

Hydrophobic photolabeling as a new method for structural characterization of molten globule and related protein folding intermediates

PATRICK R. D'SILVA AND ANIL K. LALA

Biomembrane Laboratory, Department of Chemistry and Biotechnology Center, Indian Institute of Technology Bombay, Powai, Bombay 400 076, India

(RECEIVED August 31, 1998; ACCEPTED January 29, 1999)

Abstract

Recent advances in attempts to unravel the protein folding mechanism have indicated the need to identify the folding intermediates. Despite their transient nature, in a number of cases it has been possible to detect and characterize some of the equilibrium intermediates, for example, the molten globule (MG) state. The key features of the MG state are retention of substantial secondary structure of the native state, considerable loss of tertiary structure leading to increased hydrophobic exposure, and a compact structure. NMR, circular dichroism, and fluorescence spectroscopies have been most useful in characterizing such intermediates. We report here a new method for structural characterization of the MG state that involves probing the exposed hydrophobic sites with a hydrophobic photoactivable reagent—2[³H]diazofluorene. This carbene-based reagent binds to hydrophobic sites, and on photolysis covalently attaches itself to the neighboring amino acid side chains. The reagent photolabels α -lactalbumin as a function of pH (3–7.4), the labeling at neutral pH being negligible and maximal at pH 3. Chemical and proteolytic fragmentation of the photolabeled protein followed by peptide sequencing permitted identification of the labeled residues. The results obtained indicate that the sequence corresponding to B (23–34) and C (86–98) helix of the native structure are extensively labeled. The small β -domain (40–50) is poorly labeled, Val42 being the only residue that is significantly labeled. Our data, like NMR data, indicate that in the MG state of α -lactalbumin, the α -domain has a greater degree of persistent structure than the β -domain. However, unlike the NMR method, the photolabeling method is not limited by the size of the protein and can provide information on several new residues, for example, Leu115. The current method using DAF thus allows identification of stable and hydrophobic exposed regions in folding intermediates as the reagent binds and on photolysis covalently links to these regions.

Keywords: bovine α -lactalbumin; diazofluorene; hydrophobic photolabeling; molten globule state; protein folding intermediates

One of the major difficulties in investigation of protein folding is identification and structural characterization of the intermediate states that are often transient in nature. However, in some cases it has been possible to detect equilibrium intermediates, which retain partial native-like structure. A variety of experimental methods sensitive to different levels of protein structure (Kim & Baldwin, 1990) have been used to identify the intermediates formed at equi-

librium on folding or unfolding of globular proteins. The use of these contemporary methodologies has allowed detection of equilibrium intermediate states, the most notable among them being the molten globule (MG) state (Ohgushi & Wada, 1983; Ptitsyn, 1995). This state is characterized by partial retention of the secondary structure of the native state but with loss of the tertiary structure, giving rise to a compact structure without rigid packing inside the molecule. Further, there is a substantial increase in the fluctuations of side chains leading to increased hydrophobic exposure. Although the initial characterization of such intermediates took place during studies on unfolding or folding of proteins induced by change in environmental conditions; for example, salt, pH, denaturant, current studies have indicated recognition of such intermediates by chaperones (Martin et al., 1991; Katsumata et al., 1996; Lindner et al., 1997). However, the characterization of such intermediates, though crucial to understanding the protein folding

Reprint requests to: Anil Lala, Department of Chemistry and Biotechnology, Indian Institute of Technology Bombay, Powai, Bombay 400 076, India; e-mail: anillala@chem.iitb.ernet.in.

Abbreviations: α -LA, bovine α -lactalbumin; CAPS, 3-[cyclohexylamino]-1-propanesulfonic acid; CD, circular dichroism; CNBr, cyanogen bromide; DAF, 2-[³H]diazofluorene; MG, molten globule; PTH-AA, phenyl thiohydantoin amino acid; PVDF, polyvinyl difluorobenzene; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)-ethyl]glycine.

process, has been achieved only to a limited extent primarily due to limitations of the methodologies involved. Thus, although CD and fluorescence provide a macroscopic picture, the structural details have been delineated largely by use of NMR spectroscopy using the H-D exchange of the exposed amide NH groups of the protein backbone (Kim & Baldwin, 1990).

