

Near Ultraviolet Absorption Arising from Lysine Residues in Close Proximity: A Probe to Monitor Protein Unfolding and Aggregation in Lysine-Rich Proteins

Lopamudra Homchaudhuri and Rajaram Swaminathan*

Departments of Chemistry and Biotechnology, Indian Institute of Technology Guwahati, Guwahati 781 039, Assam, India

Received June 2, 2003; E-mail: rsw@iitg.ernet.in

There is a need for an intrinsic spectral probe to monitor key events like protein unfolding and aggregation in a rapid and unambiguous manner. Protein aggregation is an important issue, but ironically there is a dearth of simple techniques to directly detect the presence of aggregates in solution. We report here the hitherto undiscovered electronic absorption around 300–350 nm in aqueous solutions (pH 7) of human serum albumin (HA), calf thymus histone, and poly-L-lysine. The above spectra were significantly absent in controls like subtilisin carlsberg, mutant barstar, and lysozyme. The possibilities that Rayleigh scattering or impurities could account for the above spectra were checked and ruled out. Based on the analysis of available three-dimensional structures from PDB and our earlier work on the lysine amino acid, an intramolecular interaction between lysine side chains in close spatial proximity was deduced to be the origin for the above spectra. The utility of Lys–Lys interaction in detecting protein unfolding and aggregation in a lysine-rich protein like calf thymus histone using near-ultraviolet absorption is demonstrated.

It is well known that the side chains of Tyr and Trp play a dominant role in near-ultraviolet electronic absorption and subsequent fluorescence among proteins.[1,2] The intrinsic fluorescence of the Trp residue has been extensively employed to probe the local structure, dynamics, and polarity surrounding the indole moiety in a variety of single and multi-tryptophan containing proteins.[3,4] However, there is a need for alternate intrinsic probes, especially to track important events like protein unfolding and aggregation. It is desirable that these probes are a) sensitive to global rather than local changes (like Trp) of conformation, and b) workable at higher protein concentrations (like those encountered in aggregation), when signals from other probes (like Trp) have attained saturation.

There have been earlier attempts to introduce intrinsic probes biosynthetically by employing tryptophan analogues like 7-azatryptophan[5–7] and 5-hydroxytryptophan.[5,8] However, these analogues have met with limited applications due to the tedious procedures involved in incorporating them into proteins. Proteins that are intrinsically fluorescent, e.g., green fluorescent protein,[9] red fluorescent protein[10,11] etc., have been utilized to covalently tag numerous target proteins in the past few years.[12] However, the need for a natural intrinsic spectroscopic probe requiring no biosynthetic incorporation or covalent tagging, which can sense global changes in protein structure, still remains.

Earlier work by us had shown that L-lysine displays new-near ultraviolet (~270 nm) absorption at high concentrations (~0.5 M) in aqueous medium at pH 7.[13] The new spectra were deduced to arise from *intermolecular* interactions between adjacent lysine side chains in the aggregates. In this paper, similar observations arising from *intramolecular* interactions between two or more lysine residues present in close spatial vicinity in human serum albumin (HA),[14] poly-L-lysine and calf thymus histone are presented. Consistent with the intramolecular nature of the interactions, these spectroscopic features were observed at dilute concentrations (~100 μM for HA). Also, as predicted from the earlier study, the lysine side chain and not the terminal amino or carboxyl groups, play a major role in the above interaction. Importantly, the present investigation showed that the above spectral features are dependent on the proximity of the lysine side chains in the protein. This property was subsequently exploited to observe the unfolding and aggregation of calf thymus histone from the absorption spectra. To the best of our knowledge, there has been no earlier report of these novel spectra.

