{-7} mol CH_2/g peptide; Fig. 2), is observed. This fact points to the notion that, on average, state U would exhibit a similar site accessibility to the reagent, suggesting a homogeneous environment for all targets of reaction. On the other hand, the average labeling yield of peptides was the highest in state A and the lowest in state N, in full agreement with the global estimates of the extent of modification (Table 1). The maps corresponding to these states show (i) variation in the extent of label incorporation among the different peptides and (ii) a qualitatively similar distribution, providing direct evidence on the native-like nature of the A state. Here, the emerging concept is that, despite the presence of a fluid phase, it would not be deprived of local organization.

We then proceeded to plot the absolute :CH_2 labeling yield of the peptides against their theoretical accessibility for both the N state and the U state (Fig. 3). In this fashion, (i) linear correlations with zero ordinate appear obvious and (ii) the same slope is observed, excluding a few 'outlier' peptides where label incorporation in excess of that expected for their ASA occurs only for the N state, where certain sites become prime targets for labeling due to binding of the reagent. As will be discussed below, this excess labeling might find its cause in the involvement of these peptides in the boundary of cavities amenable to lodging the reagent DZN, thus prolonging its residence time and therefore increasing the yield of reaction. A contrasting picture arises from the same analysis applied to peptides derived from the protein labeled in the U state, where no 'outlier' peptides appear, as expected from the 'equalization' of the microenvironments for the reaction, leaving the collisional factor as the only one influencing the :CH_2 labeling. Rather than being peculiar to α-LA, this behavior represents a general phenomenon.[27,28] In this regard, a similar correlation was also demonstrated for full-length proteins of different molecular weights (assayed in the range 14,200–66,400). The slope derived for these proteins in the U state (:CH_2 labeling/ASA: 2.7×10^{-7} to 3.6×10^{-7}, depending on the model chosen for the U state; Fig. 4) is strikingly close to the values observed for tryptic peptides of α-LA in the U state (:CH_2 labeling/ASA: 2.6×10^{-7} to 3.5×10^{-7}; Fig. 3). Nevertheless, the slope of the same set of proteins labeled in the N state (:CH_2 labeling/ASA: 5.8±

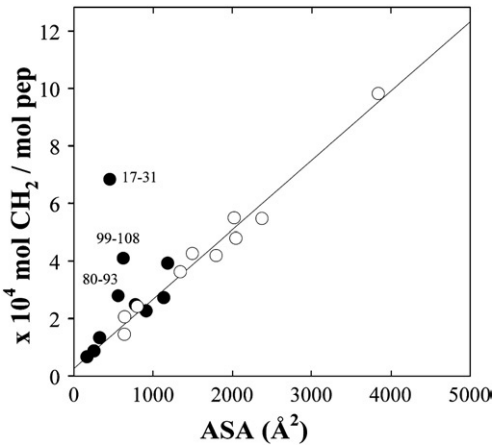


Fig. 3. Correlation plot between :CH_2 labeling yield and ASA. The radioactivity associated with each tryptic peptide derived from α-LA photolabeled in the N state (filled circles) or U state (open circles) is expressed as moles of :CH_2 incorporated per mole of peptide at 1 mM [^{3}H]DZN. For the N state, the ASA corresponding to each peptide was estimated from the crystallographic structures of α-LA (ASA_N); for the U state, an extended-chain model (ASA_{ext}) was used (similar correlations were observed for ASA_{min} or ASA_{max} models). The average numbers of CH_2 groups incorporated per mole of polypeptide per angstrom square of surface at 1 mM [^{3}H]DZN concentration (:CH_2 labeling/ASA) were as follows: $3.3\pm0.3\times10^{-7}$ (N state, excluding peptides 17–31, 80–93, and 99–108), $3.5\pm0.3\times10^{-7}$ (U state, ASA_{min} model), $2.7\pm0.3\times10^{-7}$ (ASA_{max} model), and $2.6\pm0.3\times10^{-7}$ (ASA_{ext} model). R^2 values were 0.90 (N state) and 0.98 (U state).

0.7×10^{-7}; Fig. 4) is the highest, due to the unavoidable inclusion of binding sites for DZN enhancing the overall :CH_2 labeling in folded polypeptides. At this point, a consensus value emerges for :CH_2 labeling/ASA ($\sim3\times10^{-7}$ CH_2 groups incorporated per mole of polypeptide per angstrom square of surface at 1 mM [^{3}H]DZN concentration), representing exclusively the contribution of accessibility to the :CH_2 labeling observed. It is important to underscore the significance of the linear behavior observed because it validates the use of the unspecific photoreaction of DZN to approximate an experimental measurement of the extent of solvent exposure.