We report here a novel method for characterization of MG state. The method being reported is similar in approach to the fluorescence method involving use of hydrophobic polarity sensitive probes (Semisotnov et al., 1991; Engelhard & Evans, 1995). The fluorescent probes, for example, ANS, bind to exposed hydrophobic sites leading to increase in fluorescent intensity and often a blue shift, indicating increased hydrophobic exposure in the MG state. The fluorescence method, although extremely useful, does not provide any information on the sites or domains to which the probe binds. In the current method we use a hydrophobic photoactivable reagent—2[³H]diazofluorene (DAF)—that binds to hydrophobic sites in a protein. On photolysis, DAF generates a highly reactive carbene called fluorenylidene. A carbene by definition being an electron deficient species seeks the first available electron-rich center to complete its sextet. Thus, in bulk water it will readily insert into O-H bond of water to give fluorenol. However, in a nonpolar environment like a hydrocarbon solvent, it inserts even into a C-H bond (Baron et al., 1973; Pradhan & Lala, 1987; Lala & Kaul, 1992). Fluorenylidene further equilibrates between singlet and triplet states, as its energy barrier between these two states is very low (Grasse et al., 1983; Ruzicka et al., 1992). Such a carbene can thus insert into a wide variety of bonds ranging from C-H, N-H, O-H, and S-H. Among several other points relating to design of hydrophobic photoactivable reagents, the choice of DAF as a reagent was dictated by the fact that carbene resulting on photolysis of DAF rapidly equilibrates between the singlet and triplet state (Anjaneyulu et al., 1984; Anjaneyulu & Lala, 1984; Pradhan & Lala, 1987). DAF, being hydrophobic, thus binds to hydrophobic sites in a protein, for example, a membrane-spanning protein, and on photolysis covalently links to transmembrane segments of these proteins (Pradhan & Lala, 1987; Lala & Bhat, 1990). Here we exploit the hydrophobic nature of DAF and its ability to insert into side chains of amino acids on photolysis, to map hydrophobic surface in proteins. Thus, on incubation with a protein or a folding intermediate with hydrophobic sites, DAF binds to these sites and on photolysis covalently links to the amino acid side chains present at these sites. Subsequent chemical and/or enzymatic fragmentation of the photolabeled protein followed by peptide sequencing can provide useful information about the hydrophobic sites. To evaluate our method we have used bovine α -LA, primarily as this is one of the most extensively studied protein by various techniques and the low pH-induced MG state is probably best characterized in this case (Kuwajima, 1996). Further, α -LA has been extensively studied using amide NH H-D exchange NMR studies (Alexandrescu et al., 1993; Chyan et al., 1993; Schulman et al., 1995, 1997), thus allowing us to compare the structural details on the MG state obtained by NMR with the data reported here.

Results and discussion

[³H]DAF was incubated with bovine α -LA at pH 3 in the dark and then photolyzed with the light of the wavelength greater than 350 nm. Analysis of DAF-labeled α -LA indicated that DAF was cova-