Experimental

Human serum albumin (99% purity, fatty acid and globulin free, product # A 3782), chicken eggwhite lysozyme, subtilisin carlsberg, calf thymus histone, guanidine hydrochloride (GdnCl),[14] poly-L-lysine hydrobromide (Mol. wt. 1000–4000) and L-lysine hydrochloride of the highest available purity were purchased from Sigma Aldrich Foreign Holding Company, New Delhi. The absorption ratio A_{300}/A_{340}, observed by us for HA (3.11) is in good agreement with the value (~3.0) obtained from the spectrum reported in the literature.[15] The W38FW44F mutant of barstar, henceforth referred to as barstar mutant that contains a single Trp (Trp53) was employed. It was isolated from a bacterial overexpression system described elsewhere.[16] Wild-type barstar was not employed, since it has three tryptophans. The presence of multiple tryptophan residues can under favorable conditions lead to unwanted interference from Trp–Trp π–π interactions. All other chemicals employed were of analytical grade.

Absorption measurements were performed on a double beam

Published on the web April 2, 2004; DOI 10.1246/bcsj.77.765

Hitachi spectrophotometer (Model U2001). The displayed spectra are averages of at least ten scans.

Protein concentrations were confirmed by measuring the absorbance at 280 nm and relating the measured absorbance values with the calculated molar extinction coefficients.[17] The distance between each pair of lysine residues were calculated using the three-dimensional structures procured from PDB and RasMol software (Version 2.6). As the protein crystals are well hydrated, structures determined from crystals are not much different from the structures of soluble proteins in aqueous solutions. For calculation, the distance separating the N atoms of the ε-amino group of each lysine in the pair was taken. The choice of 7.2 Å as cutoff for proximity is based on the distance constraint for hydrogen bonding between two donor N atoms and one acceptor O atom for both donors (see under 'Results and Discussion'). Samples were made in deionized water. All experiments were carried out at 22–25 °C. Experiments in the presence of GdnCl or urea were carried out after overnight incubation at room temperature in the presence of the denaturant. Other relevant experimental conditions are given in the figure legends.

Results and Discussion

Previous studies have shown that L-lysine has a distinct absorption around 270 nm at high concentrations,[13] as reproduced here in Fig. 1, curve b. If this absorption does arise from intermolecular interactions among lysine side chains present in close proximity among aggregates, then it is likely to be observed in a protein where intramolecular interactions between lysine side chains are possible. We proceeded to investigate this.

Figure 1 shows the absorption spectrum of HA (curve a) and lysozyme (curve e). The tyrosine and more dominant tryptophan absorption bands in these proteins mask any features that may exist due to interactions among lysine side chain(s) between 270 and 310 nm. On examining wavelengths beyond 315 nm, one sees that a broad tail of absorption exists for HA. The molar extinction coefficient (ε) was $\sim$1546 M^{-1} cm^{-1} at 325 nm for HA. A similar tail of absorption was evident for HA in a previous report also.[15] Lysozyme unlike HA, shows a steep drop in absorbance beginning from 304 nm with little absorption, if any, beyond 315 nm.

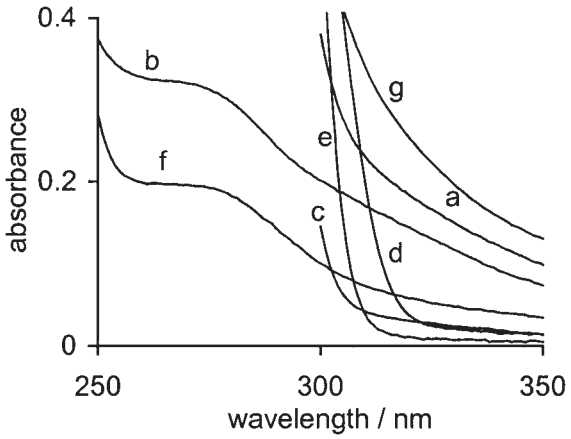

Fig. 1. Absorption: The absorption spectrum of (a) human serum albumin [1.08×10^{-4} M]; (b) L-lysine [1.0 M]; (c) subtilisin carlsberg [1.5×10^{-4} M]; (d) mutant barstar [1.5×10^{-4} M]; (e) lysozyme [1.5×10^{-4} M]; (f) poly-L-lysine hydrobromide [10 mg/mL]; and (g) calf thymus histone [11 mg/mL]. All proteins were dissolved in 0.05 M NaH_2PO_4 adjusted to pH 7.0.