A useful interpretation of :CH_2 labeling results stems from the consideration of the surface available for labeling, not only in its extent but also in its nature. Distinctive 'host' sites for the 'guest' molecule DZN, which give rise to enhanced :CH_2 labeling, can be related to its implication in the condensed hydrophobic phase, not excluding the persistence of packing defects (cavities). A prediction of local hydrophobicity based on the sequence[45] at the level of small peptides points to two main sites (around residues 17–31 and 93–108; Fig. 5), the same regions where the :CH_2 labeling yield for the N and A states shows the highest values. On the other hand, it becomes useful to compare the U/N labeling ratio (Table 2) with surface data derived from the known structure of the protein (Fig. 6a). The :CH_2 labeling of most peptides (Fig. 6b, blue or cyan) behaves as expected of a general reagent that labels surfaces primarily in proportion to their exposure. Nevertheless, for peptides 17–31 and 99–108 (in red), a very low correlation was observed. These sites are also implicated in the boundaries of the only empty cavities located in the flexible region including the loop/helix D motif (red blobs). Occupancy of these voids with DZN would not be impeded by the protein because it is known that breathing motions permit access to molecules of this size.[49,50] In fact, small molecules such as methane or cyclopropane occupy two homologous cavities in HEWL, a protein with a topology closely resembling that of α-LA.[51] Indeed, previous work from our laboratory lends support to this contention: tryptic peptides 74–96 derived from HEWL photolabeled with DZN in its N state show an enhanced labeling yield, in agreement with the observation that the backbone atoms of amino acids belonging to this peptide are the main contributors to the environment around the largest cavity present in this protein.[28] In α-LA, residues lining the boundary of these cavities (Fig. 6b, red) belong to the so-called aromatic cluster I[42,46]—sites that would be predicted to accommodate DZN—because of the favorable partition coefficient of this reagent in aromatic solvents.[24]

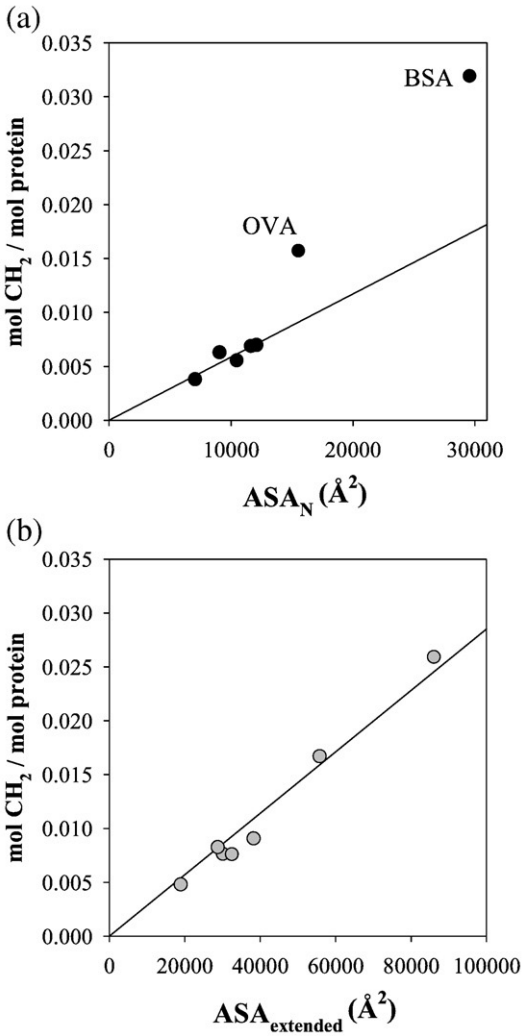


Fig. 4. Correlation plot between :CH_2 labeling yield and ASA. α-LA, bovine growth hormone, β-lactamase from *Bacillus licheniformis*, ovalbumin (OVA), and bovine serum albumin (BSA) were labeled in 40 mM Tris/HCl buffer (pH 7.4), whereas trypsin from bovine pancreas and α-chymotrypsin from bovine pancreas were labeled in 40 mM sodium acetate buffer (pH 4). All proteins were photolabeled under both N conditions (a; black circles) and U conditions (in the presence of 8 M urea) (b; gray circles). Results are expressed as moles of :CH_2 incorporated per mole of protein at 1 mM [^{3}H]DZN. For the N state, the ASA corresponding to each protein (ASA_N) was estimated from their structures: PDB IDs 1OVA,[38] 4BLM,[39] 4CHA,[40] 1TLD,[41] and 1F6S.[42] The structures of BSA and bovine growth hormone were modeled from the structure of their human homologues (PDB IDs 1GNI[43] and 1HGU,[44] respectively). In the N state, OVA and BSA show anomalous values of :CH_2 labeling (in excess of those reported for the rest of the proteins assayed), presumably due to the presence of highly reactive thiol groups in the former and the known ability of the latter to bind hydrophobic molecules. For the U state, an extended-chain model (ASA_{ext}) was used (similar correlations were observed for the ASA_{min} or ASA_{max} models). The average numbers of CH_2 groups incorporated per mole of polypeptide per angstrom square of surface at 1 mM [^{3}H]DZN concentration (:CH_2 labeling/ASA) were as follows: $5.8\pm0.7\times10^{-7}$ (N state, excluding OVA and BSA), $3.6\pm0.4\times10^{-7}$ (U state, ASA_{min} model), $2.7\pm0.3\times10^{-7}$ (U state, ASA_{max} model), and $2.7\pm0.3\times10^{-7}$ (U state, ASA_{ext} model). R^2 values were 0.79 (N state) and 0.98 (U state).

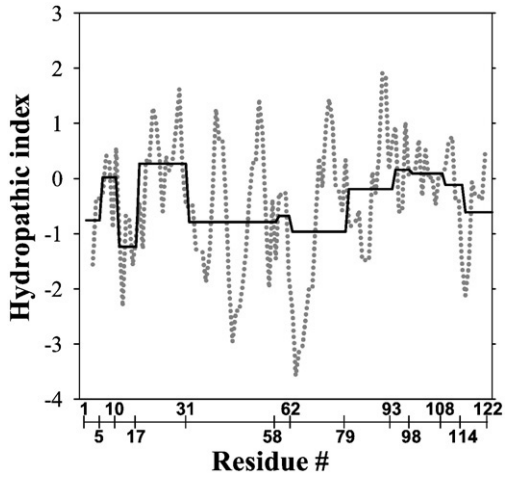


Fig. 5. Hydropathicity plot of α-LA. The hydropathic index of Kyte and Doolittle is plotted for α-LA using a moving window of five residues (dotted gray line).[45] The average hydropathic index for each tryptic peptide is also shown (continuous black line).