lently linked to α -LA, and the bound radioactivity could not be dissociated from α -LA by solvent extraction or during SDS-PAGE. In a control experiment carried out in the dark, DAF was not found to be associated with α -LA, indicating that the observed labeling was photochemical in nature. The photolabeling of α -LA was pH dependent, with a very low level of labeling at pH 7 when compared to pH 3 (Lala & Kaul, 1992). To see if DAF binding to α -LA leads to any structural changes, we recorded fluorescence spectra of DAF-bound α -LA and free α -LA at pH 3. The fluorescence spectra in either case were found to be identical. It may be added here that DAF is nonfluorescent. In another control experiment, α -LA was incubated with DAF in the presence of 7 M guanidine-HCl and photolyzed. None or negligible photolabeling by DAF was observed. These results suggest that as the pH is lowered and α -LA unfolds to an MG state, the changes associated with tertiary structure provide sufficient hydrophobic sites for DAF to bind and label it on photolysis. To obtain structural information on these hydrophobic sites, the DAF-labeled α -LA was subjected to protein sequencing. The PTH-amino acid analysis was then carried out and radioactivity associated with each cycle determined (Fig. 1). The protein could be sequenced up to 50 cycles, although definitive assignments for PTH-amino acid released could be made only up to 45 cycles, beyond which it was difficult to make all the assignments. Radioactivity associated with each cycle was counted. The major amino acid residues labeled are Val21, Trp26, Thr29, and Val42. The most heavily labeled amino acid in this part of the molecule is Trp26. The crystal structure of bovine α -LA indicates that it has two hydrophobic helices, B (23–34) and C (86–98), which are in the protein interior and close to each other (Pike et al., 1996). The two helices are connected to each other in the intervening region by small β -sheets (41–44, 47–50), a small helix (77–80), and loop regions. The H-D exchange studies carried out on guinea pig α -LA at pH 2 by NMR indicate that the most protected amides are in the hydrophobic B (23–34) and C (86–98) helices (Chyan et al., 1993). These results suggest that this region of the molecule in the MG state is most persistent and native-like in the MG state. Most typically in this region, the Trp26 has an indole NH that is highly protected from proton exchange. The high protection factor toward the H-D exchange in the MG state for this residue (Trp26) and others like Ile29, Ile30, Ile95, and Leu96 in the B- and C-helix region strongly support that these two helices are most persistent in retaining their native-like structure conformation in the MG state. The high labeling of Trp26 observed in the present work is encouraging and supports the NMR data. The fact that only one residue Val42 is strongly labeled in the β -sheet region (41–44, 47–50) of the native α -LA and that the labeling is poor compared to other residues suggests that this region is unlikely to persist in the MG state. Similar conclusions were reached by the NMR study (Chyan et al., 1993). To identify other sites in α -LA labeled by DAF, the labeled protein was subjected to CNBr fragmentation as the protein contains only one methionine residue at position 90. The CNBr fragmentation permitted isolation of the CN-1 and CN-2 fragment by SDS-PAGE and electroblotting on to PVDF membrane, which was subjected to peptide sequencing. The data obtained on sequencing CN-1 (1–90) for 29 cycles were similar to that obtained by the above-mentioned direct sequencing of the N-terminal of DAF-labeled LA (Fig. 2). The CN-2 fragment permitted sequencing from residue Cys91 to the end of the protein, i.e., residue 123. The major residues labeled in this case were Val92, Ile95, Lys98, Leu105, and Leu115 (Fig. 2). The first three of these residues belong to the C-helix in the native structure. The