It is worth mentioning that the molar extinction coefficients of tryptophan, tyrosine, phenylalanine, cysteine, and cystine are negligibly small beyond 315 nm at neutral pH.[18] Therefore the absorbance observed beyond 315 nm in HA cannot be attributed to any of the above amino acid side chains. Apart from the above, no significant electronic absorption is known to exist for the rest of the amino acid side chains at wavelengths longer than 315 nm.[18]

To verify the origin of the absorption beyond 315 nm, we performed experiments with poly-L-lysine as depicted in Fig. 1, curve f. Poly-L-lysine shows an absorption spectrum similar to that of lysine amino acid (curve b), but at a concentration that is 250 fold less than that of lysine amino acid, since the interactions between the lysine residues are intramolecular in this case.

In order to compare the absorption of different proteins in a concentration independent manner, we have tabulated the ratio A_{350}/A_{325} for the different proteins in Table 1. The wavelengths 325 and 350 nm were chosen so as to remain clearly away from tryptophan absorption. The table shows that L-lysine, HA, and histone have a similar ratio ($\sim$0.6). This indicates a similar shape in their absorption spectra between 325 and 350 nm. Poly-L-lysine (0.66) shows a ratio that is not too far from 0.6 either. Thus, overall, the spectra for the above proteins are fairly similar in shape to that of L-lysine, hinting at a common origin.

Rayleigh scattering of suspended particles like undissolved or precipitated proteins can also be mistaken for absorption. Figure 2 shows the normalized spectra of HA, calf thymus histone, L-lysine, and poly-L-lysine under identical conditions. These spectra are nearly superimposable on one another after 325 nm consistent with their common origin. The absorption that shall result owing to purely Rayleigh scattering is shown in Fig. 2 for two cases: (a) by simulation using a λ^{-4} dependence (thick line), (b) using a dilute aqueous solution of magnesium hydroxide (dashed line). It is evident that in both cases the scatter spectra are clearly below the spectra observed for proteins. In fact, the simulated curve approaches close to the protein spectra only if a λ^{-7} dependence is employed. Hence, we can conclude that the spectra observed among lysine-rich proteins do not arise from Rayleigh scattering.

It is unlikely that the tailing of absorbance up to 350 nm arises from Trp–Trp π–π stacking interactions since HA has only one tryptophan residue in the entire protein. No Trp is present

Table 1. Absorbance Ratio (350 nm/325 nm)

Sample	A_{350}/A_{325}
L-Lysine	0.57
HA	0.59
Histone	0.59
Poly-L-lysine	0.66

The absorbance ratio for lysozyme, mutant barstar, and subtilisin carlsberg are not shown as their absorbance at 350 nm are too less to be meaningfully interpreted.

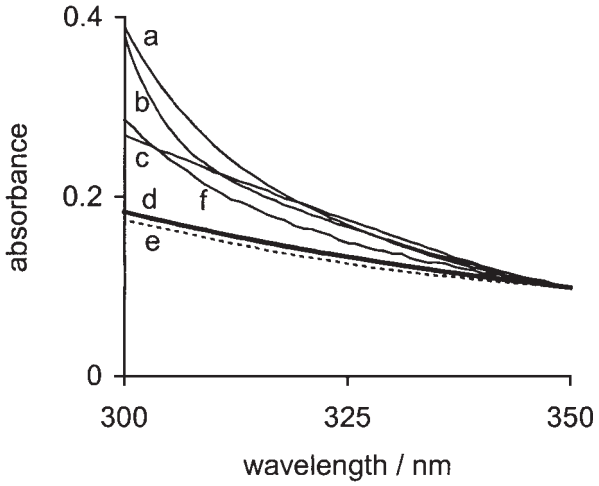

Fig. 2. Can it be due to Rayleigh scattering: The absorption spectra are as follows: (a) calf thymus histone; (b) human serum albumin; (c) L-lysine; (d) simulated scatter (thick line); (e) a dilute scattering solution of $Mg(OH)_2$ (dashed line); and (f) poly-L-lysine. For the sake of comparison, all absorption spectra except human serum albumin, were normalised to have identical absorption at 350 nm.