Mapping the A state of α-LA

The analysis presented above provides the framework for interpreting the site-specific labeling pattern of the A state of α-LA. Here, there is no standard model that could theoretically describe the solvent accessibility of the polypeptide chain. In this circumstance, the uniform level of labeling observed for the U state provides an adequate reference for comparison given that, in principle, this state would pose an upper limit on site accessibility. Therefore, any excess labeling occurring in the A state could arise as a consequence of other factors added to considerations on mere site accessibility. Thus, we analyzed the A/U labeling ratio for each peptide (Fig. 7).

It is noteworthy that no peptide in the A state appears to incorporate less label than in the U state. The distribution of excess labeling in the A state mapped onto the structure of α-LA is shown in Fig. 6c. Peptides encompassing the β-subdomain of α-LA (peptides 32–58, 59–62, and 63–79) show only a slight enhancement in :CH_2 labeling in the A state. This is fully consistent with a view of the molten globule state where the β-subdomain suffers a major disruption, as evidenced by NMR methods on different variants of α-LA,[52–57] disulfide exchange studies,[58] and high susceptibility to proteolysis.[59,60] As regards the α-subdomain, peptides that show the highest :CH_2 labeling yield in the A state (peptides 17–31, 80–93, and 95–98, encompassing α-helices B and C) are the same as those that have been proposed to participate in the hydrophobic core of this conformational state. This agrees with the notion that there is partial preservation of the secondary structure and topology of this region, stabilized by diffuse hydrophobic interactions among the helices.[52,58,61,62] In this vein, NMR studies of bovine and guinea pig α-LA carried out at pH 2 show low but significant hydrogen-exchange protection factors of the amide groups corresponding to the B helix (residues 23–34) and C helix (residues 86–98),[55,56] suggesting the persistence of these elements in the A state. It is also noteworthy to point out that our results agree well with those reported by D'Silva and Lala, who used the photoactivatable reagent 2-[^{3}H] diazofluorene to map the hydrophobic surface of this state.[63] In that work, they found a high labeling yield of residues V21, W26, T29, V92, I95, K98, and L115, belonging to the B and C helices, and the carboxy-terminal 3_{10} helix of the protein. The remaining structure inferred from our own labeling data and from those of the literature does not exclude the persistence of packing defects akin to those known to exist in the N state (Fig. 6b, red blobs). However, not all peptides belonging to the α-subdomain show an

Table 2. Comparative analysis of the labeling yield and the theoretical accessibility of each tryptic peptide derived from α-LA labeled in the N state or the U state

Peptide	U/N labeling ratio (1)	U/N ASA ratio			Ratio		
		ASA_{min}/ASA_N (2)	ASA_{max}/ASA_N (3)	ASA_{ext}/ASA_N (4)	(1)/(2)	(1)/(3)	(1)/(4)
6–10	1.82	1.88	2.48	2.47	0.97	0.73	0.74
17–31	0.70	3.35	4.30	4.54	0.21	0.16	0.15
32–58	2.50	2.43	3.16	3.25	1.03	0.79	0.77
59–62	3.09	3.24	4.09	3.93	0.95	0.75	0.79
63–79	2.01	1.54	2.01	2.10	1.30	1.00	0.96
80–93	1.97	2.75	3.63	3.64	0.72	0.54	0.54
95–98	1.66	1.95	2.53	2.53	0.85	0.66	0.66
99–108	1.04	1.85	2.37	2.40	0.56	0.44	0.43
1–5+109–114	1.85	1.37	1.78	1.97	1.35	1.04	0.94
115–122	1.46	1.27	1.61	1.73	1.15	0.91	0.84

The U/N labeling ratio for each tryptic peptide was measured as the :CH_2 labeling yield of a sample modified in the U state relative to the corresponding value for a sample of equivalent mass modified in the N state. ASA U/N ratio was calculated as the ASA of each peptide in the U state relative to the corresponding value for the N state. ASA_N was estimated from crystallographic structures of α-LA (PDB IDs 1F6S[42] and 1HFZ[46]). For the U state, ASA_{ext} corresponds to the area of the protein modeled as a fully extended chain, whereas ASA_{max} and ASA_{min} bracket the surface area between a maximum value derived from a Monte Carlo computer simulation of peptides modeled with a hard sphere potential and a minimum value calculated after the analysis of the surface of peptide fragments excised from crystallographic structures.[47,48] Peptides 1–5 and 109–114 were evaluated jointly because they coelute after RP-HPLC.