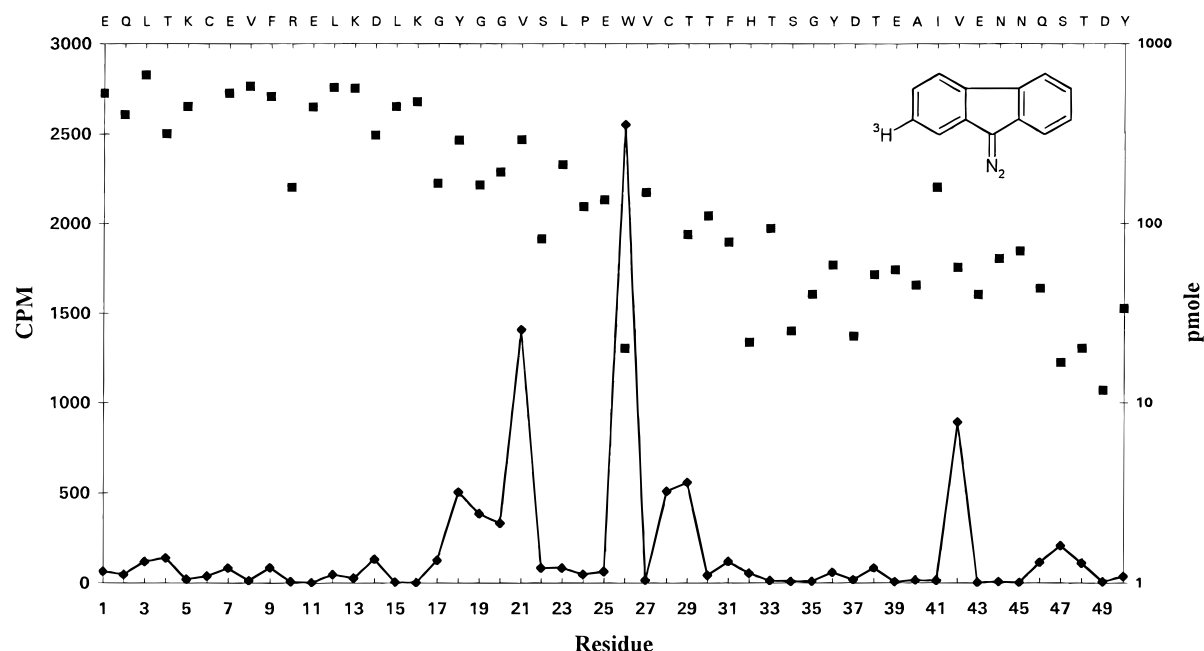


Fig. 1. Radioactivity and mass release on Edman degradation of DAF-labeled α -LA. DAF-labeled α -LA was subjected to Edman degradation on an automated gas-phase sequencer. After each Edman degradation cycle 60% of the sample was analyzed by direct injection on an on-line PTH analyzer for PTH-amino acid release (pmol, ■). The balance 40% of the sample was directed to a fraction collector and subsequently counted (cpm, ♦). A single peptide beginning with E-1 was detected. The sequence assignments could be made up to 45 cycles and are given on the top. The inset shows structure of 2-[^3H]DAF.

Trp104 present in the region between helix C (86–98) and D (105–110) helix, has its side-chain indole ring located in the hydrophobic environment provided by side chains of these two helices, is not labeled quite unlike the Trp26. Instead the neighboring

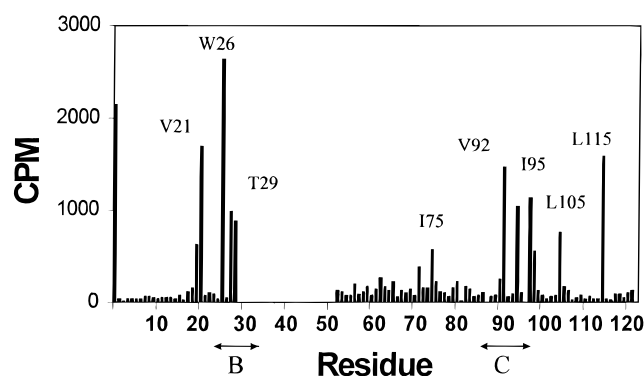


Fig. 2. Radioactivity release on Edman degradation of CNBr (CN1 1–89, CN2 90–123) and pepsin (53–123) fragments obtained from DAF-labeled α -LA. The PVDF blots for these fragments were subjected to Edman degradation. In the case of CN-1, fragment sequencing could be carried out for 29 cycles, beyond which the assignments of PTH-amino acid released were not reliable. However, in the case of CN-2 (90–123), sequencing could be carried out for 32 cycles and assignments made. In the case of the pepsin fragment, a single peptide beginning with Phe53 could be detected and sequenced for 38 cycles. After each Edman degradation cycle 60% of the sample was analyzed by direct injection on a on-line PTH analyzer for PTH-amino acid release. The balance 40% of the sample was directed to a fraction collector and subsequently counted. The position corresponding to B and C helix is marked.

Leu105 is significantly labeled. The significant labeling of Leu115 suggest that the C-terminal might retain partial native-like structure.

Polverino de Laureto et al. (1995) have recently reported use of pepsin and thermolysin to probe the MG state of α -LA. These results indicated that at pH 2, pepsin fairly selectively cleaved α -LA at Ala40–Ile41 and then at Leu52–Phe53. The rapid accessibility of these sites at pH 2 suggested that the β -sheet region (40–43, 47–50) in the native state is not maintained in the MG state. We took advantage of this selectivity to sequence balance (53–90) of the DAF-labeled LA. The pepsin treatment permitted isolation of the fragment beginning with Phe53. Sequencing of this peptide was possible up to Met90. It was observed that the overall labeling of this fragment was much lower compared to the N-terminal (1–52) and C-terminal (91–123) fragments described above. The major labeled residue was Ile75, with Ile72 showing relatively lower levels of labeling (Fig. 2).