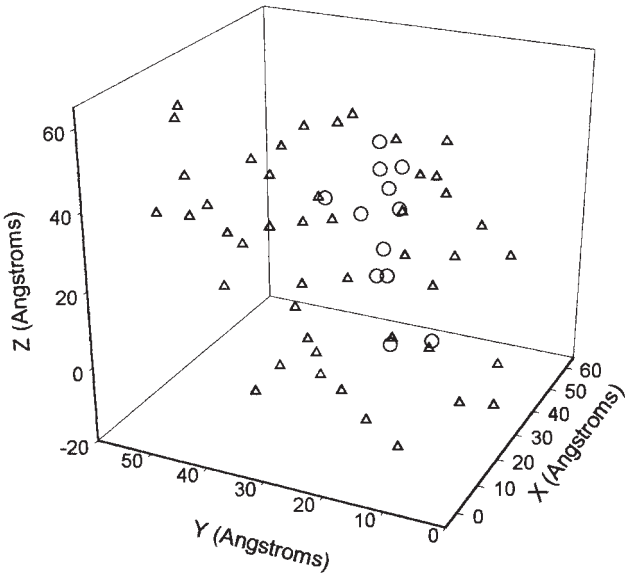

Fig. 3. Lysine clusters in human albumin: The three dimensional spatial distribution of ε-amino N atoms in HA are shown employing the crystal structure from PDB (1AO6, Chain A). The atoms that are closer to one another by less than 7.2 Å in space are indicated as open circles, while the rest are shown as open triangles.

Table 2. Spatial Distribution of Lysine Residues in the Proteins Studied

Protein	PDB code	Number of lysine's in the protein sequence	Number of lysine pairs separated by less than 7.2 Å
Lysozyme	1E8L	6	1
Barstar	1BTA	6	0
Subtilisin carlsberg	1SCA	9	0
Human serum albumin	1AO6	59	7

Please refer 'Experimental Section' for more details.

in poly-L-lysine.

It is apparent from Fig. 1 that HA displays a more intense absorption compared to lysozyme, mutant barstar and subtilisin carlsberg. It is now imperative to look at some of the facts from the three-dimensional structure of these proteins in their native state as available from PDB. From Table 2 it is evident that both subtilisin carlsberg and barstar mutant possess no proximal lysine pairs. Lysozyme has only one pair of lysine residues in close proximity (less than 7.2 Å). If the observed spectra do originate from intramolecular lysine side chain interactions, it would explain why lysozyme, mutant barstar and subtilisin carlsberg showed negligible absorption. HA which showed a comparatively higher absorption, has a relatively large number of proximal lysine pairs (~7). Figure 3 depicts the lysine pairs that exist within 7.2 Å (open circles) in HA. A couple of lysine clusters are clearly evident. However, it must be borne in mind that distances calculated from the three-dimensional structure are average indicators of the true picture in solution at any given time. In spite of this fact, based on the correlations described above, it appears likely that the observed absorption spectra are directly linked to the number of proximal lysine residues in the protein.

The large value of ε_{325} in HA (1546 $M^{-1}\,cm^{-1}$) can be attributed to the folded globular protein architecture enabling a multitude of possible intramolecular interactions between the 59 lysine residues. In contrast, poly-L-lysine is known to possess a random coil structure near neutral pH and 25 °C.[19]

To verify our conclusion, we performed two experiments with calf thymus histone where lysine residues abound. In the first experiment, the absorption spectrum of histone was recorded at identical concentrations in the presence and absence of 6 M GdnCl. The result is shown in Fig. 4. A clear decrease in the absorbance (~4 fold drop at 325 nm) is observed on unfolding suggesting that the population of spatially proximal lys residues is significantly reduced in the presence of 6 M GdnCl. A similar drop in absorbance was also observed with unfolding in the presence of 8 M urea (data not shown).