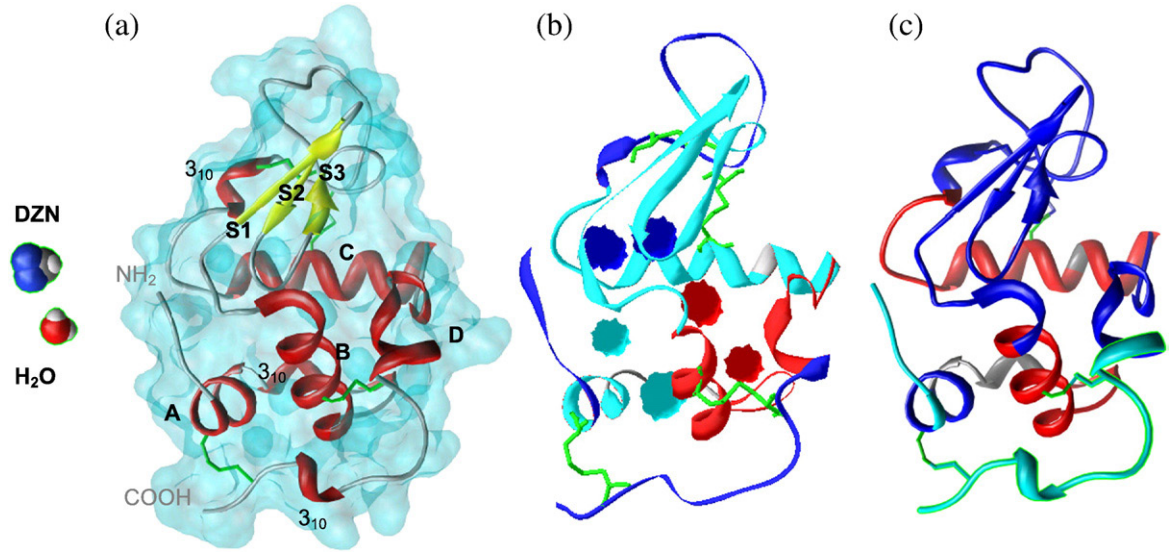


Fig. 6. Structure of α-LA (PDB ID 1F6S[42]). (a) Trace of the backbone highlighting the main secondary-structure elements. The α-subdomain (lower part) includes helix A (residues 5–11), helix B (residues 23–34), helix C (residues 86–98), loop/helix D (residues 105–110), and two small 3_{10} helices (residues 18–20 and 115–118). The β-subdomain (upper part) comprises an anti-parallel β-sheet (residues 41–44, 47–50, and 55–56) and a short 3_{10} helix (residues 77–80). The superimposed translucent surface represents the ASA (probe radius, 1.4 Å) side by side with space-filling models of water and DZN. (b) The structure was colored according to the degree of correlation between the $:CH_2$ labeling changes occurring upon unfolding (U/N labeling ratio) and the theoretical ASA increments (U/N ASA ratio) measured for tryptic peptides derived from the labeled protein (last column in Table 2). Blue peptides include those with correlation values between 0.8 and 1.0 (residues 1–5 + 109–114, 63–79, and 115–122); cyan peptides correspond to those showing values between 0.5 and 0.8 (residues 6–10, 32–58, 59–62, 80–93, and 95–98); and red peptides are those with values below 0.5 (residues 17–31 and 99–108). Colored blobs represent cavities: blue, those occupied by calcium or water; red, those able to lodge DZN; cyan, remainder. (c) The backbone was colored according to the $:CH_2$ labeling change in the A state relative to the U state measured for tryptic peptides derived from the labeled protein (A/U labeling ratio; Fig. 7). Peptides showing the highest labeling enhancement (A/U values in the range 1.8–2.1) are shown in red (residues 17–31, 80–93, and 95–98); those showing intermediate values (in the range 1.5–1.6) are shown in cyan (residues 1–5 + 109–114 and 115–122); those presenting minimal enhancements (A/U values lower than 1.4) are shown in blue (residues 6–10, 32–58, 59–62, 63–79, and 99–108). In all panels, the five disulfide bridges are shown in green.

enhanced $:CH_2$ labeling with respect to the U state: such is the case of peptides 6–10 and 99–108 (mapping to helices A and D, respectively) that would exhibit some extent of structural disorder. This view also finds support in conclusions derived from the exposure of aromatic amino acid residues measured by photochemically induced dynamic nuclear polarization NMR.[35] Independent evidence in support of this observation comes from proteolysis experiments.[59,60] These authors have demonstrated a significant susceptibility to proteolysis of the 95–105 region in the A state of α-LA (although lower than that shown by the β-subdomain). In addition, crystallographic studies on the N state have revealed that this region would also show local disorder, adopting a distorted α-helical conformation at neutral pH (helix D: residues 105–110), while around pH 4 (pretransition N–A), it would be forming a loop exposed to the solvent.[46,64] This phenomenon has been postulated as one of the first denaturation events of the protein induced by pH. In this regard, it would be feasible for the structural disruption of the β-sheet sequentially intercalated between B and C helices to cause a collapse of these helices onto each other, moving farther away the loop/helix D from the structural core.

Prospects and conclusions of the $:CH_2$ labeling method for the structural analysis of proteins

The extremely short half-life of the intermediate $:CH_2$ species (<50 ps) opens the possibility of addressing kinetic processes related to protein folding and assembly. To achieve this goal, one can take advantage of very powerful light sources such as those provided by UV laser beams or synchrotron radiation to photolyse the precursor reagent in very short time windows.

Recent preliminary results of our laboratory indicate that a promising alternative to the use of radiotracers is the development of accurate means to quantify the extent of reaction of the polypeptide chain by modern mass spectrometry techniques (the addition of each $:CH_2$ implies an increment of 14 mass units). This novel approach would be cost-effective and easy to automate and would only require minimal amounts of sample, making it appealing to proteomics studies.