The fact that DAF labels α -LA severalfold higher at pH 3 than pH 7 indicates the structural changes in the native state leading to formation of the MG state at low pH involve partial disruption of the densely packed protein interior, which permit DAF to partition and on photolysis label these hydrophobic sites (Lala & Kaul, 1992). An analysis of the DAF-labeled sites indicate that bulk of labeled residues correspond to the B- and C-helix (Fig. 3). These results suggest that the space between these two helices that is occupied by side chains of residues that are not part of the secondary structure elements in the native structure undergoes considerable tertiary structure change in the MG state, making room for DAF to bind and label residues in the B–C-helix interface. However, the secondary structure in the B–C-helix region must be retained to observe labeling of several residues in this region. Similar conclusions have been reached earlier by NMR studies

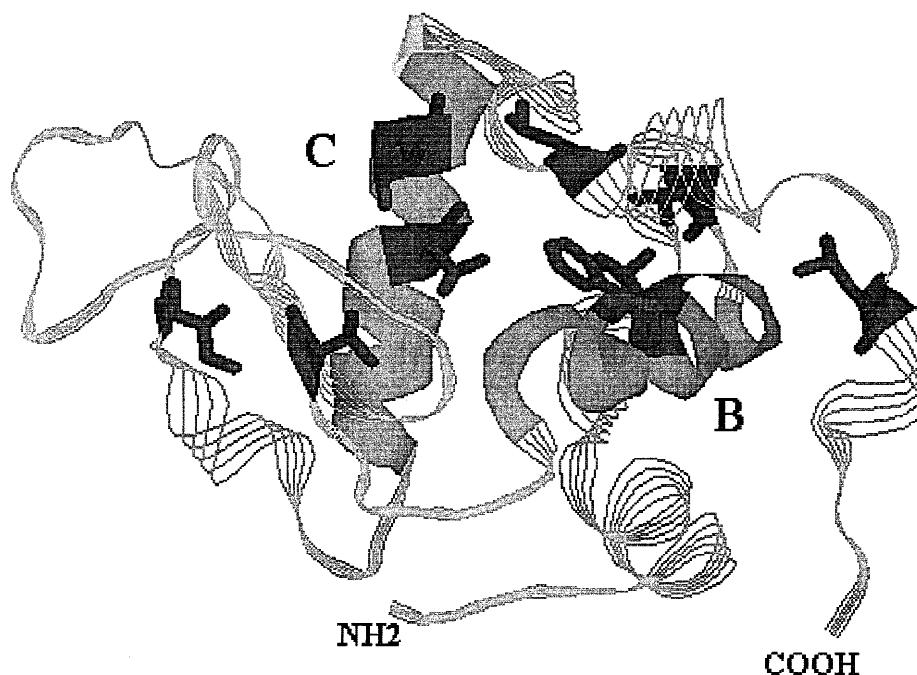


Fig. 3. The DAF-labeled residues in MG state of bovine α -LA. The residues labeled by DAF are highlighted in the structure of bovine α -LA (Pike et al., 1996). The coordinates for bovine α -LA (1hfz) were obtained from the Brookhaven Protein Data Bank. The bulk of the strongly labeled residues (21, 26, 29, 42, 75, 92, 95, 105, 115; side chains shown) belong to the B and C helix (shaded) region.

(Chyan et al., 1993; Schulman et al., 1995, 1997). Many of the residues labeled here have also been identified by H-D exchange NMR studies (Alexandrescu et al., 1993; Chyan et al., 1993), suggesting that the DAF-labeling method can be used to detect and structurally characterize the MG and MG-like equilibrium folding intermediates. However, unlike the NMR method, the photolabeling method is not limited by the size of the protein, and can provide information on several new residues, for example, Leu115 in α -LA, which may not have been detected by NMR due to limitation of spectral resolution.

Materials and methods

Bovine α -LA from Sigma was purified on a Sephadex G-100 column chromatography and was found to be homogeneous on analysis by SDS-PAGE. 2-[^3H]Diazofluorene (specific activity 491 mCi/mmol) was prepared from 2-[^3H]fluorenone according to the procedure reported earlier (Pradhan & Lala, 1987). Automated Edman degradation of photolabeled peptides was carried out using a Shimadzu gas-phase protein sequencer (PPSQ10) connected to an on-line PTH analyzer. The samples loaded as elctroblots on PVDF membrane.

Photolabeling of α -LA

Four hundred micrograms (28.2 nmol) of α -LA was taken in 5 mL of 50 mM glycine buffer containing 48 mM NaCl (pH 3) and an ethanolic solution of [^3H]DAF (70.5 nmol, 34.61 μCi) was added to it, the final ethanol concentration being kept less than 2% (v/v). The sample was incubated for 10 min and then photolyzed for 1 min in a Rayonet Mini-reactor fitted with four 3,500 Å lamps.