In the other experiment, we monitored the absorbance of histone as its concentration in the aqueous buffer was increased from 0.7 mg/mL to 10 mg/mL. The absorbance at 325 nm was plotted against the protein concentration, as shown in Fig. 5. A wavelength of 325 nm was chosen so as to remain clearly away from tryptophan absorption. A distinct break in the linearity is observed. The biphasic transition clearly indicates the formation of histone aggregates. As revealed from the slopes, this transition to a more aggregated protein occurs

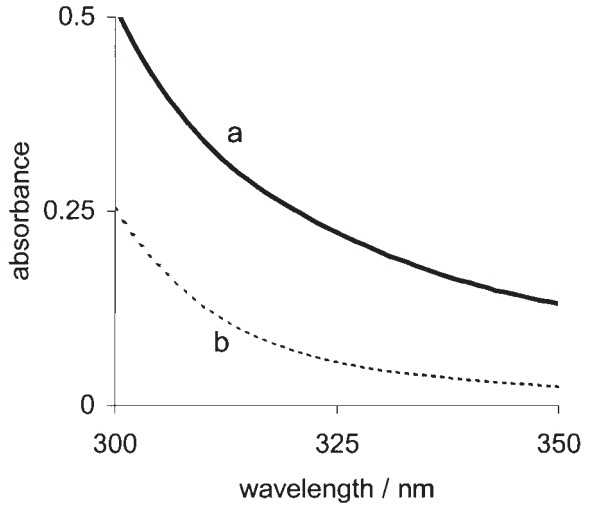

Fig. 4. Monitoring protein unfolding: The absorption spectrum of calf thymus histone [11 mg/mL] is shown at pH 7 in the absence (curve 'a') and presence (curve 'b') of 6 M GdnCl. Other conditions are similar to Fig. 1.

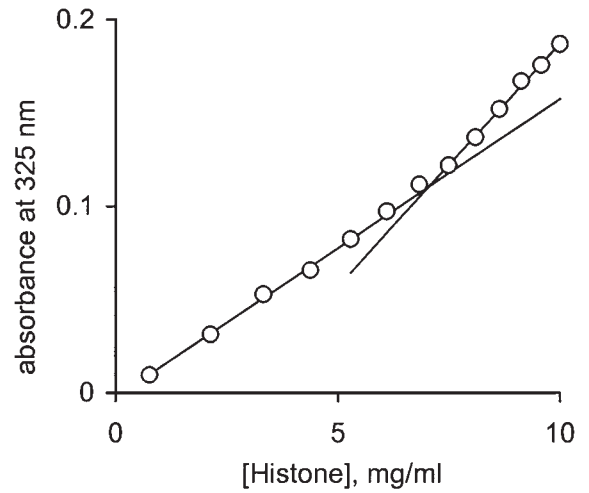

Fig. 5. Monitoring protein aggregation: The absorbance of calf thymus histone at 325 nm is plotted against its concentration. Five consecutive points from either end were chosen to obtain the fitted straight lines. The lines intersect at a concentration of 7.1 mg/mL. Other conditions are similar to Fig. 1.

at a concentration of 7.1 mg/mL. It is evident that, after aggregation, the Lys–Lys absorption shows a more rapid increase with concentration (shown by a higher slope), due to a increase in the population of Lys–Lys pairs post aggregation.

The results in Fig. 4 clearly argue against the role of impurities in producing the observed lysine spectra, as it is unlikely that addition of GdnCl or urea at constant pH will influence the extinction coefficient of the impurity. A nonlinear dependence on concentration (Fig. 5) is also difficult to account for by an impurity. In summary, these experiments demonstrate a simple and economical method to monitor protein aggregation and unfolding in lysine-rich proteins.