After the experimental parameter $:CH_2$ labeling has been shown to provide a quantitative estimate of ASA, the realization of this technique to address the solvent accessibility of the polypeptide chain in its different conformational states appears particularly

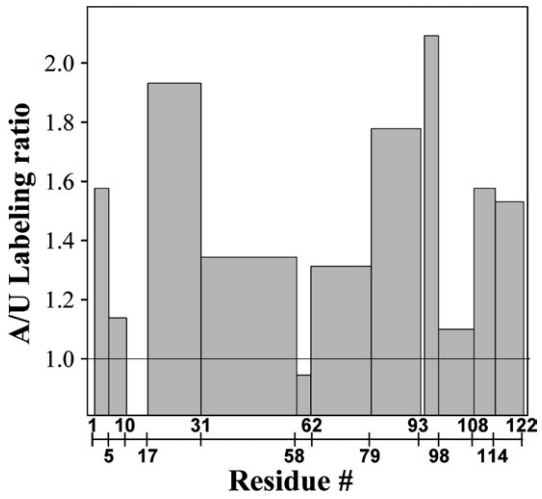


Fig. 7. Regiospecific :CH_2 labeling pattern of state A of α-LA. The A/U labeling ratio for each tryptic peptide was measured as the :CH_2 labeling yield of a sample modified in the A state relative to the corresponding value for a sample of equivalent mass modified in the U state.

attractive, especially so in the case of nonnative states that become intractable by other biophysical or biochemical methods. In this regard, the external surface will be the prime target of :CH_2 labeling, as well as inner regions that constitute a special cage for the reagent DZN (cavities) and noncompact apolar phases in nonnative states.

Materials and Methods

Materials

[^{3}H]Formaldehyde (1–5 mCi, 90.0 mCi/mmol) was purchased from New England Nuclear or ARC; 37% (wt/vol) formaldehyde was obtained from E. Merck; and formamide, urea, guanidine hydrochloride (GdnHCl), and bovine α-LA (type I) or apo α-LA (type III) were obtained from Sigma Chemical Co. α-LA was used without further purification, and urea was recrystallized from ethanol before use. L-1-Tosylamido-2-phenylethyl chloromethyl ketone trypsin was obtained from Worthington. Acetonitrile (E. Merck) and trifluoroacetic acid (TFA) (Riedel de Haën) were of HPLC grade. All other reagents and chemicals used were of analytical grade. α-LA concentration was determined by its UV absorption at 280 nm on a Jasco 7850 spectrophotometer using an extinction coefficient of $E_{1\%}=20.1$.[65] α-LA and their tryptic peptides were separated using an ÄKTA purifier (Amersham Pharmacia Biotech) and Rainin Dynamax HPLC systems.

Sample preparation

Samples of α-LA in Table 1 were incubated under strong denaturing conditions (8 M urea or 6 M GdnHCl; U state) at pH 2 (A state) or in a 5 mM ethylenediaminetetraacetic acid–50 mM Tris/HCl buffer at pH 7 (apo state). Reduction and carbamidomethylation of α-LA (CM α-LA sample) were performed in accordance with established protocols.[31] α-LA in Fig. 1 was equilibrated for 1 h under nitrogen under different conditions of pH (2–7) or urea (0–8 M) in degassed 40 mM glycine, 40 mM sodium acetate, and 40 mM Tris/HCl buffer (buffer A).

Circular dichroism

Spectra of α-LA (70–140 μM) were recorded on a Jasco J-810 spectropolarimeter using quartz cylindrical cuvettes of 1-mm or 10-mm path lengths for the far-UV (200–250 nm) and near-UV (250–310 nm) regions, respectively. Data were converted into molar ellipticity $[\theta]_M$ (deg cm^2 dmol^{-1}) using a mean residue weight value of 115.5 g/mol for α-LA.

Fluorescence spectroscopy

Measurements were carried out with an Aminco Bowman Series II spectrofluorometer, using a quartz cuvette of 3-mm path length, on α-LA samples (50–140 μM) incubated for 1 h with ANS (44–51 μM). The fluorescence emission in the range 450–550 nm was collected using an excitation wavelength of 380 nm and bandwidths of 4 nm.

Synthesis of DZN

[^{3}H]DZN (0.2–1 mCi/mmol) was synthesized in a microscale setup, as described previously.[25] [^{3}H]DZN concentration in aqueous solution was estimated by measuring (i) the absorbance at 320 nm ($\varepsilon_{320}=180$ M^{-1} cm^{-1}[25]) and (ii) the concentration of radioactivity after complete photolysis of the reagent.

Photolabeling of α-LA

[^{3}H]DZN (0.3–3.3 mM, 0.2–1 mCi/mmol) was dissolved with α-LA (0.14–0.35 mM) in capped quartz cuvettes of 1-cm path length. Photolysis was carried out in a water bath at 20 °C using a UV light source (Philips HPA 1000 lamp placed 12 cm from the samples) filtered below 300 nm (Oriel 59044 long-pass filter). UV irradiation was extended for ~4 half-lives ($t_{1/2\ DZN}=10.3$ min in our photolysis setup). The workup procedure consisted of addition of urea (up to 8 M), followed by dialysis against 0.07 mM Na_2CO_3 (pH 7.5–8.0) and freeze drying. Finally, the samples were dissolved in 6 M GdnHCl in 0.05% aqueous TFA, purified by reversed-phase (RP) HPLC on a C4 column (Vydac 214TP510), developed with a linear gradient of acetonitrile/water (0–60% in 10 min) in 0.05% TFA, and followed by an isocratic elution with 60% acetonitrile in 0.05% TFA (5 min) at a flow rate of 3 mL/min. Elution was monitored by UV absorption at 280 nm. The protein was collected, freeze dried, and redissolved in 20 mM sodium phosphate buffer (pH 7.4) to measure the radioactivity on a Pharmacia 1214 Rackbeta liquid scintillation counter. The extent of :CH_2 labeling was expressed as the average number of moles of CH_2 per mole of protein by taking into account the sample concentration and the specific radioactivity of the reagent normalized at 1 mM [^{3}H]DZN.