After photolysis, part of the sample was analyzed on a 15% SDS-PAGE gels, according to the procedure of Laemmli (1970). The gel was then stained in 0.1% Coomassie blue and destained in a de-staining solution containing 10% acetic acid and 40% methanol in water. The gel was then sliced into 2-mm pieces and digested in to 250 μL of 30% H_2O_2 at 60° for 8–10 h. The digested samples were counted in a toluene-based scintillation mixture. Another part (200 μg) of the sample was desalted on a Sep-Pak 1-mL C_{18} cartridge (Water Millipore, Bedford, Massachusetts). The column was thoroughly washed with 0.1% TFA in water and then eluted with 0.1% TFA in acetonitrile: water (80:20). The eluted fractions were checked for radioactivity and the radioactive fraction was lyophilized and then dissolved in SDS-PAGE sample buffer. The photolabeled α -LA was separated on a 16.5% Tris-tricine gel system according to Schagger and Von Jagow (1987) and immediately electroblotted on to a PVDF membrane (Immobolin-P^{SQ}, Millipore Corp.) according to Mozdzanowski and Speicher (1992) using 10 mM CAPS containing 20% methanol in transfer buffer. The PVDF membrane was stained for 0.5 min in 0.1% Coomassie blue and destained in 50% methanol/water. The radioactive α -LA band were excised and analyzed by sequencing.

Reduction, carboxymethylation, and CNBr fragmentation of DAF-labeled α -LA

Five hundred micrograms of photolabeled α -LA dissolved (2.5 mg/mL) in 0.1 M Tris, 6 M guanidine-HCl, pH 8.6 containing 0.18 M β -mercapto ethanol. The sample was then purged with argon for 2 min and incubated at 37°C for 12 h in a sealed test tube. Iodoacetamide prepared in 0.1 M Tris, 6 M guanidine-HCl, pH 8.6 was added with a final concentration of 1.5% with rapid mixing to the reduced α -LA and allowed to stand for 30 min. It

was then exhaustively dialyzed against 1% formic acid. The sample was concentrated to dryness. One hundred and forty-two micrograms of this material was taken in 70% formic acid, and CNBr was added at a 500:1 molar ratio of CNBr to methionine residue. The preparation was then digested for 48 h at 25°. The sample was then diluted 20 times by adding 1 to 2 mL of water and lyophilized. This process of hydration and lyophilization was repeated three times for the sample, then lyophilized to dryness. The concentrated samples were then analyzed on a 16.5% Tris-tricine gels and electroblotted on to a PVDF membrane as described above. The Coomassie blue staining pattern on PVDF was similar to that observed on SDS-PAGE gel staining. Two bands corresponding to M_r of CN-1 and CN-2 were clearly visible on the PVDF membrane. These bands were excised after destaining and subjected to automated Edman degradation.

Proteolytic fragmentation of DAF-labeled α -LA. One hundred and forty-two micrograms (10 nmol) of DAF-labeled α -LA in 150 μ L of 10 mM HCl containing 0.1 M NaCl (pH 2.0) and digested at 25° with pepsin at a protease:substrate ratio of 1:100 (w/w) according to the procedure of Polverino de Laureto et al. (1995). The proteolysis was stopped by adding 0.2% ammonia (v/v) solution (20 μ L). The sample was concentrated to dryness, dissolved in SDS-PAGE sample buffer, run on a 16.5% Tris-tricine gel, and electroblotted on to a PVDF membrane using 10 mM CAPS containing 30% methanol in transfer buffer. The band corresponding to residues 53–123 was clearly visible on the PVDF membrane and was subjected to automated Edman degradation.

Acknowledgments

The protein sequencing was carried out under a National facility program supported by Department of Science and Technology and Department of Biotechnology. This work was supported by a grant-in-aid from Department of Science and Technology, New Delhi.