We now come to the nature of forces that can support Lys–Lys interactions in the protein. One hypothesis is the presence of a bridging water molecule between each of the >N–H groups of the Lys–Lys pair. The two lone pairs in the O atom of the bridging water molecule can each act as hydrogen bond acceptor for one H atom each from each of the two nitrogens. Such an arrangement: a) is not uncommon in proteins;[20] b) requires the distance between the two N atoms to be around ~7 Å; c) is not affected if the ε-amino group exists as $-NH_3^+$ since the N lone pair is not involved. In fact, our data with lysine amino acid indicate only a minor change in the absorption in the pH range from 2 to 11.[21] More work is required to prove this interaction.

The major question that remains to be answered is the nature of the chromophore involved. Here one should avoid speculations; it is imperative that we perform quantum chemical calculations a) to support possible structures and b) to determine the oscillator strength for the transition theoretically. These are in progress. It must be mentioned that a faint blue luminescence (quantum yield < 0.01) was also observed from the proposed lys bands.

Taken together, the results presented here suggest that the newly observed near-ultraviolet absorptions beyond the Trp bands owe their origin to nearby lysine side chains in proteins. These new spectral features were utilized to 1. monitor protein aggregation, when aggregation brings a bunch of lysine residues in separate protein molecules in close proximity; 2. observe protein unfolding, when unfolding leads to a decrease in the population of proximal lysine pairs. We hope that it may also find application in tracking a lysine-rich protein during its purification or isolation.

RS wishes to thank Prof. J. B. Udgaonkar for allowing him to purify barstar mutant. Financial support from the Department of Science and Technology, New Delhi is acknowledged.

References

1 J. M. Beechem and L. Brand, *Annu. Rev. Biochem.*, **54**, 43 (1985).

2 M. R. Eftink, "Topics in Fluorescence Spectroscopy—Vol. 6: Protein Fluorescence," ed by J. R. Lakowicz, Plenum Pub. Corp, New York (2001), pp. 1–15.

3 J. R. Lakowicz, "Principles of Fluorescence Spectroscopy," Kluwer Academic/Plenum, New York (1999).

4 R. Swaminathan, G. Krishnamoorthy, and N. Periasamy, *Biophys. J.*, **67**, 2013 (1994).

5 C. Y. Wong and M. R. Eftink, *Protein Sci.*, **6**, 689 (1997).

6 J. Guharay and P. K. Sengupta, *Biochem. Biophys. Res. Commun.*, **219**, 388 (1996).

7 P. Soumillion, L. Jespers, J. Vervoort, and J. Fastrez, *Protein Eng.*, **8**, 451 (1995).

8 T. M. Laue, D. F. Senear, S. Eaton, and J. B. A. Ross, *Biochemistry*, **32**, 2469 (1993).

9 M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, *Science*, **263**, 802 (1994).

10 G. S. Baird, D. A. Zacharias, L. A. Gross, R. C. Hoffman, K. K. Baldridge, and R. Y. Tsien, *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11984 (2000).

11 L. A. Gross, L. A. Baird, R. C. Hoffman, K. K. Baldridge, and R. Y. Tsien, *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11990 (2000).

12 R. Y. Tsien, *Annu. Rev. Biochem.*, **67**, 509 (1998).

13 L. Homchaudhuri and R. Swaminathan, *Chem. Lett.*, **2001**, 844.

14 Abbreviations: HA, human serum albumin, GdnCl, guanidine hydrochloride.

15 N. Hagag, E. R. Birnbaum, and D. W. Darnall, *Biochemistry*, **22**, 2420 (1983).

16 R. Swaminathan, U. Nath, J. B. Udgaonkar, N. Periasamy, and G. Krishnamoorthy, *Biochemistry*, **35**, 9150 (1996).

17 C. R. Cantor and P. R. Schimmel, "Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function," W. H. Freeman and Company, New York (1980).

18 "Practical Handbook of Biochemistry and Molecular Biology," ed by G. D. Fasman, CRC Press, London (1992), pp. 79–83.

19 K. Rosenheck and P. Doty, *Proc. Natl. Acad. Sci. U.S.A.*, **47**, 1775 (1961).

20 G. A. Arteca and X. Luo, *J. Mol. Struct.: THEOCHEM*, **501–502**, 479 (2000).

21 L. Homchaudhuri and R. Swaminathan, unpublished results.