Reduction, carbamidomethylation, and tryptic digestion of α-LA

After the cleanup procedure described in Photolabeling of α-LA, reduction and carbamidomethylation of α-LA were performed in accordance with Waxdal *et al.*[31]

Peptide mapping

Tryptic peptides derived from α-LA samples labeled in the N, A, or U state were separated by size-exclusion chromatography on a Superdex Peptide HR 10/30 column (GE Healthcare) using a 4 M urea–50 mM Tris/HCl (pH 7.0) buffer at a flow rate of 0.5 mL/min. Elution was monitored by both UV absorption at 215 nm and radioactivity associated with fractions collected at 30-s time intervals. Collected peaks were further separated by RP-HPLC on a C18 column (Vydac 218TP54) using a linear gradient of acetonitrile/water (0–60% in 60 min) in 0.05% TFA at a flow rate of 1 mL/min. Elution was monitored by UV absorption at 215 and 280 nm and by the radioactivity associated with fractions collected at 1-min time intervals. Peptide identification was achieved by amino acid analysis on an Applied Biosystems 420A. The $:CH_2$ labeling yield of peptides was calculated by the ratio of the radioactivity associated with each peak and the amount of peptide eluted.

Molecular modeling

The ASAs of α-LA and its peptides in the N and U states were calculated with MSP[66] using a probe radius of 1.4 Å. The analysis of the N state was based on crystallographic structures available for the wild-type protein [Protein Data Bank (PDB) ID 1F6S[42]] and for an M90V mutant variant (PDB ID 1HFZ[46]), whereas that of the U state was based on an extended-chain model (ASA_{ext}) constructed with MacroModel.[67] Additional ASA estimates of the U state (ASA_{max} and ASA_{min}) were obtained through a Web resource‡.[47,48] Cavities were calculated with MSP on the structure PDB ID 1F6S (devoid of Ca^{2+} and water) using a probe radius of 1.3 Å. Figure 6b and c were built with Swiss-PdbViewer[68] and Chimera,[69] respectively, running on a personal computer under Windows XP. MacroModel and MSP were run on SGI workstations (Indigo XS24Z R4000 and O2 R10000 under Irix 5.3 or 6.5, respectively).

Acknowledgements

P.O.C., J.J.C., and J.M.D. are career investigators of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). G.E.G. was a recipient of a graduate student fellowship from the University of Buenos Aires and is currently an Estenssoro fellow at the YPF Foundation. D.B.U. is a professional technician at CONICET.

This research has been supported by grants to J.M.D. from the University of Buenos Aires, CONICET, and the Agencia Nacional de Promoción Científica y Tecnológica.

References

1. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins*, **21**, 167–195.
2. Ptitsyn, O. B. (1995). Molten globule and protein folding. *Adv. Protein Chem.* **47**, 83–229.
3. Arai, M. & Kuwajima, K. (2000). Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**, 209–282.
4. Lee, B. & Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.
5. Richards, F. M. (1977). Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.
6. Makhatadze, G. I. & Privalov, P. L. (1990). Heat capacity of proteins: I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* **213**, 375–384.
7. Privalov, P. L. & Makhatadze, G. I. (1990). Heat capacity of proteins: II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects. *J. Mol. Biol.* **213**, 291–385.
8. Livingstone, J. R., Spolar, R. S. & Record, M. T., Jr (1991). Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. *Biochemistry*, **30**, 4237–4244.
9. Loladze, V. V., Ermolenko, D. N. & Makhatadze, G. I. (2001). Heat capacity changes upon burial of polar and nonpolar groups in proteins. *Protein Sci.* **10**, 1343–1352.
10. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
11. Geierhaas, C. D., Nickson, A. A., Lindorff-Larsen, K., Clarke, J. & Vendruscolo, M. (2007). BPPred: a Web-based computational tool for predicting biophysical parameters of proteins. *Protein Sci.* **16**, 125–134.
12. Englander, S. W. & Krishna, M. M. G. (2001). Hydrogen exchange. *Nat. Struct. Biol.* **8**, 741–742.
13. Truhlar, S. M., Croy, C. H., Torpey, J. W., Koeppe, J. R. & Komives, E. A. (2006). Solvent accessibility of protein surfaces by amide $H/^2H$ exchange MALDI-TOF mass spectrometry. *J. Am. Soc. Mass Spectrom.* **17**, 1490–1497.
14. Englander, S. W. (2006). Hydrogen exchange and mass spectrometry: a historical perspective. *J. Am. Soc. Mass Spectrom.* **17**, 1481–1489.
15. Tullius, T. D. & Dombroski, B. A. (1985). Iron (II) EDTA used to measure the helical twist along any DNA molecule. *Science*, **230**, 679–681.
16. Tullius, T. D., Dombroski, B. A., Churchill, M. E. A. & Kam, L. (1989). Hydroxyl radical footprinting: a high resolution method for mapping protein–DNA contacts. In *Recombinant DNA Methodology* (Wu, R., Grossman, L. & Moldave, K., eds), Academic Press, San Diego, CA.
17. Jain, S. S. & Tullius, T. D. (2008). Footprinting protein–DNA complexes using the hydroxyl radical. *Nat. Protoc.* **3**, 1092–1100.
18. Maleknia, S. D., Brenowitz, M. & Chance, M. R. (1999). Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. *Anal. Chem.* **71**, 3965–3973.
19. Maleknia, S. D., Ralston, C. Y., Brenowitz, M. D., Downard, K. M. & Chance, M. R. (2001). Determination of macromolecular folding and structure by synchrotron X-ray radiolysis techniques. *Anal. Biochem.* **289**, 103–115.
20. Maleknia, S. D. & Downard, K. M. (2001). Unfolding of apomyoglobin helices by synchrotron radiolysis and mass spectrometry. *Eur. J. Biochem.* **268**, 5578–5588.
21. Takamoto, K. & Chance, M. R. (2006). Radiolytic protein footprinting with mass spectrometry to probe