References

- Alexandrescu AT, Evans PA, Pitkeathly M, Baum J, Dobson CM. 1993. Structure and dynamics of the acid-denatured molten globule state of α -lactalbumin: A two-dimensional NMR study. *Biochemistry* 32:1707–1718.
- Anjaneyulu PSR, Lala AK. 1984. On the importance of indiscriminate insertion for photoactivable reagents—Photolysis of diazofluorene in mixed organic solvents. *Indian J Chem* 23B:802–807.
- Anjaneyulu PSR, Pradhan D, Lala AK. 1984. Design and synthesis of new fluorescent photoaffinity labels to study membrane structure. *Proc Indian Acad Sci (Chem Sci)* 93:1229–1235.
- Baron WJ, DeCamp MR, Hendrick ME, Jones M Jr, Levin RH, Sohn MB. 1973. Carbenes form diazo compounds. In: Jones M Jr, Moss RA, eds. *Carbenes*, vol. 1. New York: John Wiley & Sons. pp 79–84.
- Chyan C-L, Wormald C, Dobson CM, Evans PA, Baum J. 1993. Structure and stability of the molten globule state of guinea-pig α -lactalbumin: A hydrogen exchange study. *Biochemistry* 32:5681–5691.
- Engelhard M, Evans PA. 1995. Kinetics of interaction of partially folded proteins with a hydrophobic dye: Evidence that molten globule character is maximal in early folding intermediates. *Protein Sci* 4:1553–1562.
- Grasse PB, Brauer BE, Zupancic JJ, Kaufmann KJ, Schuster GB. 1983. Chemical and physical properties of fluorenylidene: Equilibration of the singlet and triplet carbenes. *J Am Chem Soc* 105:6833–6845.
- Katsumata K, Okazaki A, Tsurupa GP, Kuwajima K. 1996. Dominant forces in the recognition of a transient folding intermediate of α -lactalbumin by GroEL. *J Mol Biol* 264:643–649.
- Kim PS, Baldwin RL. 1990. Intermediates in the folding reactions of small proteins. *Annu Rev Biochem* 59:631–660.
- Kuwajima K. 1996. The molten globule state of α -lactalbumin. *FASEB J* 10:102–109.
- Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685.
- Lala AK, Bhat S. 1990. Hydrophobic photolabeling in membranes: The human erythrocyte glucose transporter. *Biotechnol Appl Biochem* 12:586–594.
- Lala AK, Kaul P. 1992. Increased exposure of hydrophobic surface in molten globule state of α -lactalbumin. Fluorescence and hydrophobic photolabeling studies. *J Biol Chem* 267:19914–19918.
- Lindner RA, Kapur A, Carver JA. 1997. The interaction of the molecular chaperone, α -crystallin, with molten globule states of bovine α -lactalbumin. *J Biol Chem* 272:27722–27729.
- Martin J, Langer T, Boteva R, Schramel A, Horwich AL, Hartl FU. 1991. Chaperonin-mediated protein folding at the surface of groEL through a “molten globule”-like intermediate. *Nature* 352:36–42.
- Mozdzanowski J, Speicher DW. 1992. Microsequence analysis of electroblotted proteins. I. Comparison of electroblotting recoveries using different types of PVDF membranes. *Anal Biochem* 207:11–18.
- Ohgushi M, Wada A. 1983. Molten-globule state: A compact form of globular proteins with mobile side-chains. *FEBS Lett* 164:21–24.
- Pike AC, Brew K, Acharya KR. 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.
- Polverino de Laureto P, De Filippis V, Di Bello M, Zamboni M, Fontana A. 1995. Probing the molten globule state of α -lactalbumin by limited proteolysis. *Biochemistry* 34:12596–12604.
- Pradhan D, Lala AK. 1987. Photochemical labeling of membrane hydrophobic core of human erythrocytes using a new photoactivable reagent 2[³H]diazofluorene. *J Biol Chem* 262:8242–8251.
- Puitsyn OB. 1995. Molten globule and protein folding. *Adv Protein Chem* 47:83–229.
- Ruzicka J, Leyva E, Platz MS. 1992. Laser flash photolysis of 9-diazofluorene in low-temperature glasses. *J Am Chem Soc* 114:897–905.
- Schagger H, Von Jagow G. 1987. Tris/tricine buffer system for separation of small polypeptides. *Anal Biochem* 166:368–379.
- Schulman BA, Kim PS, Dobson CM, Redfield C. 1997. A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat Struct Biol* 4:630–634.
- Schulman BA, Redfield C, Peng ZY, Dobson CM, Kim PS. 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.
- Semisotnov GV, Rodionova NA, Razgulyaev OI, Uversky VN, Gripas AF, Gilmanshin RI. 1991. Study of the “molten globule” intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* 31:119–128.