‡ http://roselab.jhu.edu/utils/unfolded.html

the structure of macromolecular complexes. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 251–276.
22. Brunner, J. (1993). New photolabeling and crosslinking methods. *Annu. Rev. Biochem.* **62**, 483–514.
23. Delfino, J. M., Schreiber, S. L. & Richards, F. M. (1993). Design, synthesis and properties of a photoactivatable membrane-spanning phospholipidic probe. *J. Am. Chem. Soc.* **115**, 3458–3474.
24. Richards, F. M., Lamed, R., Wynn, R., Patel, D. & Olack, G. (2000). Methylene as a possible universal footprinting reagent that will include hydrophobic surface areas: overview and feasibility: properties of diazirine as a precursor. *Protein Sci.* **9**, 2506–2517.
25. Craig, P. O., Ureta, D. B. & Delfino, J. M. (2002). Probing protein conformation with a minimal photochemical reagent. *Protein Sci.* **11**, 1353–1366.
26. Liu, M. T. H. (1982). The thermolysis and photolysis of diazirines. *Chem. Soc. Rev.* **11**, 127–140.
27. Ureta, D. B., Craig, P. O., Gómez, G. E. & Delfino, J. M. (2007). Assessing native and non-native conformational states of a protein by methylene carbene labeling: the case of *Bacillus licheniformis* beta-lactamase. *Biochemistry*, **46**, 14567–14577.
28. Gómez, G. E., Cauerhff, A., Craig, P. O., Goldbaum, F.A. & Delfino, J. M. (2006). Exploring protein interfaces with a general photochemical reagent. *Protein Sci.* **15**, 744–752.
29. Nuss, J. E. & Alter, G. M. (2004). Denaturation of replication protein A reveals an alternative conformation with intact domain structure and oligonucleotide binding activity. *Protein Sci.* **13**, 1365–1378.
30. Kuwajima, K. (1996). The molten globule state of α-lactalbumin. *FASEB J.* **10**, 102–109.
31. Waxdal, M. J., Konigsberg, W. H., Henley, W. L. & Edelman, G. M. (1968). The covalent structure of a human gamma G-immunoglobulin: II. Isolation and characterization of the cyanogen bromide fragments. *Biochemistry*, **7**, 1959–1966.
32. Yutani, K., Ogasahara, K. & Kuwajima, K. (1992). Absence of the thermal transition in apo-alpha-lactalbumin in the molten globule state. A study by differential scanning microcalorimetry. *J. Mol. Biol.* **228**, 347–350.
33. Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. (1972). In *The Enzymes* (Boyer, P. D., ed.), *The Enzymes*, vol. 7, pp. 665–868. Academic Press, New York, NY.
34. Shimizu, A., Ikeguchi, M. & Sugai, S. (1993). Unfolding of the molten globule state of α-lactalbumin studied by ^{1}H-NMR. *Biochemistry*, **32**, 13198–13203.
35. Mok, K. H., Nagashima, T., Day, I. J., Hore, P. J. & Dobson, C. M. (2005). Multiple subsets of side-chain packing in partially folded states of alpha-lactalbumins. *Proc. Natl Acad. Sci. USA*, **102**, 8899–8904.
36. Chakraborty, S., Ittah, V., Bai, P., Luo, L., Haas, E. & Peng, Z. (2001). Structure and dynamics of the α-lactalbumin molten globule: fluorescence studies using proteins containing a single tryptophan residue. *Biochemistry*, **40**, 7228–7238.
37. Turro, N. J., Cha, Y. & Gould, I. R. (1987). Reactivity and intersystem crossing of singlet methylene in solution. *J. Am. Chem. Soc.* **109**, 2101–2107.
38. Stein, P. E., Leslie, A. G. W., Finch, J. T. & Carrell, R. W. (1991). Crystal structure of uncleaved ovalbumin at 1.95 Å resolution. *J. Mol. Biol.* **221**, 941–959.
39. Knox, J. R. & Moews, P. C. (1991). Beta-lactamase of *Bacillus licheniformis* 749/C. Refinement at 2 Å resolution and analysis of hydration. *J. Mol. Biol.* **220**, 435–455.
40. Tsukada, D. M. & Blow, D. M. (1985). Structure of alpha-chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol.* **184**, 703–711.
41. Bartunik, L. J., Summers, L. J. & Bartsch, H. H. (1989). Crystal structure of bovine beta-trypsin at 1.5 Å resolution in a crystal form with low molecular packing density. Active site geometry, ion pairs and solvent structure. *J. Mol. Biol.* **210**, 813–828.
42. Chrysina, E. D., Brew, K. & Acharya, K. R. (2000). Crystal structures of apo- and holo-bovine α-lactalbumin at 2.2-Å resolution reveal an effect of calcium on inter-lobe interactions. *J. Biol. Chem.* **275**, 37021–37029.
43. Petitpas, I., Gruene, T., Bhattacharya, A. A. & Curry, S. (2001). Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *J. Mol. Biol.* **314**, 955–960.
44. Chantalat, L., Jones, N. D., Korber, F., Navaza, J. & Pavlovsky, A. G. (1995). The crystal structure of wild type growth-hormone at 2.5 Å resolution. *Protein Pept. Lett.* **2**, 333.
45. Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132.
46. Pike, A. C. W., Brew, K. & Acharya, K. R. (1996). Crystal structures of guinea-pig, goat and bovine α-lactalbumin highlight the enhanced conformational flexibility of regions that are significant for its action in lactose synthase. *Structure*, **4**, 691–703.
47. Creamer, T. P., Srinivasan, R. & Rose, G. D. (1995). Modeling unfolded states of peptides and proteins. *Biochemistry*, **34**, 16245–16250.
48. Creamer, T. P., Srinivasan, R. & Rose, G. D. (1997). Modeling unfolded states of proteins and peptides: II. Backbone solvent accessibility. *Biochemistry*, **36**, 2832–2835.
49. Karplus, M. (2003). Molecular dynamics of biological macromolecules: a brief history and perspective. *Biopolymers*, **68**, 350–358.
50. Lakowicz, J. R. & Weber, G. (1973). Quenching of protein fluorescence by oxygen. Detection of structural fluctuations in proteins on the nanosecond time scale. *Biochemistry*, **12**, 4171–4179.
51. Otting, G., Liepinsh, E., Halle, B. & Frey, U. (1997). NMR identification of hydrophobic cavities with low water occupancies in protein structures using small gas molecules. *Nat. Struct. Biol.* **4**, 396–404.
52. Wijesinha-Bettoni, R., Dobson, C. M. & Redfield, C. (2001). Comparison of the denaturant-induced unfolding of the bovine and human α-lactalbumin molten globule. *J. Mol. Biol.* **312**, 261–273.
53. Baum, J., Dobson, C. M., Evans, P. A. & Hanley, C. (1989). Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry*, **28**, 7–13.
54. Schulman, B. A., Kim, P. S., Dobson, C. M. & Redfield, C. (1997). A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat. Struct. Biol.* **4**, 630–634.
55. Forge, V., Wijesinha, R. T., Balbach, J., Brew, K., Robinson, C. V., Redfield, C. & Dobson, C. M. (1999). Rapid collapse and slow structural reorganisation during the refolding of bovine alpha-lactalbumin. *J. Mol. Biol.* **288**, 673–688.
56. Chyan, C., Wormald, C., Dobson, C. M., Evans, P. A. & Baum, J. (1993). Structure and stability of the molten globule state of guinea-pig α-lactalbumin: a hydrogen exchange study. *Biochemistry*, **32**, 5681–5691.
57. Schulman, B. A., Redfield, C., Peng, A., Dobson, C. M. & Kim, P. S. (1995). Different subdomains are most

protected from hydrogen exchange in the molten globule and native states of human α-lactalbumin. *J. Mol. Biol.* **253**, 651–657.
58. Wu, L., Peng, Z. & Kim, P. S. (1995). Bipartite structure of α-lactalbumin molten globule. *Nat. Struct. Biol.* **2**, 281–285.
59. Polverino de Laureto, P., De Filippis, V., Di Bello, M., Zambonin, M. & Fontana, A. (1995). Probing the molten globule state of α-lactalbumin by limited proteolysis. *Biochemistry*, **34**, 12596–12604.
60. Polverino de Laureto, P., Scaramella, E., Frigo, M., Wondrich, F. G., de Filippis, V., Zambonin, M. & Fontana, A. (1999). Limited proteolysis of bovine α-lactalbumin and characterization of protein domains. *Protein Sci.* **8**, 2290–2303.
61. Peng, Z. & Kim, P. S. (1994). A protein dissection study of a molten globule. *Biochemistry*, **33**, 2136–2141.
62. Peng, Z., Wu, L. C. & Kim, P. S. (1995). Local structural preferences in the α-lactalbumin molten globule. *Biochemistry*, **34**, 3248–3252.
63. D'Silva, P. R. & Lala, A. K. (1999). Hydrophobic photolabeling as a new method for structural characterization of molten globule and related protein folding intermediates. *Protein Sci.* **8**, 1099–1103.
64. Harata, K. & Muraki, M. (1992). X-ray structural evidence for a local helix–loop transition in α-lactalbumin. *J. Biol. Chem.* **267**, 1419–1421.
65. Kronman, M. J. & Andreotti, R. E. (1964). Inter and intramolecular interactions of α-lactalbumin: I. The apparent heterogeneity at acid pH. *Biochemistry*, **3**, 1145–1151.
66. Connolly, M. L. (1993). The molecular surface package. *J. Mol. Graphics*, **11**, 139–141.
67. Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C. *et al.* (1990). MacroModel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **11**, 440–467.
68. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
69. Huang, C. C., Couch, G. S., Pettersen, E. F. & Ferrin, T. E. (1996). Chimera: an extensible molecular modeling application constructed using standard components. *Pac. Symp. Biocomput.* **1**, 724.